ECMWF COPERNICUS REPORT

Copernicus Atmosphere Monitoring Service

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

Period December 2020 – February 2021

Issued by: KNMI

Date: 22 June 2021 final

Ref: CAMS84_2018SC3_D1.1.1_DJF2021

IMPLEMENTED BY ECMWF

Validation report of the CAMS near-real-time global atmospheric composition service: Period December 2020 – February 2021

EDITORS:
M. Schulz (MetNo), Q. Errera (BIRA-IASB), M. Ramonet (LSCE),
N. Sudarchikova (MPG), H.J. Eskes (KNMI)

AUTHORS:
S. Basart (BSC), A. Benedictow (MetNo), Y. Bennouna (CNRS-LA),
A.-M. Blechschmidt (IUP-UB), S. Chabrillat (BIRA-IASB),
Y. Christophe (BIRA-IASB), E. Cuevas (AEMET), A. El-Yazidi (LSCE),
H. Flentje (DWD), P. Fritzsche (DWD), K. M. Hansen (AU), U. Im (AU),
J. Kapsomenakis (AA), B. Langerock (BIRA-IASB),
A. Richter (IUP-UB), V. Thouret (CNRS-LA), A. Wagner (MPG),
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:
Schulz, M., Q. Errera, M. Ramonet, Sudarchikova, N., H. J. Eskes, S. Basart, A. Benedictow, Y. Bennouna, A.-M. Blechschmidt, S. Chabrillat, Christophe, Y., E. Cuevas, A. El-Yazidi, H. Flentje, P. Fritzsche, K.M. Hansen, U. Im, J. Kapsomenakis, B. Langerock, A. Richter, V. Thouret, A. Wagner, T. Warneke, C. Zerefos, Validation report of the CAMS near-real-time global atmospheric composition service: Period December 2020 – February 2021, Copernicus Atmosphere Monitoring Service (CAMS) report, CAMS84_2018SC3_D1.1.1_DJF2021.pdf, June 2021, doi:10.24380/f540-kb09.

STATUS:
Version 1.0, final

DATE:
22 June 2021

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 March 2021, with a focus on December 2020 - February 2021 (DJF-2021). Updates of this document appear every 3 months (Sudarchikova et al., 2021). A detailed description of the measurement datasets used is provided in Eskes et al. (2021). Automated verification plots are made available through the CAMS global evaluation server, https://global-evaluation.atmosphere.copernicus.eu.

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 configurations are 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 atmospheric composition data assimilation. Also, the pre-operational delayed-mode analyses and high-resolution forecasts of CO_2 and CH_4 are assessed in this report.

Air quality and atmospheric composition

Tropospheric ozone (O_3)

The CAMS o-suite ozone is validated with surface and free tropospheric ozone observations from the GAW and ESRL networks, ozone sondes, IAGOS aircraft profiles and IASI tropospheric ozone retrievals. 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 ±20% for the Arctic in the recent years (Fig. S.1). Over Antarctica the o-suite biases are observed within ±20% for the recent years, whereas the control run shows larger negative biases. For DJF 2021, good agreement is found for both runs over the NH mid latitudes, Arctic (MNMBs are within 10%) and Antarctica (MNMBs are within ± 5%) in the free troposphere.

The time series show no indication of substantial changes in tropospheric ozone biases after the upgrade to Cy47R1 on October 6, 2020.

Due to the persisting restrictions associated to the COVID-19 crisis, IAGOS operations remain impacted this period although more observations are available compared to other quarters of 2020. Ozone is well represented in the low troposphere by both runs with slight overestimations. On average, a positive MNMB of less than 15% is found up to 2000 m and a correlation of more than 80%. In the free troposphere, the MNMB is about 10%, however the correlation is much

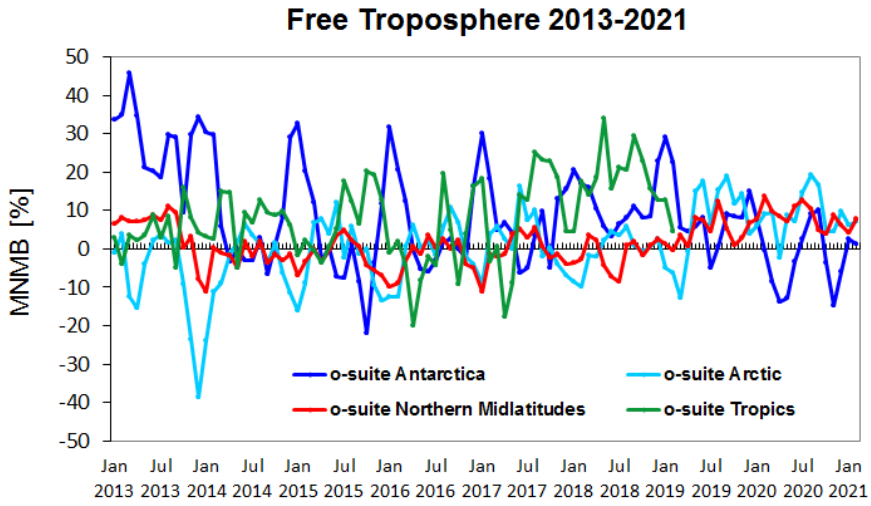


Figure S.1: Time series of MNMB of ozone in the o-suite, compared against ozone sondes, averaged over different latitude bands, period January 2013 to February 2021. The free troposphere is defined here as the layer between 750 and 300 hPa.

lower than in the lowest layers. In the UTLS region the bias is larger, and ozone is mostly overestimated by the o-suite with an MNMB of about 25%.

The comparisons of IAGOS cruise data show that the o-suite overestimates ozone with a relative bias in the range 50-100% over all regions. Like for the o-suite, an overestimation is also found for the control run but often with larger bias than the o-suite as well as more pronounced differences between the regions of the world.

The comparison with GAW surface stations shows that O_3 surface mixing ratios in DJF are overestimated with MNMBs within 10% for Europe (except for Hohenpeissenberg station with bias of about 30%) and within 15% for Asian stations. Correlation coefficients for European stations are between 0.55 and 0.81 and between 0.1 and 0.95 for Asian stations. The validation with ESRL station observations likewise shows overestimations of ozone mixing ratios over the Artic (order 20%), USA (between 0 and 30%) and Tropical stations (order 10%), close to zero bias over New Zealand and underestimations of ozone mixing ratios over Antarctica (range -37% to -5%). Correlation coefficients for Arctic and tropical stations are higher than 0.8 (except Point Barrow station), between 0.42 and 0.78 for the USA stations and 0.75 at Lauder station in New Zealand.

The validation with IASI satellite data shows that o-suite run is in good agreement with observations with a difference within ±10%. The model run captured well record-low ozone values over the Arctic in March and over Antarctic during September-November 2020 (higher bias up to 20% can be seen in October). The control run shows relatively good results with a bias within 10% from autumn 2020 onwards, linked to the update of the modelling of stratospheric ozone in 47R1. An overestimation of about 20% is observed in the control run over the southern high latitudes during October-December 2020 (Figure S.2) which may be related to the very stable vortex, strong ozone depletion and long duration of the ozone hole season in 2020, which deviates from the climatology in the hybrid linear stratospheric ozone scheme and is corrected in the assimilation (o-suite).

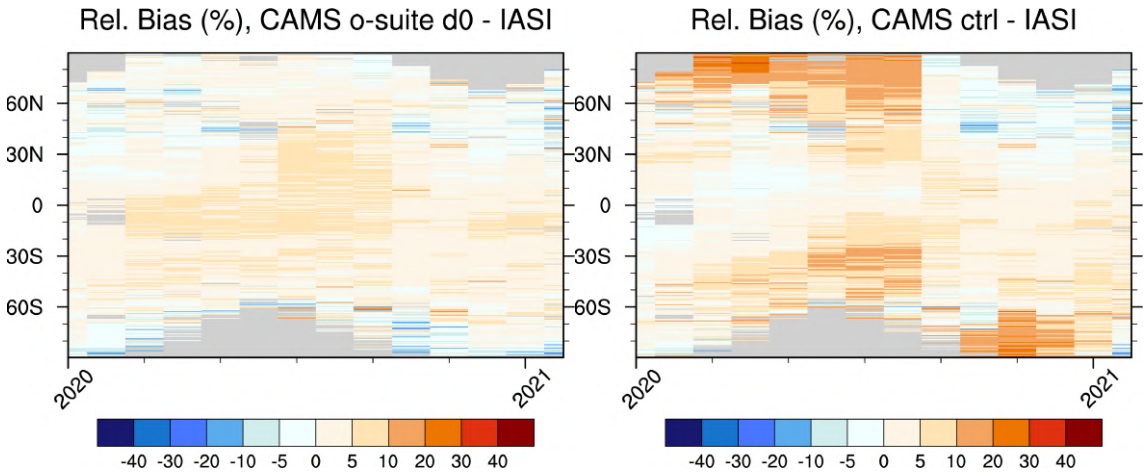


Figure S.2: Relative difference between the model runs and IASI Metop-B O_3 total column as function of latitude and time from January 2020 to February 2021: o-suite analysis (left) and control run (right). Grey colour indicates missing values. The upgrade of 6 October results in a clear improvement of stratospheric ozone in the model, as demonstrated by the control run plot (right).

The validation with IASOA surface observations shows that the CAMS simulations do not capture ozone depletion events in March – June in 2015 – 2020 at any of the sites. These events are related to halogen chemistry reactions that are not represented in the CAMS simulations. The simulations are on average in good agreement with the observations apart from the spring depletion events. For the period December 2020 – February 2021 the measurements are not quality controlled. The CAMS simulations captured the surface O_3 levels and variability very well at Svalbard and VRS with low biases (around zero) and high correlations (0.83 and 0.78).

The comparison with surface ozone observations in megacities in China show significant correlations with $0.3 < r < 0.7$. The o-suite mostly underestimates surface ozone values in Chinese cities.

Tropospheric Nitrogen dioxide (NO_2)

Model validation with respect to Sentinel-5P TROPOMI and GOME-2C NO_2 data shows that tropospheric NO_2 column distributions are well reproduced by the NRT CAMS runs, indicating that emission patterns and NOx photochemistry are generally well represented. However, the model runs are positively biased over anthropogenic emission hotspots, the seasonality is overestimated over East-Asia and Eastern-US and modelled shipping signals are more pronounced than in the satellite retrievals (Figure S.3). With respect to months affected by COVID-19 lockdown time periods, the o-suite generally fails to reproduce observed reductions in tropospheric NO_2 for large areas over China and South Asia but performs a bit better for Europe.

The comparison with surface NO_2 observations in megacities in China shows that the o-suite strongly overestimates surface NO_2 with MNMBs between 20% and 75% for latitudes between 30°N-40°N. Significant correlations are found, in the range $0.32 < r < 0.73$. The NO_2 overestimation from the CAMS NRT runs may explain the O_3 underestimation in Chinese megacities.

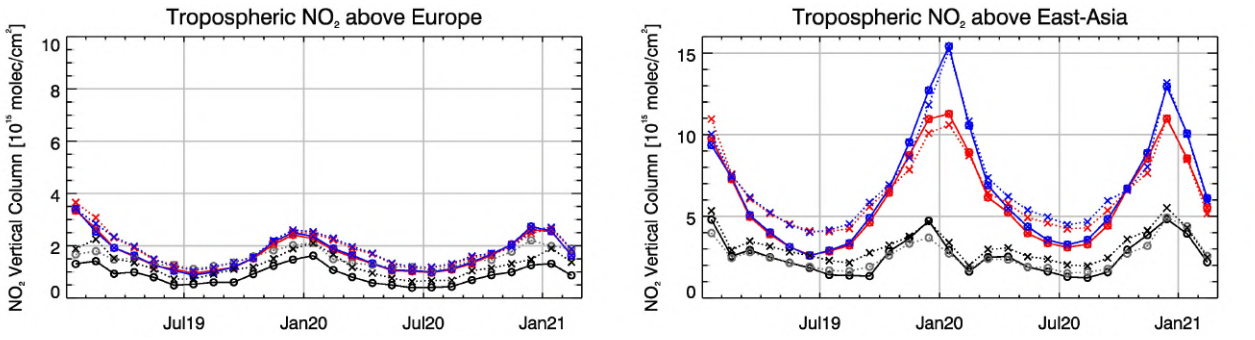


Figure S.3: Time series of average tropospheric NO_2 columns [10^{15} molec cm^{-2}] from (black and grey) satellite retrievals, (blue) control and (red) o-suite between January 2019 and February 2021. The lines with circles show results for TROPOMI (in black the IUP-Bremen product and in grey the operational offline product), the dotted lines with crosses show comparisons for GOME-2C.

Carbon Monoxide (CO)

Model validation with respect to GAW network surface observations, 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. Compared to GAW stations in DJF-2021, the o-suite run is well reproducing the observations for European stations (biases between 0% and -12%), for Asian stations and for the station located in the Southern Hemisphere (bias close to zero). The control run shows larger negative biases for European stations (up to -22%) and for Asian stations (up to -32%). The comparisons with EEA AirBase surface observations in Europe shows high temporal correlations, small biases over Spain, Belgium, Germany, Austria, Switzerland and Cyprus and larger negative biases over Estonia (-35%), the Czech Republic (-40%), Poland (down to -60%) and Bulgaria (-95%).

The model upgrade (46R1, 60 to 137 levels) implemented in July 2019 changes the overall biases in both the troposphere and stratosphere compared to NDACC FTIR observations (Figure S.4). The negative bias for the tropospheric columns increased from -2% before July 2019 to -7% in the most recent quarter and is larger than the reported 3% measurement uncertainty. The stratospheric column bias also changed to -7% in DJF 2021 compared to values well above +10% before July 2019. The 6 October 2020 upgrade (47R1) shows similar biases to the period before, while the 46R1 configuration shows significantly lower CO in both the troposphere and stratosphere.

The TCCON data in DJF were available only for one site, Nicosia. The o-suite shows good agreement with the measurements. The control run underestimates CO values and does not capture the seasonality.

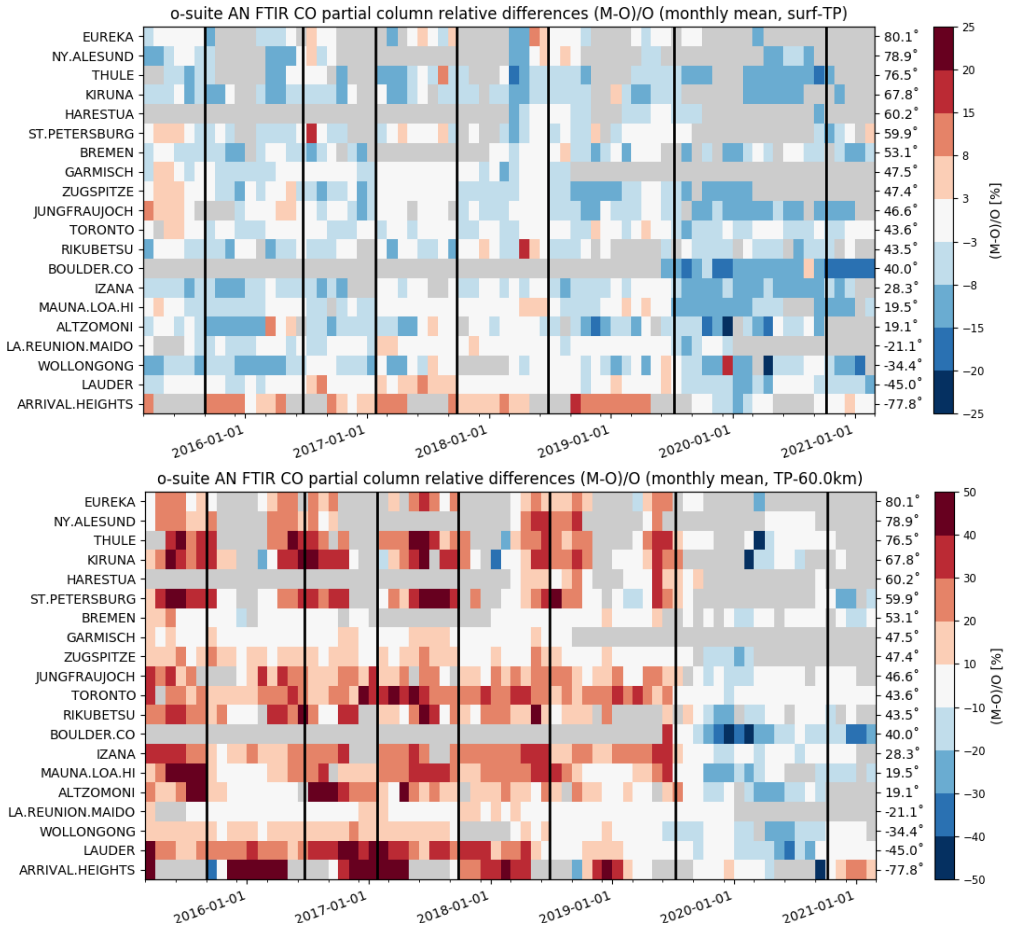


Figure S.4: Monthly-mean bias from 2015 up to the period DJF 2021 for tropospheric CO columns (top, mean relative difference in %) and stratospheric CO columns (bottom). The o-suite upgrades are indicated in black vertical lines, stations are sorted by latitude. The overall uncertainty for the CO measurements is approximately 3% on the tropospheric columns and 10% for the stratospheric columns. The o-suite analysis averaged bias in tropospheric columns increased to -6% for SON/DJF compared to -2% bias before the model update in July 2019. The bias in the stratosphere reduced to -8.5% compare to +18% before July 2019 and is comparable to the measurement's uncertainty.

According to IAGOS observations, CO is mostly underestimated over Frankfurt by both the o-suite and the control run. Largest underestimations appear in the lowest layers, while upper layers show a better agreement in general. The agreement with observations is better for the o-suite than for control run in the low to mid-troposphere. On average the MNMB ranges between -15% and 5% for the o-suite and between -30% and -10% for the control run in these layers. The correlation appears nearly constant from the surface to the free troposphere with values between 50 and 60%.

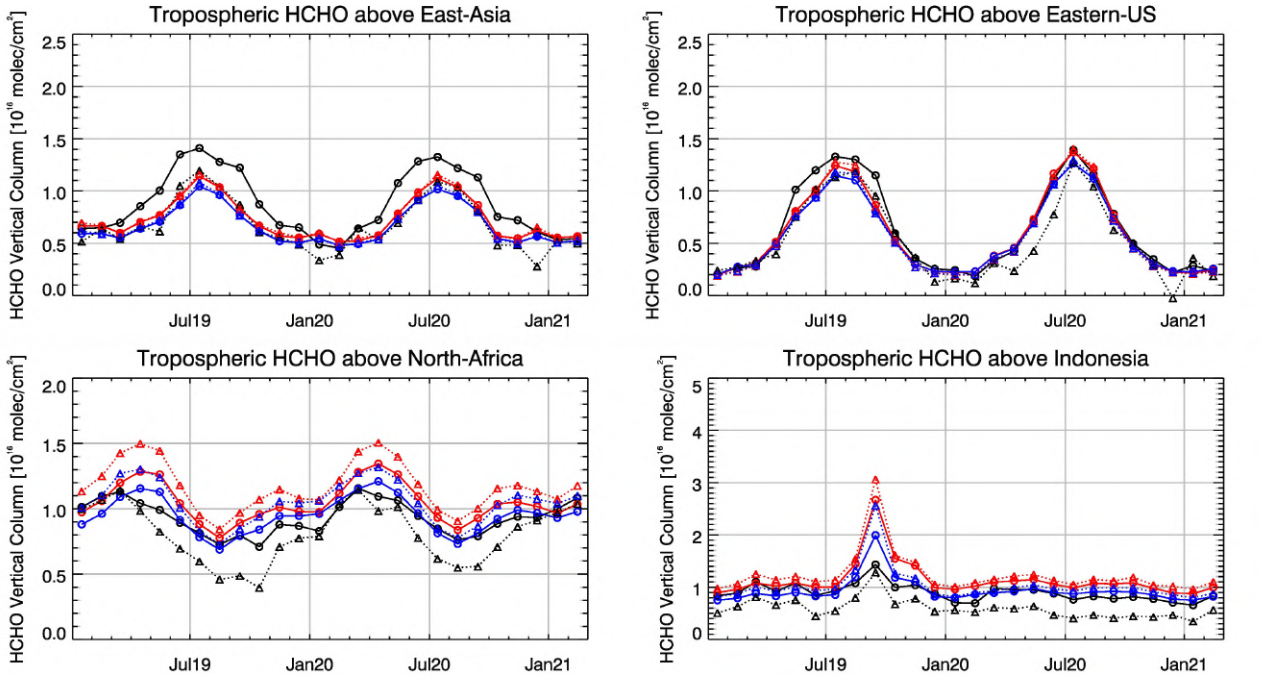


Figure S.5: Time series of average tropospheric HCHO columns [10^{16} molec cm^{-2}] from (black and grey) satellite retrievals, (blue) control and (red) o-suite model results for the period January 2019 to February 2021. The solid lines with circles show comparisons based on TROPOMI, the dotted lines with triangles show comparisons for GOME-2B. 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).

The comparison of IAGOS cruise data shows that the o-suite mostly underestimates CO with a bias absolute value below 15% for most regions of the world. The bias from the control run behaves similarly, with in general a larger bias than the o-suite.

For the evaluation with MOPITT satellite data in DJF-2021, the o-suite run shows good agreement with a negative bias within -10%, with some regional exceptions where the bias reaches -20%. The best agreement can be seen over the European, East and South Asian regions and Siberian fire region, with a bias within 5%. The evaluation with IASI satellite data shows biases in a range of ± 30 %, reflecting differences between the MOPITT and IASI retrievals.

Formaldehyde

Model validation with respect to Sentinel-5P TROPOMI and GOME-2B HCHO data 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 (Fig. S.5). The TROPOMI-based comparisons mainly show a very good agreement, while the comparison to GOME-2B shows a positive bias over main emission regions of HCHO.

The time series do not show evidence of a jump in regional HCHO concentration levels after the Cy47R1 upgrade.

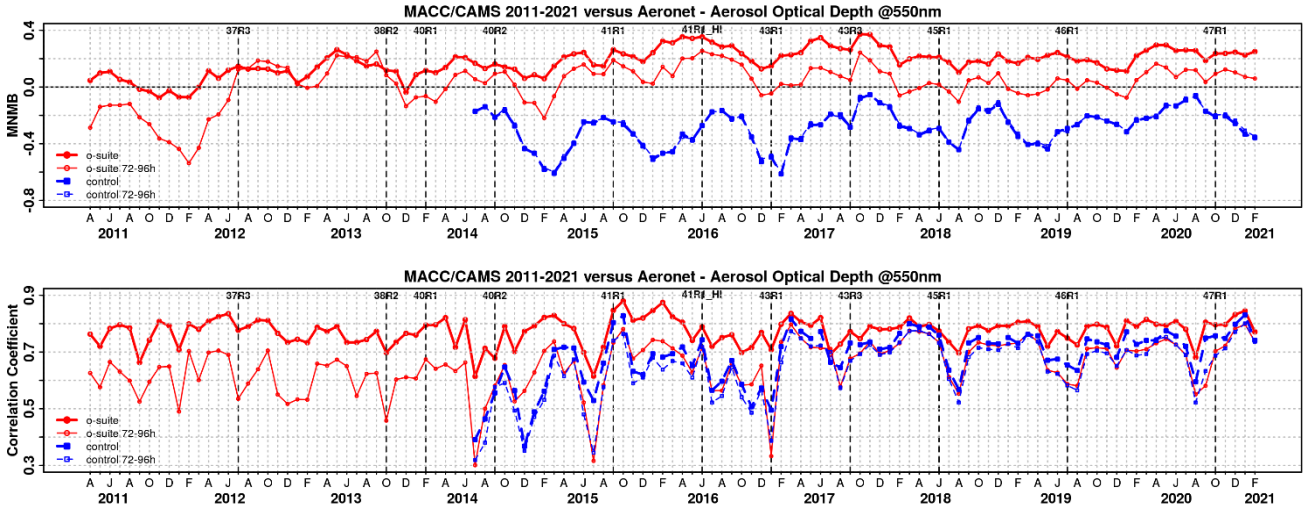


Figure S.6. Aerosol optical depth at 550nm in IFS 00Z model simulations for April 2011 – February 2021 against daily matching Aeronet Version3 level 1.5 data. a) Modified normalized mean bias (MNMB); o-suite (thick red curve); o-suite at last forecast day (light red curve); Control (blue dashed); Control at last forecast day (light blue dashed); b) Corresponding correlation coefficient. Model version changes are marked as vertical bars.

Aerosol

We estimate that the o-suite aerosol optical depth showed an average positive bias in the latest three months of +24%, measured as modified normalized mean bias against daily Aeronet (V3 level 1.5) sun photometer data. The 3-day forecasted aerosol distribution shows 18% less aerosol optical depth (AOD) than that from the initial forecast day, as shown in Fig. S.6-a. Spatiotemporal correlation, shown in Fig. S.6-b, was a little higher in DJF 2021 than in the winter 2020. The simulation reproduces approximately 47% of the day-to-day AOD variability across all Aeronet stations. The o-suite forecast at +3 days shows slightly lower correlation, because of imperfect forecasted meteorology and fading impact of the initial assimilation of MODIS AOD and MODIS fire info on model performance.

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 the latest upgrade in October 2020, the largest contributions to global AOD come from organics, sulphate and sea salt, dust decreased globally.

The aerosol Ångström exponent (AE) contains information about the size distribution of the aerosol, and implicitly about composition. The o-suite AE became more positive indicating a change to slightly more fine particles since the model upgrade to version 47R1 in October 2020, along with a decrease in correlation. MNMB Bias in AE increased from unbiased before to +35% since October 2020. The change is probably linked to a change in the sea salt parameterisation with smaller particles coming from the marine source. However, sulphate AOD has also increased.

PM10 and PM2.5, as defined by the IFS aerosol model, are evaluated against NRT data – including suburban sites, globally and a climatology. NRT data suggest on average during DJF 2020/2021 for North America, Europe, and China a PM10 MNMB bias -31%, +2%, -4% respectively and for PM2.5 a

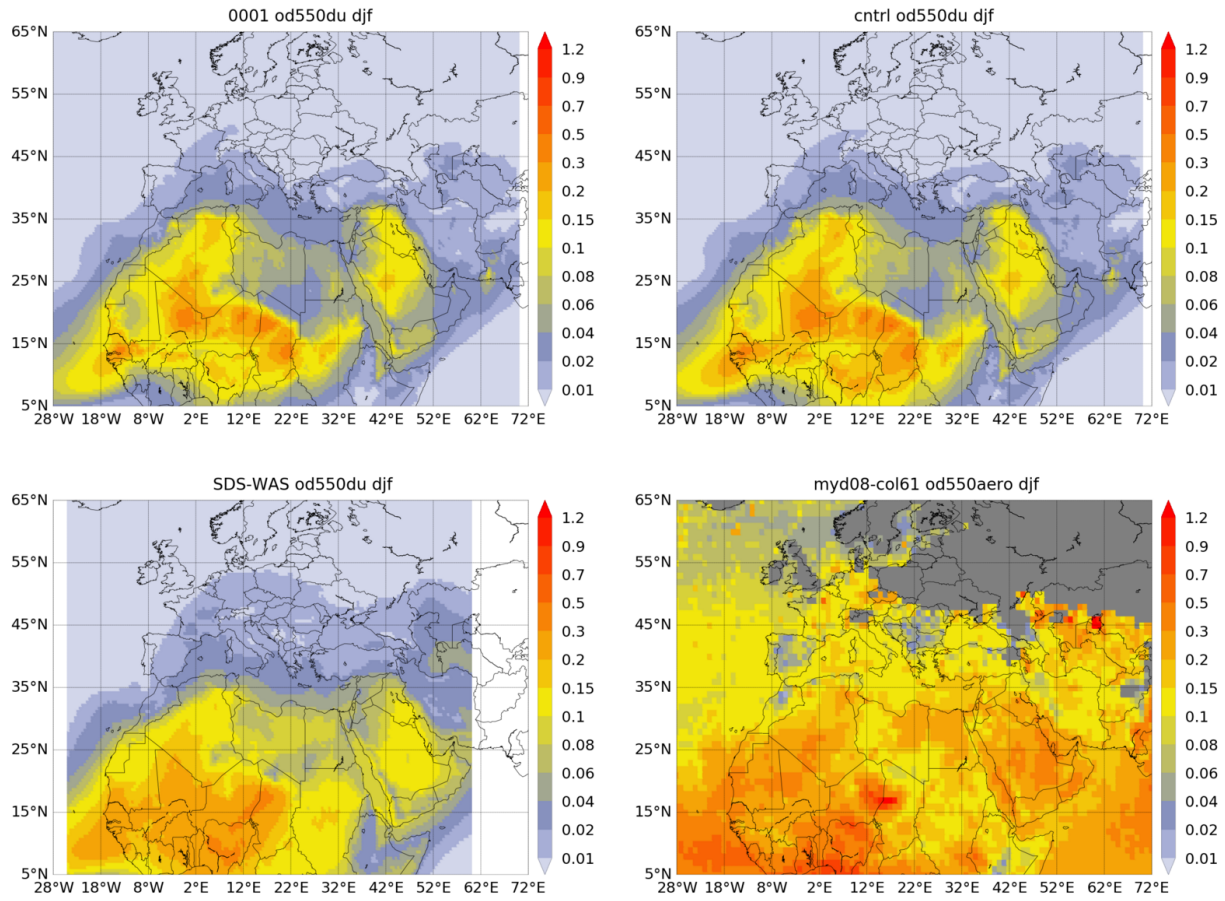


Figure S.7: Averaged DOD 24h forecast from o-suite (top left) and control (top right), DOD of the multi-model SDS-WAS Median product (bottom left) as well as AOD from MODIS/Aqua Collection 6.1 Level 3 combined Dark Target and Deep Blue product (bottom right) for the study period.

MNMB bias of -4%, -10% and +5% respectively. Regional and local variations of this bias are large. A second long-term IFS comparison is made against a climatological (period 2000-2009) average from 160 more remote, rural and background sites in North America and Europe. Note that observed PM levels from that period where higher than what is observed today. The fraction of PM10 and PM2.5 simulated data within a factor 2 of observed values stayed similar since July 2019 at levels of 40-60%, with higher levels in summer months. With the model version upgrade in July 2019 the PM2.5 has improved significantly, and both PM2.5 and PM10 have a clearer seasonal variation. The time series for the October 2020 upgrade (Cy47R1) is too short to draw conclusions.

During this season, the MODIS satellite observations (Figure S.7) show a low dust activity (AOD < 0.3). The highest AOD values are observed in the Bodélé. Higher AOD values (> 0.3) in the Guinea Gulf are associated with biomass burning from the Savannah fires. In the Middle East, the AOD seasonal average is under 0.3 with maximum AOD in the south-eastern Arabia Peninsula. Both CAMS runs show maximum DOD values in the Bodélé and surrounding regions in Chad and Mali, as well as Mauritania, Dakar, and Algeria. They appear overestimated compared to the SDS-WAS multi-model ensemble and the MODIS observations (Figure S.7). The Atlantic transport and values in Tunisia, Libya and Egypt seem to be underestimated. In the Middle East, both CAMS runs show a

maximum in northern-central Saudi Arabia and Iraq; the rest of the regions shows DOD < 0.06. From December to February, the o-suite (control) reproduces the daily variability of AERONET dust-filtered observations with a correlation coefficient of 0.61 (0.56) averaged over all AERONET sites, which is lower than the SDS-WAS multi-model product which has a correlation coefficient of 0.84. Regarding the mean bias (MB), the o-suite tends to underestimate the AERONET observations with a MB of -0.06 (and -0.05 for the control run) in comparison with the SDS-WAS multi-model that shows lower underestimations (MB of -0.03). In the Sahel, the o-suite assimilation enhances the underestimations observed in the control run (MB of -0.21 for control and -0.24 for o-suite) even though the o-suite better reproduces the observed daily variability (with a correlation of 0.53 for o-suite in comparison to 0.34 for control). In the Sahel, a larger discrepancy is observed with the SDS-WAS Multi-model result.

The comparison of the o-suite against German ceilometer data reveals a step to a larger low bias of the backscatter in the lowest layer at 1 km, starting with cycle 47r1 (Oct 2020). A moderate correlation (r = 0.4-0.9) between modelled and measured vertical profiles is found, while the model variance is generally smaller than the observation variance, except for Saharan dust days. Nine days with high Saharan dust load in February 2021 caused a clear overestimation of the monthly mean profile, even higher for the o-suite than the control run. The Saharan dust plumes are well captured by the model in space and time.

System performance in the Arctic

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

For the period from December 2020 to January 2021 the simulations of the surface ozone concentrations are on average in good agreement with the observations with high correlations and low bias (1-20%) at four of the sites, but a large overestimation (40%) at Barrow, Alaska. There is an overestimation of the ozone concentrations in the Arctic free troposphere (MNMB up to 10%) and UTLS (MNMB up to 20%).

System performance in the Mediterranean

During winter, both CAMS runs reproduce the daily variability of AERONET AOD observations, although a general overestimation in the whole Mediterranean Basin is observed. The correlation coefficient decreases from (0.78, 0.68 and 0.39) for control to (0.84, 0.71 and 0.41) and MB slightly increases from (-0.01, 0.03 and 0.04) for control to (0.04, 0.07 and 0.11) for o-suite respectively for the Western, Central and Eastern Mediterranean. The overestimations are linked to an enhanced aerosol background that is not directly linked to natural contributions. At surface levels, both CAMS runs show a higher correlation coefficient in north-western Europe (above 0.7) in comparison with the 3-hourly EEA PM10 and PM2.5 observations. For PM10, both CAMS runs show underestimations (MB under -4 $\mu g/m^3$) except in Central Europe and central-western Mediterranean which appear

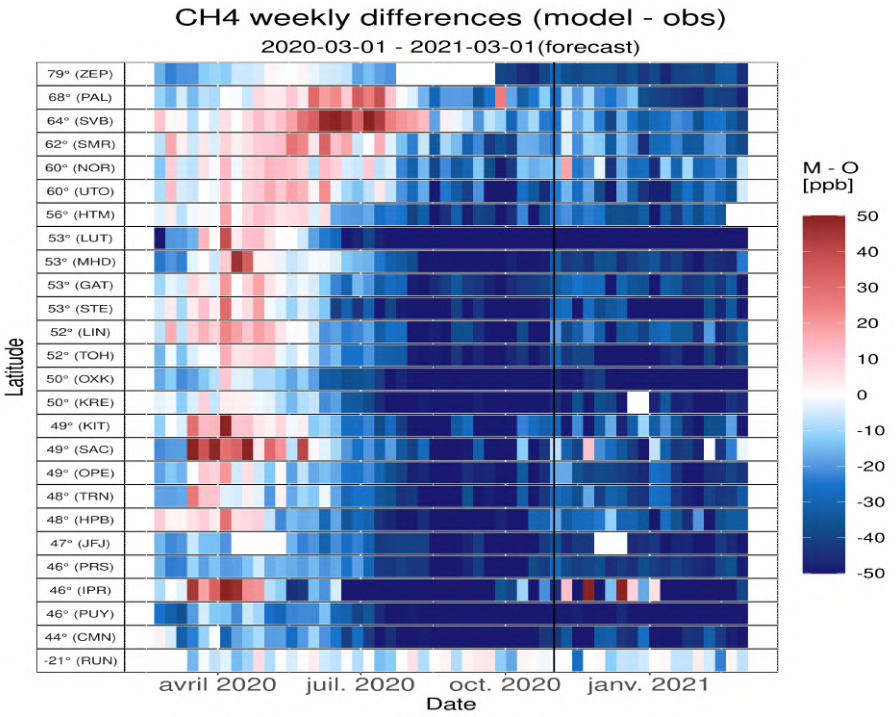


Figure S.8: Mosaic plot of CH_4 biases (in ppb) of the CAMS analysis, compared to surface station observations for the period March 2020 to February 2021. Each coloured vertical line represents a weekly mean.

overestimated (MB above 4 µg/m^3). Although the PM2.5 comparison shows smaller differences, the overestimations observed in PM10 in the o-suite in Central Europe are also detected in the PM2.5 o-suite comparison (MB above 4 µg/m^3). During winter, lower PM values are observed coincident with higher precipitation amounts. February shows two high PM10 and PM2.5 peaks that are linked to dust outbreaks, associated with concentrations above 50 µg/m^3 for PM10 in southern and central European sites. Overestimations are observed for sites mainly influenced by maritime aerosols.

The model is compared to surface O_3 observations from the AirBase network. Our analysis shows that model mostly overestimates surface ozone values with MNMBs varying between -23% and 51% depending on the station. Temporal correlation coefficients between simulated and observed surface ozone for both the o-suite and control run are highly significant over the entire Mediterranean from Gibraltar to Cyprus.

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.

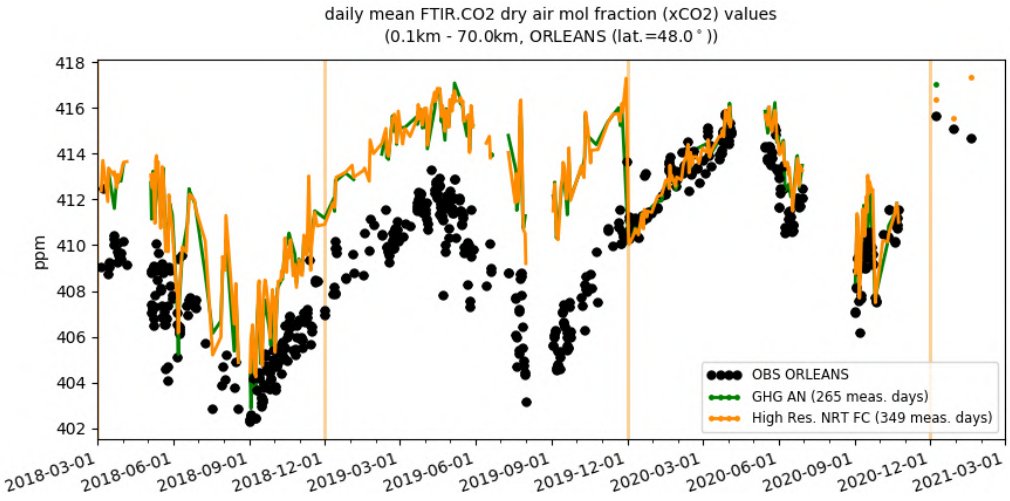


Figure S.9: Comparison of the CO_2 CAMS configurations with TCCON CO_2 at Orleans, for the period March 2018 to February 2021.

Since the new experiments started in November 2020 both ICOS (surface, Fig. S.8) and TCCON (total columns) observations showed CH_4 negative biases ranging from -10 to -50 ppb. The bias was positive at most ICOS sites in North hemisphere between April and July 2020 depending on the latitude and is also decreased in June for the northern hemisphere TCCON sites. NDACC partial columns indicate also negative biases in the troposphere, but slightly positive biases in the stratosphere.

For CO_2 both surface and total column measurements show biases generally within ±1%, with a maximum in late spring or summer at the surface, and later (September) for the TCCON sites (Fig. S.9). Higher biases are observed in autumn 2020, when the CAMS experiments fail to reproduce the high CO_2 enhancements observed for several days especially in Northern Europe.

Tropospheric Water Vapour (H_2O)

Overall, water vapour values and variability are well represented by the two runs in the low troposphere over Frankfurt with small positive biases (< 10%) and high correlation values (> 90%). The agreement is worse in the upper layers, with larger biases and smaller correlation and mostly negative biases in the UTLS.

The comparison of IAGOS cruise data show that the results of the two runs are very similar for all regions with an overall underestimations of water vapour and a bias in absolute value mostly smaller than 50%.

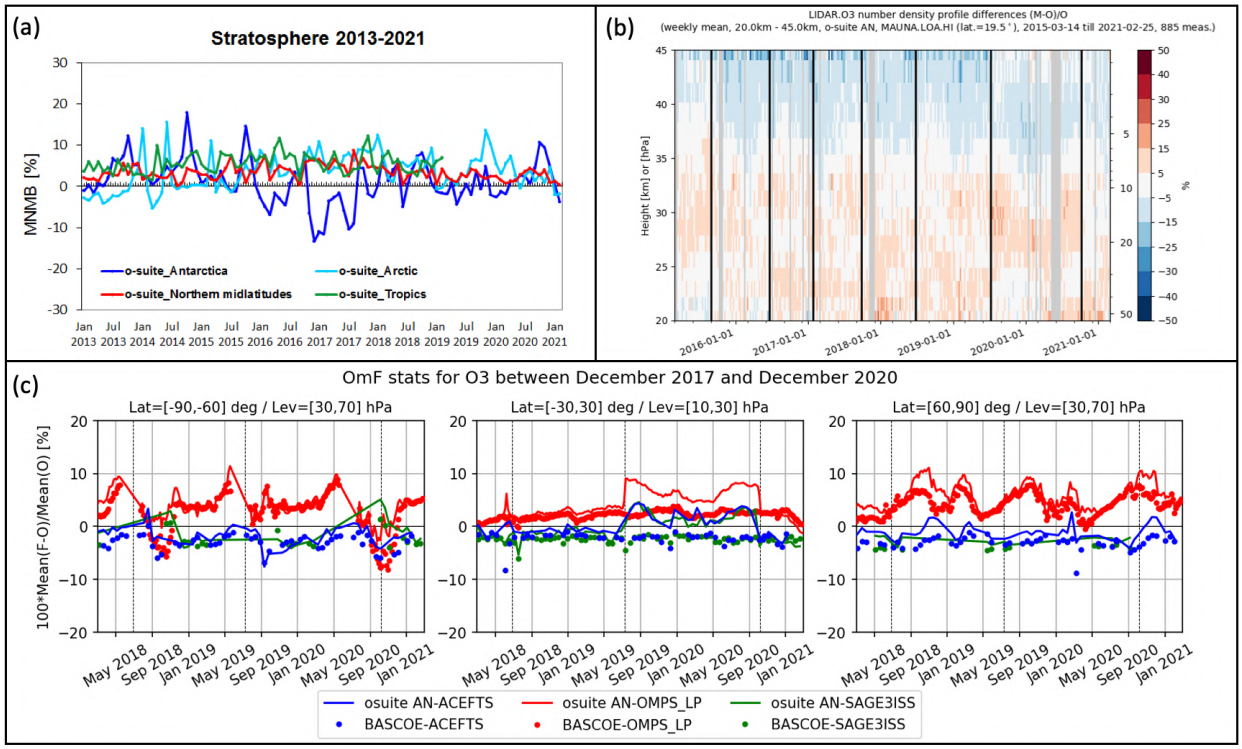


Figure S.10: (a): MNMBs (%) of ozone in the stratosphere from the o-suite against aggregated ozone sonde data in the Arctic (light blue), Antarctic (dark blue) northern midlatitudes (red) and tropics (green) from 2013 to February 2021.The stratosphere is defined as the altitude region between 60 and 10 hPa in the tropics and between 90 and 10 hPa elsewhere. (b): Comparison of the weekly mean profile bias between the O_3 mixing ratios of the 1-d forecast and the NDACC LIDAR at Mauna Loa. (c): Time series comparing model runs to observations for the period 2017-03-01 to 2021-02-28 in three latitude bands and three pressure layers (left: 90°S-60°S between 30 and 70 hPa, centre: 30°S-30°N S between 10 and 30 hPa and right: 60°N-90°N between 30 and 70 hPa) for the o-suite analyses (solid lines) and BASCOE (dotted lines) against observations from OMPS-LP v2.5 (red), ACE-FTS v3.6 (blue) and SAGE-III v5.1 (green). Shown is the normalized mean bias (model-observation)/observation (%).

Ozone layer and UV

Ozone partial columns and vertical profiles

The Autumn 2020, Winter 2021 period was marked by two noticeable facts. The first one was the exceptionally deep ozone hole that happened above Antarctica. The second fact was the implementation of the new CY47r1 o-suite on October 6, 2020. Stratospheric ozone from the CAMS o-suite seems not have been affected by the first fact and the agreement with independent observations was similar to previous years (as seen in the time series against ACE-FTS for the South Pole region). The CY47R1 upgrade, on the other hand, allowed an overall improvement in the agreement with independent observations, especially in the tropical region and in the pressure range between 10-30 hPa.

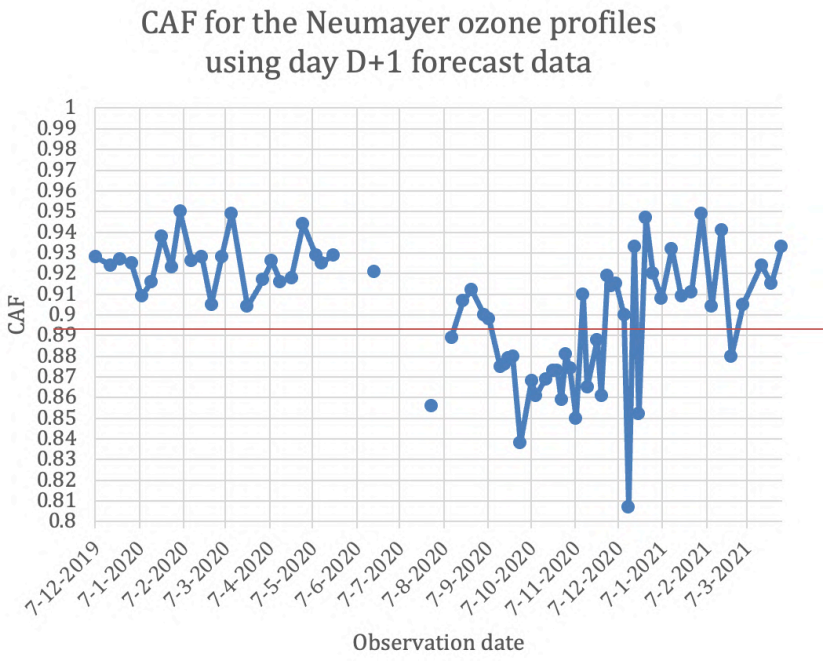


Figure S.11. Time series of the CAF values calculated between CAMS model 1-day forecast data and balloon O_3 sonde measurements at Neumayer station in Antarctica. The orange line corresponds to the KPI target minimum value of 0.9. Time series extends to 1 April 2021.More detailed comparisons between CAMS stratospheric ozone and independent observations is given in Sect. 8 where it includes observations from ozone sonde profiles, ground-based instruments from NDACC (Network for the Detection of Atmospheric Composition Change, http://www.ndacc.org) and vertical profiles from 3 satellite instruments (OMPS-LP, ACE-FTS and SAGE-III). Furthermore, o-suite analyses are also compared with the analyses delivered by the independent BASCOE assimilated analyses.

In summary, the o-suite O_3 partial pressures are slightly overestimated in all latitude bands (MNMBs between -0.5 to 4.3%, and close to zero in DJF-2021) compared to ozone sondes. Comparisons with the NDACC observations show a generally good agreement, with small performance differences between AN and 1d forecasts. The comparison with independent satellite profiles is generally in good agreement for the considered period, usually within 10%.

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. NO_2 has also been evaluated but the results show that this product is not mature and shows large biases.

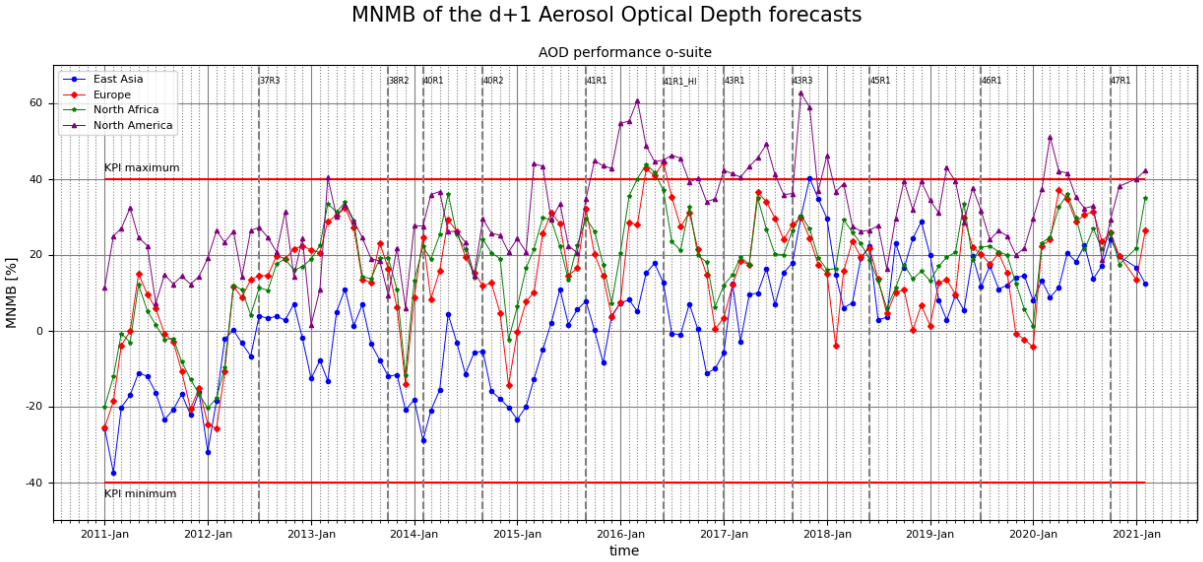


Figure S.12. Modified Normalized Mean Bias (MNMB) for the d+1 Aerosol Optical Depth (AOD) at four main regions, based on observations from AERONET stations. The horizontal red lines show allowed minimum and maximum values for this KPI. The vertical dashed lines correspond to the dates when new CAMS o-suite upgrades became operational.

Events

Dusty February 2021 in Europe and the North Tropical Atlantic: During February 2021, several intense Africa dust events influenced Europe and the Canary Islands. On 5-6 February, African dust was transported from northeast Algeria causing reddish skies in large parts of Europe. This event coincided with snow and rain, giving the Pyrenees and the Alps a brown appearance. On 15-19 February, a Saharan dust outbreak severely affected the Canary Islands before heading toward continental Europe, reaching as far north as Scandinavia. Finally, on 21 February a new dust event with origin in Algeria hit the Western Mediterranean. The comparison with MODIS shows how the AOD o-suite can reproduce the spatial and temporal distribution of the observed AOD plumes. However, DOD appears underestimated, particularly in long-range transport regions in Europe. During the extreme aerosol events, data assimilation enhances the predicted AOD peaks compared to the control run but the dust OD is not adjusted proportionally.

Key performance indicators for the CAMS o-suite

Two statistical measures with chosen threshold values, plus a production requirement were introduced as Key Performance Indicators (KPIs) for the CAMS global o-suite.

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. The o-suite data delivery for the reporting period December 2020 – February 2021 was excellent, with an on-time percentage of 100%.

The ozone profile Common Area Fraction (CAF) is a measure of the level of agreement between the CAMS ozone profile and corresponding measurements from balloon sondes. For the KPI these are

evaluated for the observations at the Neumayer station in Antarctica. Here a CAF value of 1 indicates a perfect match between the measured and forecasted profiles. For the KPI a challenging target minimum value of 0.9 is chosen. Values of CAF for collocated measured and modelled ozone profiles for the period from end of 2019 to the first months of 2021 are shown in Fig. S.11. The gaps in the middle of the time series correspond to cases where the balloon did not reach the 30 hPa level, which is the minimum height required for a meaningful comparison. The CAMS o-suite fulfils the ozone CAF predefined requirement mainly at the first half and at the end of the examined period. In the period December 2020 – February 2021, CAF values generally lie above the 0.9 value threshold. Local ozone hole conditions (September-December), which are accompanied by strong gradients in the ozone profile, pose an additional challenge for the analysis.

The Modified Normalized Mean Bias of the Aerosol Optical Depth (MNMB AOD) is an indicator of the agreement (in terms of statistical bias) between the o-suite 1-day forecast AOD of CAMS and corresponding estimations from AERONET stations. Values are reported separately for East Asia, Europe, North Africa, and North America. The KPI target is for MNMB AOD values to lie within the range ±40%. Results, reported monthly, are shown in Fig. S.12. The time series begins in 2011, allowing for all model versions since then to be evaluated (vertical dashed lines in the figure correspond to commencement of successive model versions). Overall, the results lie within the KPI limits, except for some cases in North America in the middle of the time series. In the second half of the time series, results from the four regions converge to positive values, but below the KPI limit.

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, Eskes et al. (2021), 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). 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, accessible through the Atmosphere Data Store on http://atmosphere.copernicus.eu/data.

A summary of the system and its recent changes is given in section 2. Subsequent sections give 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 Cy47R1 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-2, OMI, OMPS, TROPOMI	Sonde, lidar, MWR, FTIR, OMPS, ACE-FTS, SAGE3-ISS and BASCOE analyses
O_3, UT/LS	MLS	IAGOS, ozone sonde
O_3, free troposphere	Indirectly constrained by limb and nadir sounders	IAGOS, ozone sonde, IASI
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	TROPOMI, SCIAMACHY, GOME-2, MAX-DOAS
HCHO		TROPOMI, GOME-2, MAX-DOAS
SO_2	GOME-2, TROPOMI (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.

The CAMS global evaluation server
https://global-evaluation.atmosphere.copernicus.eu
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://www.mpimet-cams.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 Aeronet verification of CAMS NRT forecasts: https://atmosphere.copernicus.eu/charts/cams_aeronet_ver/?facets=undefined&time=2019020100,0,2019020100&site=ARM_Graciosa
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

The CAMS global evaluation server, https://global-evaluation.atmosphere.copernicus.eu, became available in Summer 2019. This server combines many of the individual verification results shown on the other CAMS web pages listed in Table 1.2 and presents the comparisons through a uniform interface.

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 (the o-suite and control run)

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.

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

Forecast system	Exp. ID	Brief description	Upgrades (e-suite ID)	Cycle
O-suite	0001	Operational CAMS DA/FC run	20210519-present 20201006-20210518 20190709-20201006 20180626-20190708 20170926-20180625 20170124-20170926 20160621-20170124 20150903-20160620 20140918-20150902	47R2 47R1 46R1 45R1 43R3 43R1 41R1 41R1 40R2
Control	hj7b hdir h7c4 gzhy gsyg gnhb gjjh geuh g4o2	control FC run without DA	20210519-present 20201006-20210518 20190709-20201006 20180626-20190708 20170926-20180625 20170124-20170926 20160621-20170124 20150901-20160620 20140701-20150902	47R2 47R1 46R1 45R1 43R3 43R1 41R1 41R1 40R2
GHG run	hd7v	Tco399L137 NRT CO_2, CH_4 analyses (~25km)	20201101-present	47R1
	he9h	High resolution Tco1279 (~9km) NRT CO_2, CH_4 forecast	20201101-present	47R1
	h72g	Tco399L137 NRT analyses	20191201-20201031	46R1
	h9sp	Tco1279 forecast	20191201-20201031	46R1
	gwx3 gqiq	GHG analysis Tco399 (~25km)	20181201-20191130 20170101-20181130	45R1 43R1
	gznv gqpe	High resolution Tco1279 (~9km) NRT CO_2, CH_4 forecast	20181201-20191130 20170101-20181130	45R1 43R1
	ghqy gf39	High resolution T1279, NRT CO_2 and CH_4 without DA	20160301-20170621 20150101-20160229	

Table 2.2: Satellite retrievals of reactive gases and aerosol optical depth that are actively assimilated in the o-suite in Cy47R1.

Instrument	Satellite	Provider	Version	Type	Status
MLS	AURA	NASA	V4	O3 Profiles	20130107 -
OMI	AURA	NASA	V883	O3 Total column	20090901 -
GOME-2	Metop-A	Eumetsat	GDP 4.8	O3 Total column	20131007 - 20181231
GOME-2	Metop-B	Eumetsat	GDP 4.8	O3 Total column	20140512 -
GOME-2	Metop-C	Eumetsat	GDP 4.9	O3 Total column	20200505 -
SBUV-2	NOAA-19	NOAA	V8	O3 21 layer profiles	20121007 - 20201005
OMPS	Suomi-NPP	NOAA / EUMETSAT		O3 13-layer profiles	20170124 – 20190409 20201006-
OMPS	NOAA-20	NOAA / EUMETSAT		O3 13 layer profiles	20201006-20201215
TROPOMI	Sentinel-5P	ESA		O3 column	20181204-
IASI	MetOp-A	LATMOS/ULB Eumetsat	-	CO Total column	20090901 - 20180621 20180622 - 20191118
IASI	MetOp-B	LATMOS/ULB Eumetsat	-	CO Total column	20140918 - 20180621 20180622 -
IASI	MetOp-C	Eumetsat		CO total column	20191119 -
MOPITT	TERRA	NCAR	V5-TIR V7-TIR V7-TIR Lance V8-TIR	CO Total column	20130129 - 20160124 - 20180626 20180626 20190702
OMI	AURA	KNMI	DOMINO V2.0	NO2 Tropospheric column	20120705 -
GOME-2	MetOp-A	Eumetsat	GDP 4.8	NO2 Tropospheric column	20180626 - 20200504
GOME-2	MetOp-B	Eumetsat	GDP 4.8	NO2 Tropospheric column	20180626 -
GOME-2	MetOp-C	Eumetsat	GDP 4.9	NO2 Tropospheric column	20200505-
GOME-2	MetOp-A	Eumetsat	GDP 4.8	SO2 Total column	20150902 -
GOME-2	MetOp-B	Eumetsat	GDP 4.8	SO2 Total column	20150902-20200414
GOME-2	MetOp-C	Eumetsat	GDP 4.9	SO2 Total column	20200505-
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 -

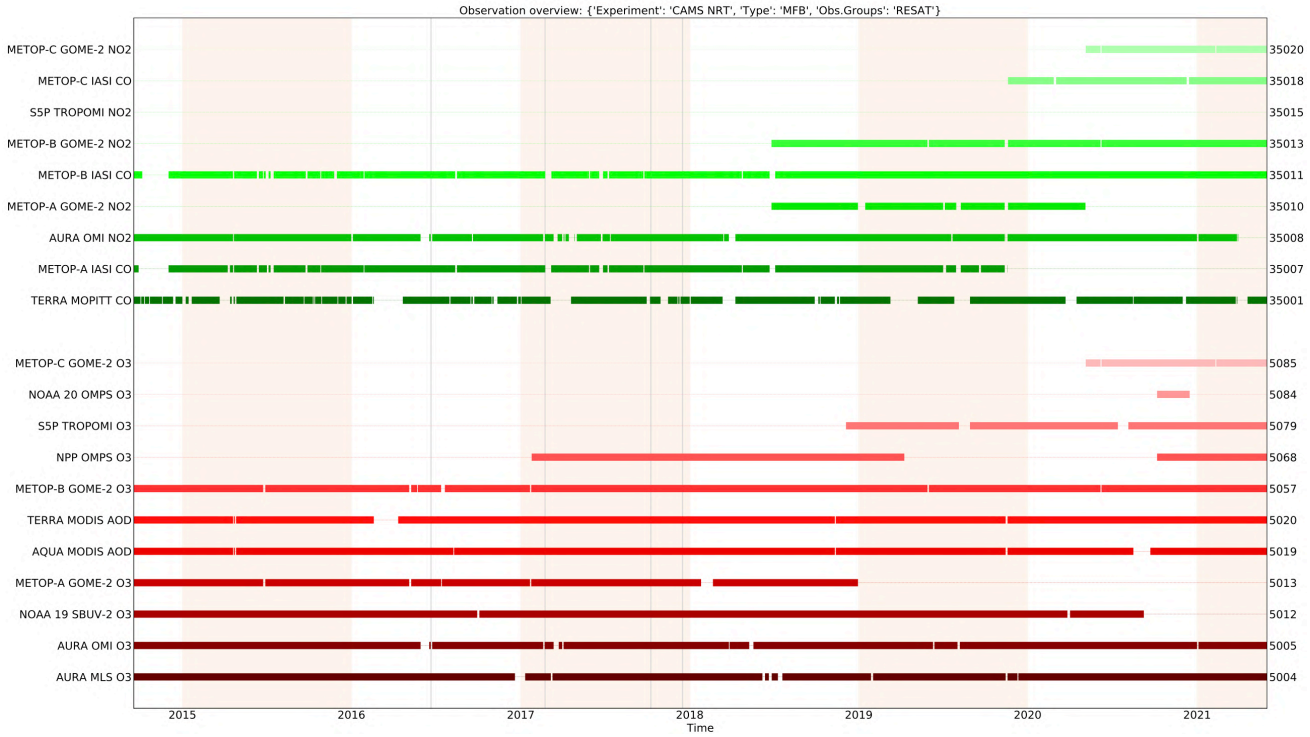


Figure 2.1: Satellite observation usage in the real-time analysis, for ozone, CO, aerosol AOD and NO_2, from October 2014 onwards. Top rows (in green): products assimilated using averaging kernels. Sentinel-5P TROPOMI ozone is assimilated since Dec. 2018 and other products from TROPOMI are monitored.

Further details on the different model runs and their data usage can be found at http://atmosphere.copernicus.eu/documentation-global-systems.

2.1.1 The CAMS 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.

A short summary of the main model specifications:

- 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). Anthropogenic emissions changed to CAMS_GLOB v2.1 with the July 2019 update.
- NRT fire emissions are taken from GFASv1.2 (Kaiser et al. 2012).

The aerosol model includes 14 prognostic variables (Remy et al., 2019).

- 3 size bins each for sea-salt and desert dust
- 2 bins (hydrophibic and hydrophobic) each for organic matter and black carbon
- 1 bin for sulphate
- 2 bins (fine and coarse) for nitrate (New since 46R1)
- 1 bin for ammonium (New since 46R1)

The SO_2 precursor for sulphate aerosol no longer exists as a separate prognostic in the aerosol scheme, which since 46R1 couples directly to the SO_2 in the chemistry scheme instead. Likewise, the precursors for the new nitrate and ammonium aerosol (nitric acid and ammonia) are also part of the chemistry scheme rather than the aerosol scheme.

Aerosol total mass is constrained by the assimilation of MODIS and PMAp AOD (Benedetti et al. 2009). A variational bias correction is currently applied for the PMAp AOD based on the approach used also elsewhere in the IFS (Dee and Uppala, 2009).

A history of updates of the o-suite is given in Table 2.3, and is documented in earlier MACC-VAL and CAMS 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 2.3.

The upgrade of the system to Cy45r1 took place on 26 June 2018. This upgrade is also relevant for this report (for the period up to 8 July), and the validation for this upgrade is described in Eskes et al., 2018b/2018c.

Note on the fire emissions between mid-August to 6 October:
Aqua MODIS products were unavailable following an issue with the satellite between 16 August and 8 September. The Aqua MODIS FRP observations were “blacklisted” in GFAS when the data went offline and, unfortunately, a switch in the suite was not set correctly when observations became available again in early September. This means that the fire emissions used in the CY46R1 o-suite were reduced due to the missing Aqua observations. The largest impact is in the tropics. The hourly GFAS suite picked up the Aqua MODIS FRP correctly so the CY47R1 o-suite is not affected.

Table 2.3: 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
2019.07.09	Update to cycle 46R1_CAMS, T511L137
2020.10.06	Update to cycle 47R1_CAMS, T511L137

2.1.2 Short description of the CAMS upgrade (47r1) of 6 October 2020

The last major upgrade of the CAMS global system relevant for this report is based on IFS version cy47r1_CAMS. For the aerosol and reactive trace gas components the upgrade took place on 6 October 2020. For the greenhouse gases the upgrade to 47R1 took place on 1 November 2020. see https://atmosphere.copernicus.eu/cycle-47r1 or https://confluence.ecmwf.int/display/COPSRV/Current+global+production+suites.

The validation for this 47r1 upgrade is described in Eskes et al. 2020: https://atmosphere.copernicus.eu/sites/default/files/2020-10/CAMS84_2018SC2_D3.2.1-202009_esuite.pdf

The meteorological changes can be found on the ECMWF-IFS CY47R1 page, https://confluence.ecmwf.int/display/COPSRV/Implementation+of+IFS+cycle+47r1.

The atmospheric composition content of the 47R1 cycle includes the following aspects:

Assimilation:

- No changes compared to 46R1

Observations:

- TROPOMI volcanic SO_2:
    - Activation on 6 Oct 2020 (for SO2 > 5DU)
- OMPS O_3 layers from NOAA-20 and NPP activated on 20201006
- SBUV/2 NOAA-19 O_3 layers retired on 20201005
- No other changes compared to the observations used in 46R1.

Emissions:

- Updated emissions inventories: CAMS_GLOB_ANT v4.2 (anthropogenic) and volcanic outgassing (based on Carn et al., 2017).
- Updated to GFASv1.4 biomass-burning emissions.
- Excluded agricultural waste burning from CAMS_GLOB_ANT, avoiding double-counting with GFAS.
- Improved diurnal cycle (CO, NO, SO_2, NH_3) and vertical profile for anthropogenic emissions (SO_2, all over sea).

Other model changes:

- Hybrid Linear Ozone (HLO) scheme (a Cariolle-type linear parameterisation of stratospheric ozone chemistry using the multi-year mean of the CAMS reanalysis as mean state).
- New sea-salt emission scheme based on Albert et al. (2016), providing better agreement with measured sea-salt size distribution.
- Updated dust source function, reducing excess dust in the Sahara, Middle East and other regions, and restoring missing dust over Australia.
- Revised coefficients in UV processor, based on ATLAS3 spectrum.

2.1.3 Control

The control run (relevant expver = **gzhy,** since 26/06/2018; expver = **h7c4** since 09/07/2019; expver = **hdir** since 06/10/2020) applies the same settings as the respective o-suites, based on the coupled IFS-CB05 system with CAMS aerosol, except that data assimilation is not switched on. The meteorology in the control run is initialized with the meteorological fields from the o-suite.

2.1.4 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 at a resolution of TL1279, i.e. ~16 km horizontal, and with 137 levels. This system runs in real time 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, 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:

- "**hd7v**" NRT CO_2, CH_4 analyses from 1 November 2020, with a resolution Tco399 (~25km) and 137 vertical levels. Cycle 47R1.
- "**he9h**" NRT CO_2, CH_4 forecasts from 1 November 2020, with high resolution Tco1279 (~9km) and 137 vertical levels. Cycle 47R1.
- "**h72g**" NRT CO_2, CH_4 analyses from 1 December 2019, with a resolution Tco399 (~25km) and 137 vertical levels. Cycle 46R1.
- "**h9sp**" NRT CO_2, CH_4 forecasts from 1 December 2019, with high resolution Tco1279 (~9km) and 137 vertical levels. Cycle 46R1.
- "**gqpe**" (43R1) from January 2017, and "**gznv"** (45R1) from 1 December 2018 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 runs on a TCO399 grid (equivalent to around 25km) and 137 vertical levels and is available from January 2017 ("**gqiq**", 43R1) and 1 December 2018 ("**gwx3**", 45R1). 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).
- "**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.

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 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 multi-model dust optical depth (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 data are available from the CAMS Atmosphere Data Store (ADS), https://atmosphere.copernicus.eu/data, in netcdf and grib2 format.

2.4 Availability and timing of CAMS products

The o-suite data delivery for the period December 2020 - February 2021 (DJF-2021) was excellent, with an on-time delivery percentage of 100%. See table 2.6 for detailed statistics from 2014 onwards. 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. Since 21 June 2016 two CAMS forecasts are produced each day.

Table 2.6: Timeliness of the o-suite from December 2014. 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
Dec 2018 - Feb 2019	98.9%	D+0, 09:03 D+0, 21:04	D+0, 09:06 D+0, 21:06	D+0, 09:08 D+0, 21:10

Mar-May 2019	100%	D+0, 09:07 D+0, 21:05	D+0, 09:10 D+0, 21:09	D+0, 09:12 D+0, 21:11
Jun-Aug 2019	99.5%	D+0, 09:19 D+0, 21:14	D+0, 09:22 D+0, 21:17	D+0, 09:27 D+0, 21:19
Sep-Nov 2019	98.9%	D+0, 09:14 D+0, 21:07	D+0, 09:23 D+0, 21:20	D+0, 09:26 D+0, 21:24
Dec 2019 - Feb 2020	99.4%	D+0, 09:00 D+0, 20:58	D+0, 09:03 D+0, 21:02	D+0, 09:12 D+0, 21:08
Mar-May 2020	100%	D+0, 08:55 D+0, 20:57	D+0, 08:58 D+0, 21:01	D+0, 09:00 D+0, 21:05
Jun-Aug 2020	100%	D+0, 08:58 D+0, 20:55	D+0, 09:03 D+0, 20:59	D+0, 09:05 D+0, 21:02
Sep-Nov 2020	100%	D+0, 08:50 D+0, 20:50	D+0, 08:58 D+0, 20:53	D+0, 09:04 D+0, 21:01
Dec 2020 - Feb 2021	100%	D+0, 08:54 D+0, 20:52	D+0, 09:00 D+0, 20:55	D+0, 09:04 D+0, 21:00

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 January 2013 to February 2021 (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.

Please note that recent scientific findings (Thompson et al., 2017; Witte et al., 2017; 2018, Stauffer, et al. in preparation 2020;) show a drop-off in Total Ozone at various global ozone stations in comparison with satellite instruments. This drop-off amounts between 5-10% for stratospheric ozone. Changes in the ECC ozone instrument are associated with the drop-off, but no single factor has been identified as cause yet. For tropospheric ozone (<50 hPa) no alternations are reported, but some systematic effect cannot be ruled out. Data availability is thus recently limited.

MNMBs for the o-suite are mostly within the range ±20%, for all months, in all zonal bands, except for the Tropics and Antarctica, where larger positive MNMBs up to ±45% appear, see Fig. 3.1.4. During the last year (February 2020 to February 2021) MNMBs are within – 2% to 20% over the Arctic, within 15% over the northern mid-latitudes and within -17 to 10% over Antarctica, see Fig. 3.1.1.-3.1.4.

In the UTLS (Fig 3.1.3, right), MNMBs are mostly positive over all regions and are within 20%.

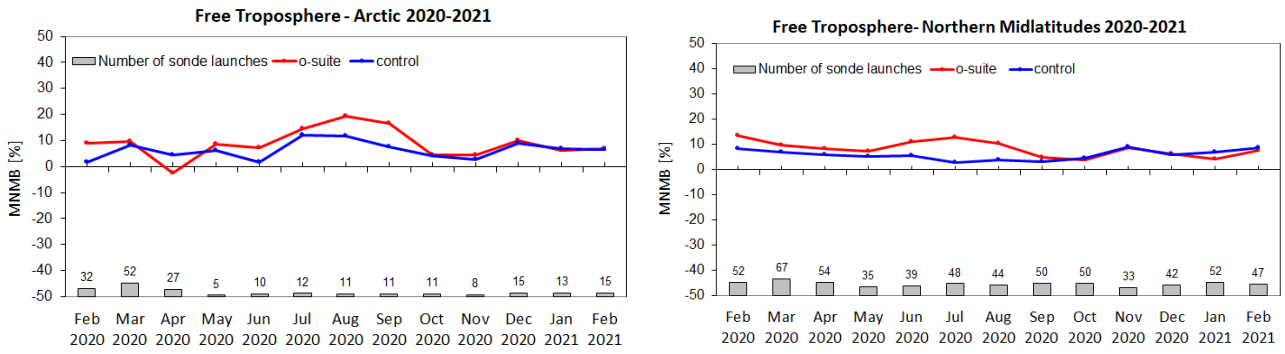


Figure 3.1.1: MNMBs (%) of ozone in the free troposphere (between 750 and 300 hPa) from the IFS model runs against aggregated sonde data over the Arctic (left) and the Northern mid latitudes (right) for the period February 2020 – February 2021. The numbers indicate the amount of individual number of sondes.

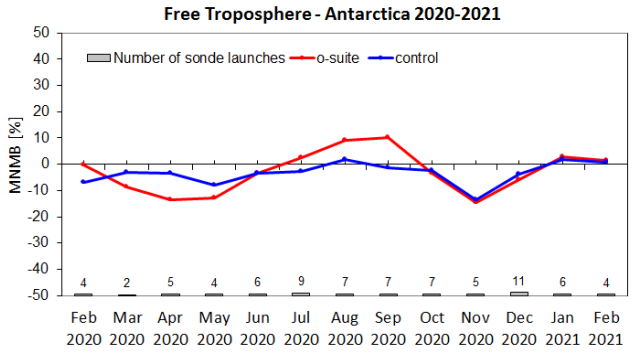


Figure 3.1.2: MNMBs (%) of ozone in the free troposphere (between 750 and 300 hPa) from the IFS model runs against aggregated sonde data over Antarctica for the period February 2020 – February 2021. The numbers indicate the amount of individual number of sondes.

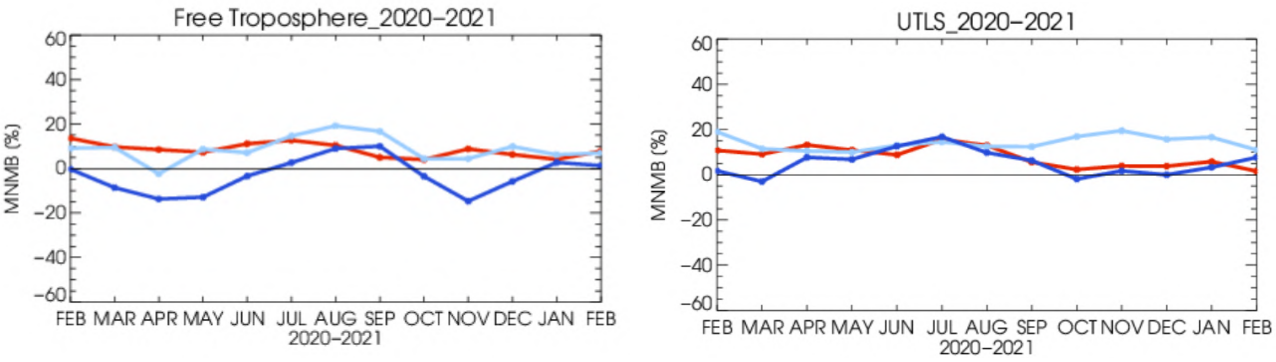


Figure 3.1.3: MNMBs (%) of ozone in the free troposphere (left, between 750 and 300 hPa) and UTLS (right, between 300 and 100 hPa) from the IFS model runs against aggregated sonde data over Antarctica (blue), Arctic (light blue) and Northern mid-latitudes (red). Period: February 2020 – February 2021.

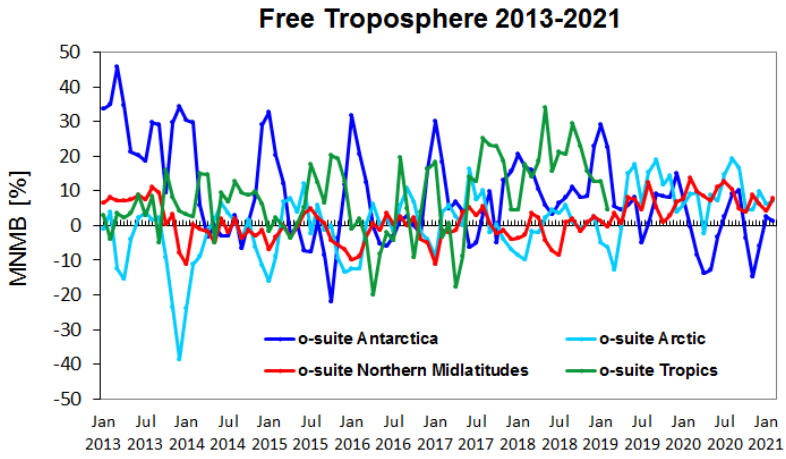


Figure 3.1.4: Time series of MNMB of ozone in the o-suite, period January 2013 – February 2021, compared against ozone sondes, averaged over different latitude bands. The free troposphere is defined as the layer between 750 and 300 hPa. There was a problem with sonde data in the tropics in 2020, see https://tropo.gsfc.nasa.gov/shadoz/.

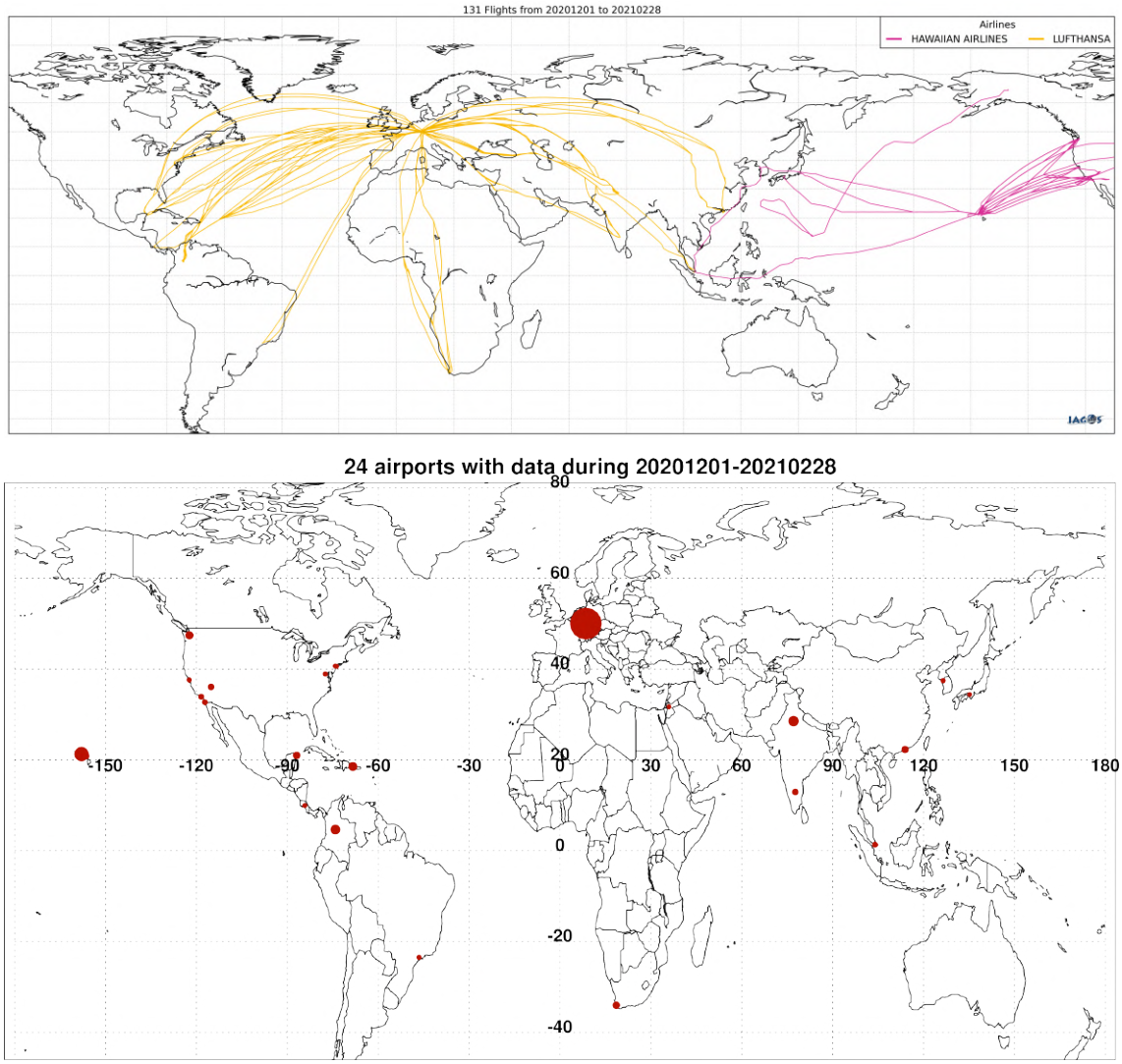


Figure 3.2.1.: Map of the flights (top) and the visited airports (bottom) during the period December 2020 - February 2021, by the IAGOS-equipped aircraft. The size of the plotting circle represents the number of profiles available.

3.2 Ozone validation with IAGOS data

The daily profiles of ozone measured at airports around the world are shown on the CAMS website at http://www.iagos-data.fr/cams/nrt_profiles.php. For the period from December 2020 - February 2021, 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.

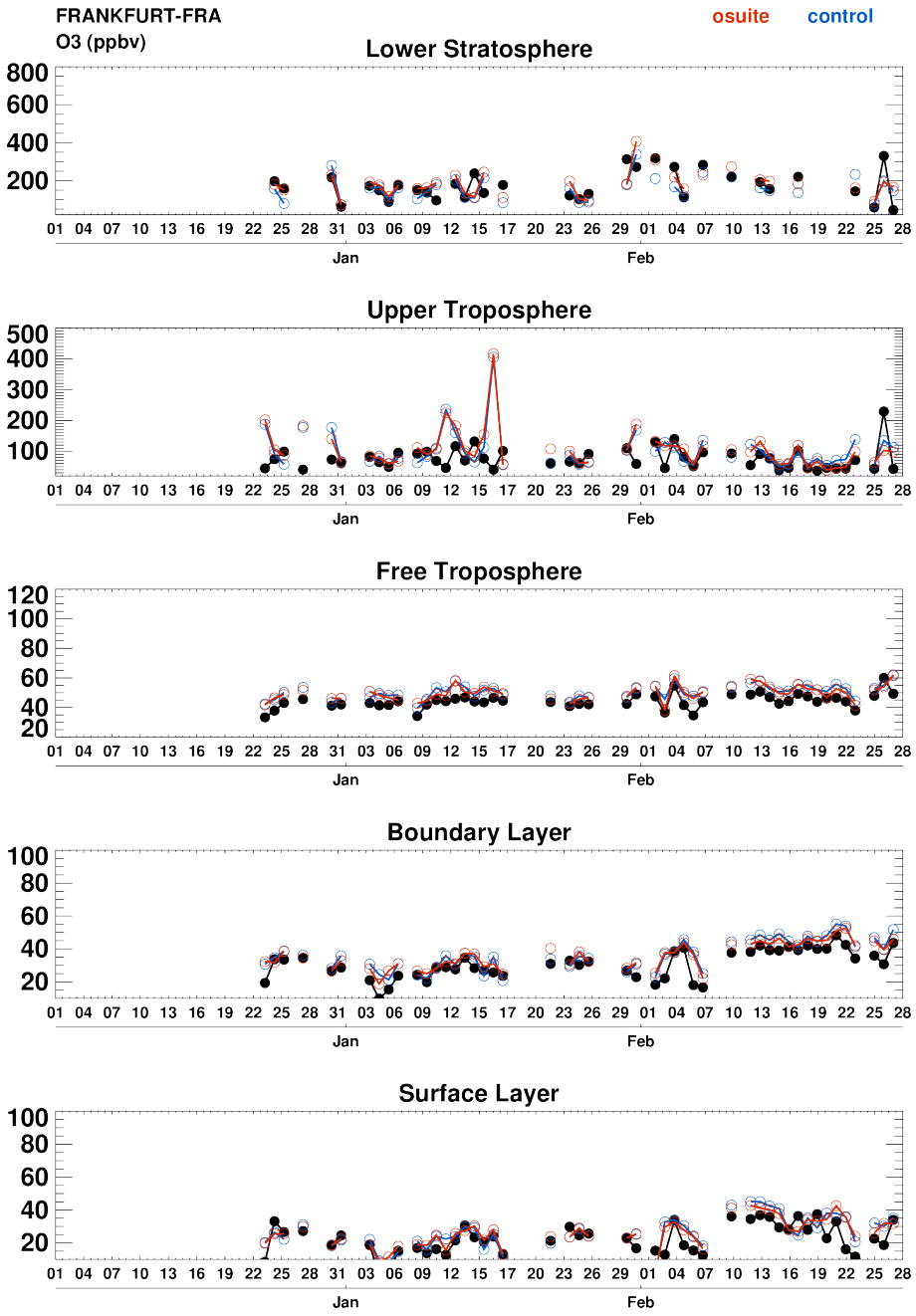


Figure 3.2.2: Time series of daily mean ozone over Frankfurt during DJF 2021 for 5 layers: Surface Layer, Boundary Layer, Free Troposphere, Upper Troposphere and Lower Stratosphere. IAGOS is shown in black, the o-suite in red and associated control run in blue. Units: ppbv.

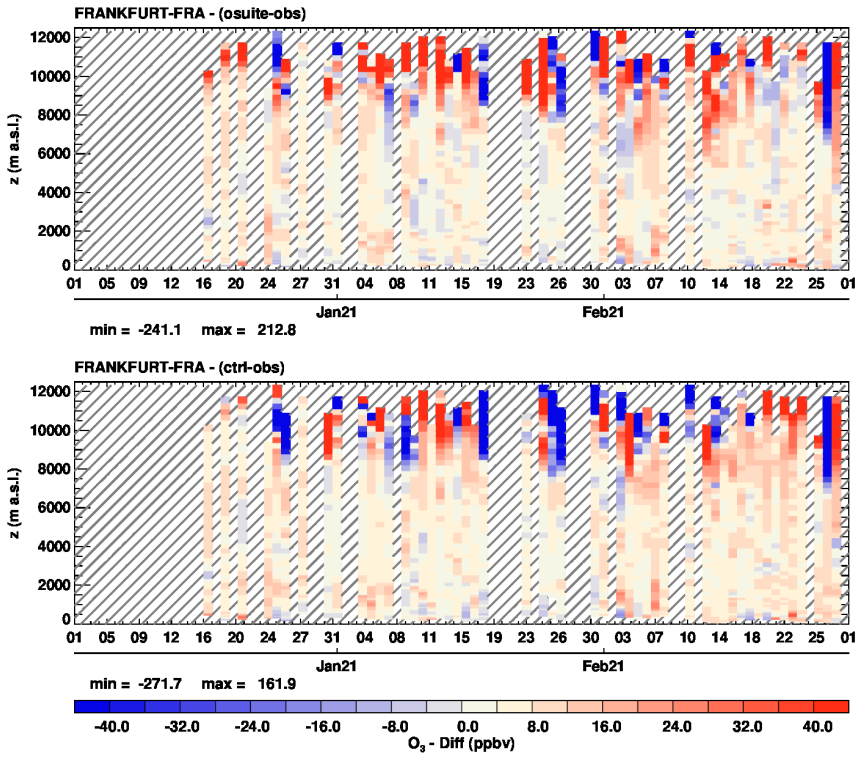


Figure 3.2.3: Time series of the absolute differences (model – IAGOS observations) in daily profiles for ozone over Frankfurt during DJF 2021. The top panel corresponds to o-suite the bottom panel to control run. Units: ppbv.

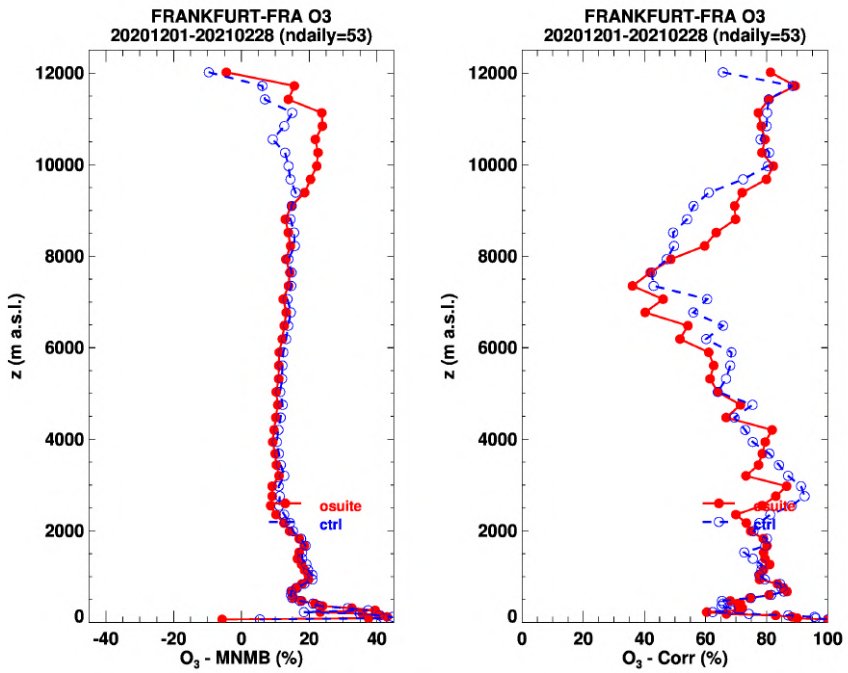


Figure 3.2.4: Model scores (MNMB and Correlation coefficient) for ozone at Frankfurt calculated over the period DJF 2021. The left panel corresponds to MNMB and the right panel to Correlation coefficient. The o-suite is shown in red and associated control run in blue. Units: %.

Six aircraft are equipped with the IAGOS system during this period. With these aircrafts, operating fully over the three-month period, we can expect a total of about 1260 flights. The actual number of flights within the period was 131 (262 profiles) giving a performance of 10%. These flights are shown in Fig. 3.2.1 (top). The low performance of this period is mainly due to a persistent reduction of air traffic and to maintenance related issues in the COVID-19 crisis context. In total for this DJF period, 55% of the operational flights had usable measurements of ozone and 53% of the flights had usable CO. Ozone and CO data are only delivered by the two aircrafts from Lufthansa operating from Frankfurt. Fig. 3.2.1 (bottom) shows the available airports, with a plotting circle scaled to the highest number of flights at an airport.

Comparison with profile data

Europe

Figure 3.2.2 presents ozone time series at Frankfurt during the full period December 2020 to February 2021 for 5 atmospheric layers. Time series of the profile differences (in ppbv) are also presented in Fig. 3.2.3. At Frankfurt, ozone data are available nearly continuously starting from mid-December until the end of February, with some gaps at the end of January and the second week of February.

In the low to mid-troposphere the two runs behave similarly. Ozone is well represented in the low troposphere by both runs (except close to the surface where sampling is poor), with on average a positive MNMB of less than 15% is found up to 2000 m and a correlation of about 80% (Fig. 3.2.3 and 3.2.4). In the mid troposphere, on average the MNMB from the models remains close to about 10% but the correlation coefficient is decreases to 40% at 8000 m (Fig. 3.2.3 and 3.2.4). Above 9000 m the bias is larger and the results from the two runs differ, ozone is mostly overestimated by the o-suite with an average MNMB of more than 25% (Fig. 3.2.3 and 3.2.4), whereas control run does not present a systematic behaviour with frequent underestimations (Fig. 3.2.3).

In December and January, according to the time series no notable event is found in the low troposphere for ozone at the Frankfurt airport (Fig. 3.2.2).

Compared to the 1991-2020 reference average, in Europe the month of February 2021 was substantially warmer over southern part and conversely colder over the northern (https://climate.copernicus.eu/surface-air-temperature-february-2021). However, some parts of Europe have met important variations of temperatures during the month of February 2021 with a change in the prevailing atmospheric circulation pattern (https://public.wmo.int/en/media/news/extreme-weather-hits-usa-europe). Germany have experienced a rapid and record change from a cold and icy first half of February to an unusual spring warmth (https://gcos.dwd.de/EN/press/press_release/EN/2021/20210226_the_weather_in_germany_in_february_2021.html?nn=732034). In the second week of February, monthly records were broken with temperature falling below -20°C at many locations, and heavy snow over low elevations. Within a few days, this was followed by unusually mild weather in central Europe. Large and record temperature difference were recorded in the space of less than a week according to DWD.

On the time series, in the low troposphere an increase in ozone is observed around 10 February from about 15 to nearly 40 ppbv in the surface layer which remain constant until the 22 of February

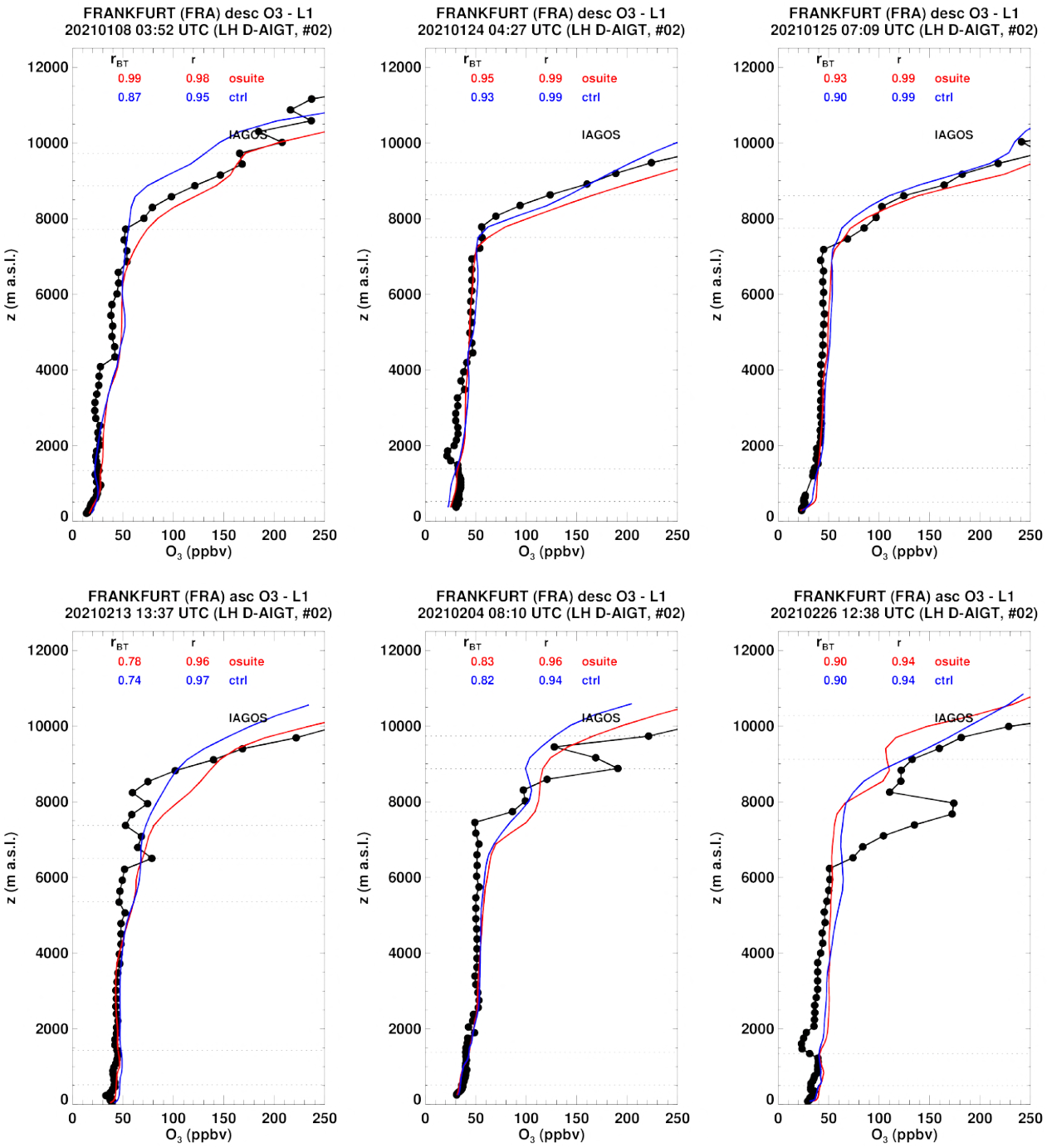


Figure 3.2.5.a: Selection of individual profiles for ozone from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over Frankfurt during DJF 2021. Units: ppbv.

where ozone values drop again to former levels (Fig. 3.2.2). This pattern is likely related to the increased formation of ozone with the warm and sunny weather mentioned before, which is well represented by the models with slight overestimations. The good performance of the models can also be seen in some of the profiles presented in Fig. 3.2.5.b.

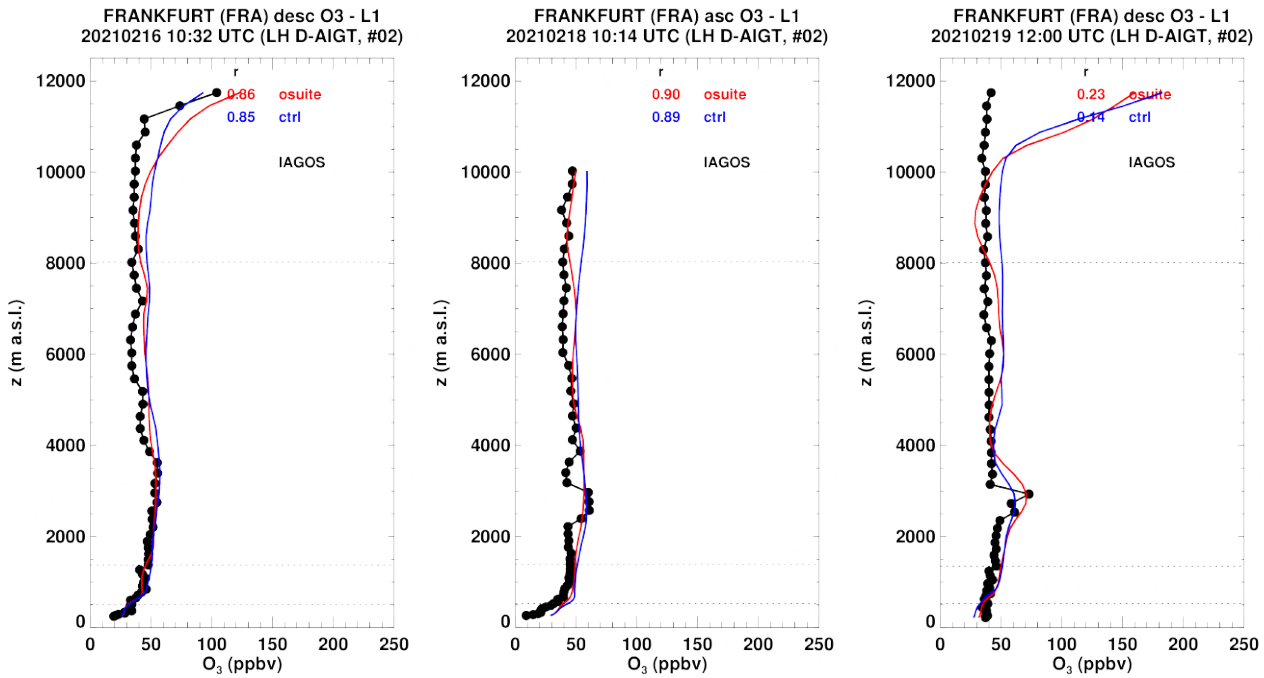


Figure 3.2.5.b: Selection of individual profiles for ozone from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over Frankfurt during DJF 2021. Units: ppbv.

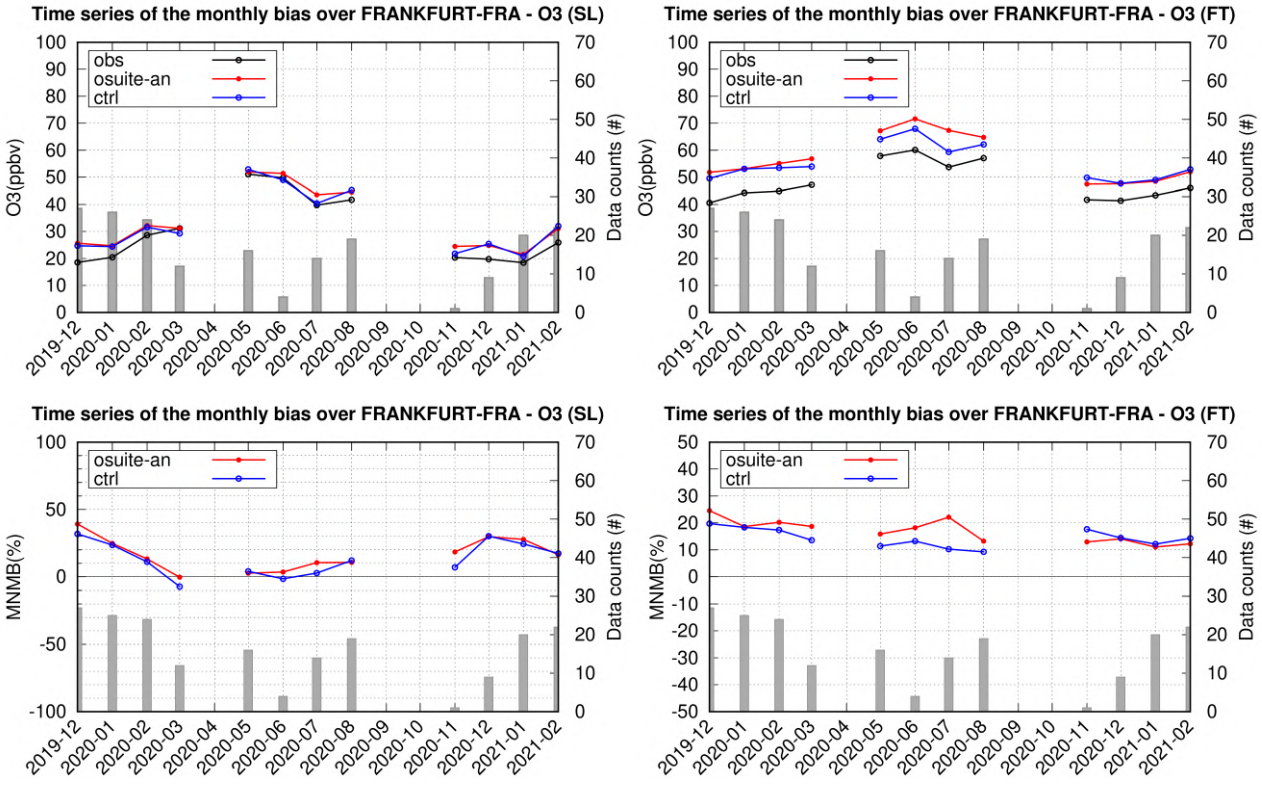


Figure 3.2.6: Top panels: monthly time series for ozone from IAGOS from o-suite analysis (red) and control run (blue) at Frankfurt in different atmospheric layers (left: surface layer, right, free troposphere) during the period December 2019 - February 2021. Bottom panels: corresponding MNMB time series for the two runs. The histogram bars indicate the number of available profiles per month.

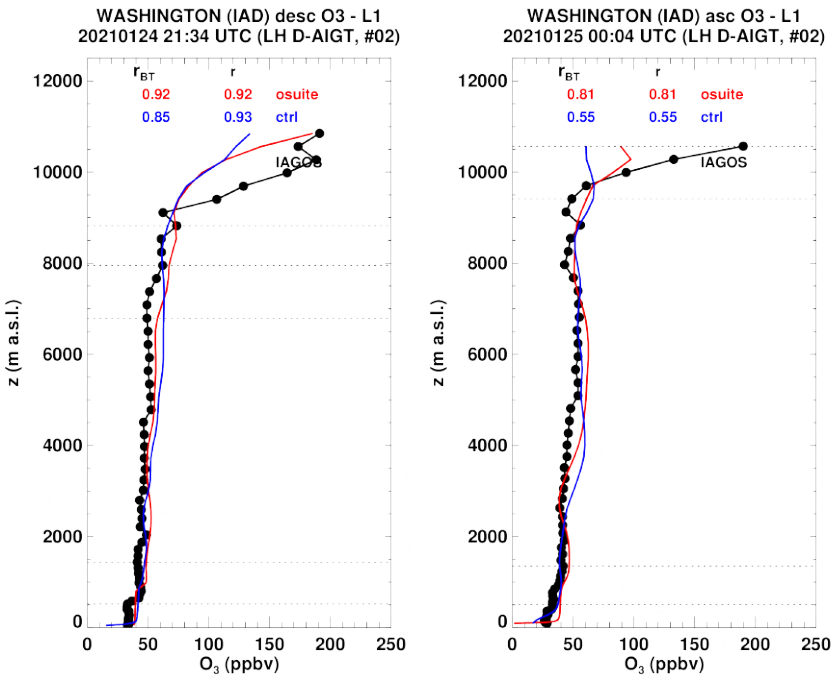


Figure 3.2.7: Selection of individual profiles for ozone from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over North America during DJF 2021. Units: ppbv.

Several events of tropopause descent are observed along the DJF period at altitudes below 8000 m. A few examples are shown in Fig. 3.2.5.a, and in most cases the low tropopause is well detected by the models.

To compare the model performance with the previous DJF period at Frankfurt, Fig. 3.2.6 presents the time series of the monthly values for surface and free tropospheric ozone together with the associated time series of the MNMB from the two runs for a period starting back from December 2019. In the surface layer, MNMB values for this DJF period are very similar to those observed in the previous DJF period for the two runs. In the free troposphere, MNMB values from the o-suite and control run in DJF 2021 appear slightly smaller compared with that of DJF 2020.

North America

Over North America ozone profiles are available only at Washington (Fig. 3.2.7). Theses profiles are nearly constant from the surface to the free troposphere with 50 ppbv. The behaviour of the two runs is very similar and ozone is in general well represented.

Eastern and Southeastern Asia

Over Eastern and Southeastern Asia, ozone profiles are available at the airports: Hong-Kong, Bangalore, and Singapore (Fig. 3.2.8). The two runs behave rather similarly in the lowest layers with mostly small overestimations. In the free troposphere remain similar with sometime a slightly better performance from the o-suite. In the UTLS the results from the two runs can differ and the agreement is usually worse.

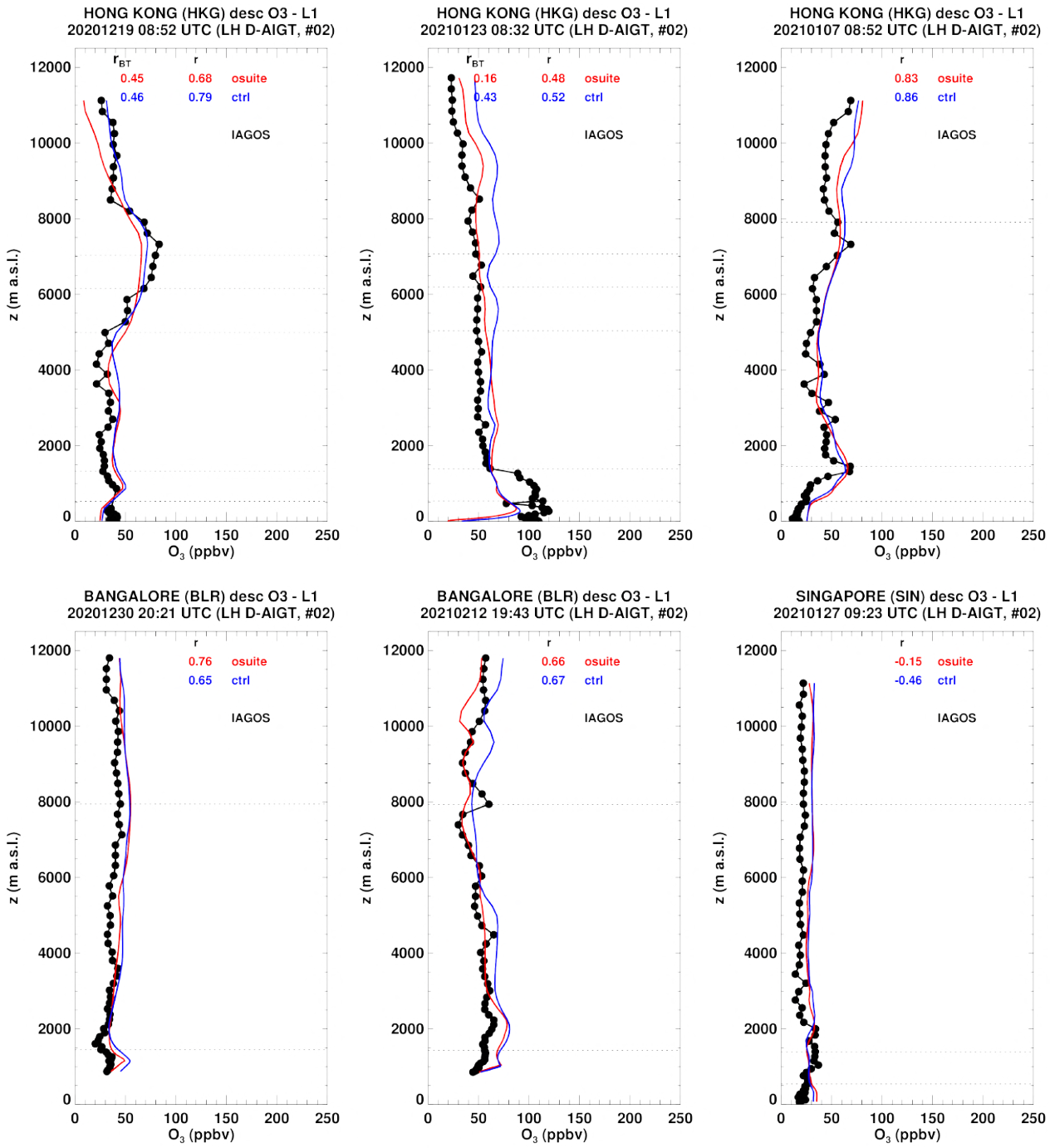


Figure 3.2.8: Selection of individual profiles for ozone from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over East and South-East Asia during DJF 2021. Units: ppbv.

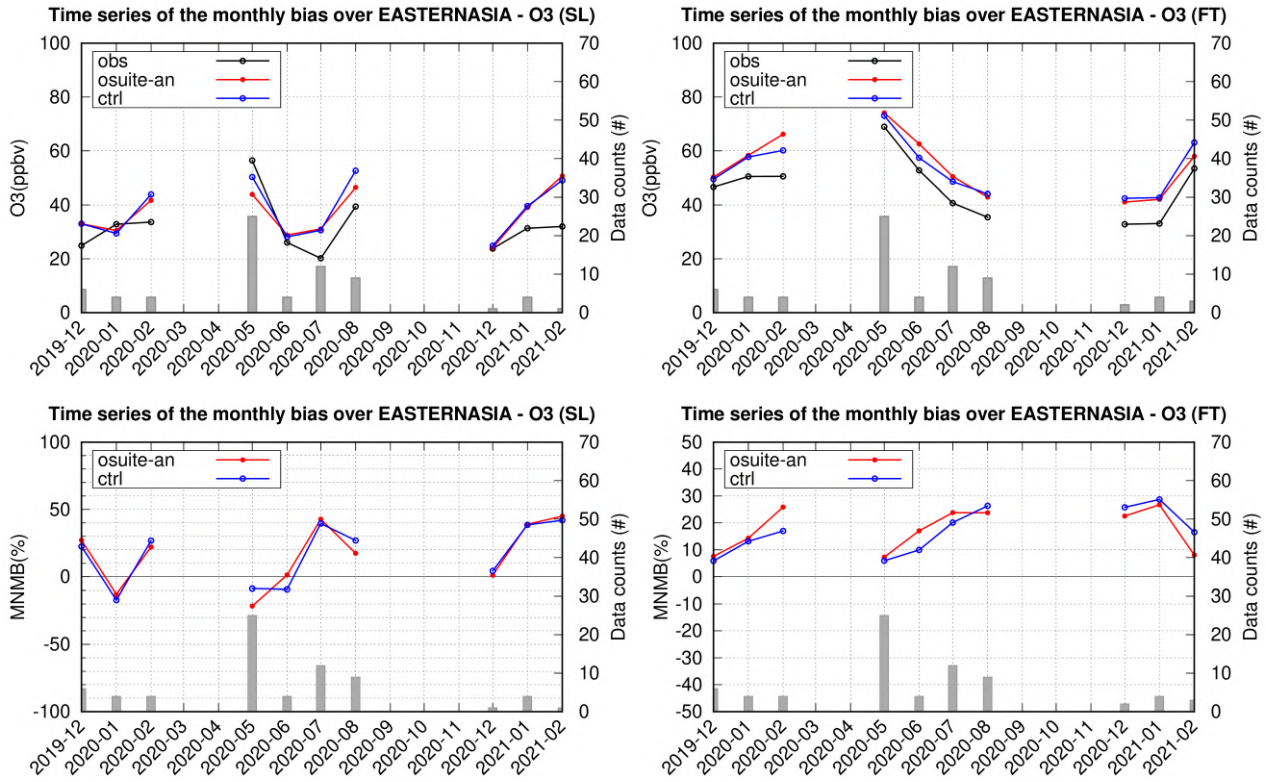


Figure 3.2.9: Top panels: monthly time series for ozone from IAGOS from o-suite analysis (red) and control run (blue) over Eastern and Southeastern Asia in different atmospheric layers (left: surface layer, right, free troposphere) during the period December 2019 - February 2021. Bottom panels: corresponding MNMB time series for the two runs. The histogram bars indicate the number of available observations.

Fig. 3.2.9 presents the time series of the monthly values for surface and free tropospheric ozone together with the associated time series of the MNMB from the two runs for a period starting back from December 2019. In the surface layer, likely due to poor sampling and different available airports, MNMB values do not show a clear and systematic behaviour for the two DJF period considered, therefore no conclusion can be drawn on the comparison between the two quarters. In the free troposphere, MNMB values from the o-suite and control run in DJF 2021 appear larger than those of DJF 2020, but again this could also be an artifact from poor and different sampling.

India

Several ozone profiles are available at Delhi during DJF 2021 as shown on some examples in Fig. 3.2.10. The results from both runs are similar for the profiles obtained at Delhi during this period. Small overestimations are found in the low to mid-troposphere with a slightly better performance from the o-suite in the lowest layers. In the UTLS the performance is not always in good agreement with observations which is conditioned by the good representation of the tropopause.

South Africa

A few profiles are available at Cap Town during this DJF period (Fig. 3.2.11). In these profiles the two runs behave similarly from the surface to the UTLS. Ozone is generally well represented from the lowest layers and free-troposphere, except for underestimated maxima in the mid-troposphere. In the UTLS ozone is generally overestimated.

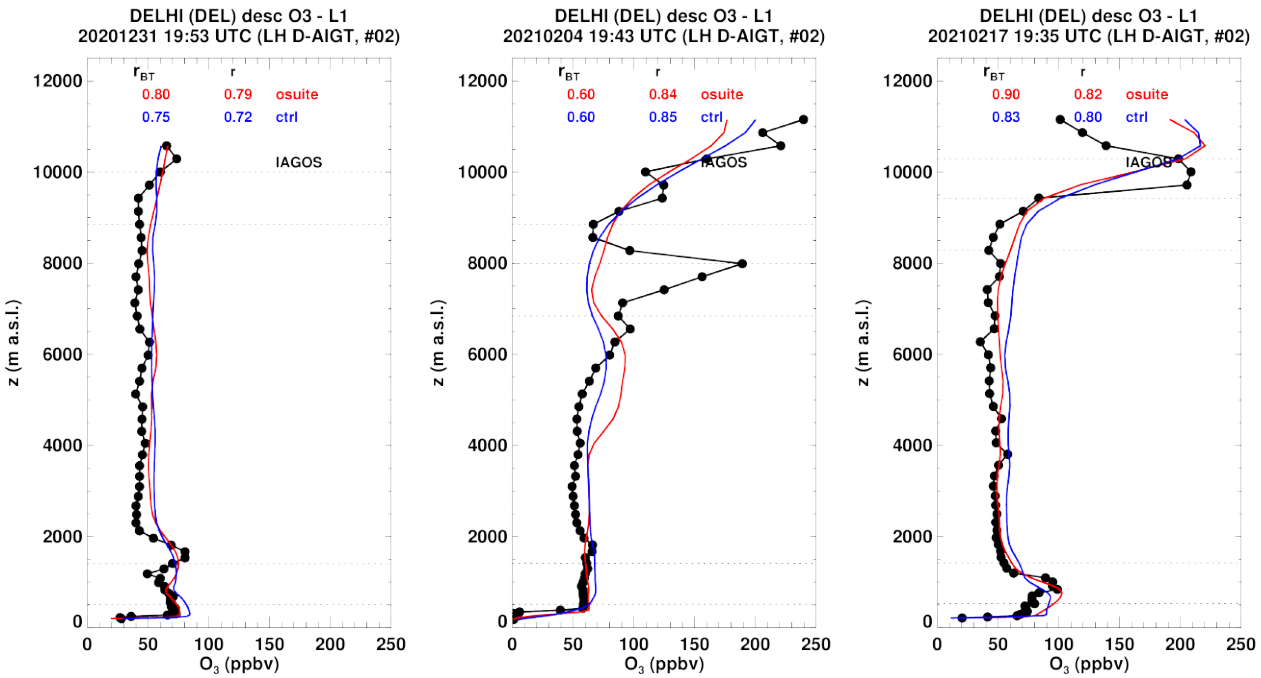


Figure 3.2.10: Selection of individual profiles for ozone from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over India during DJF 2021. Units: ppbv.

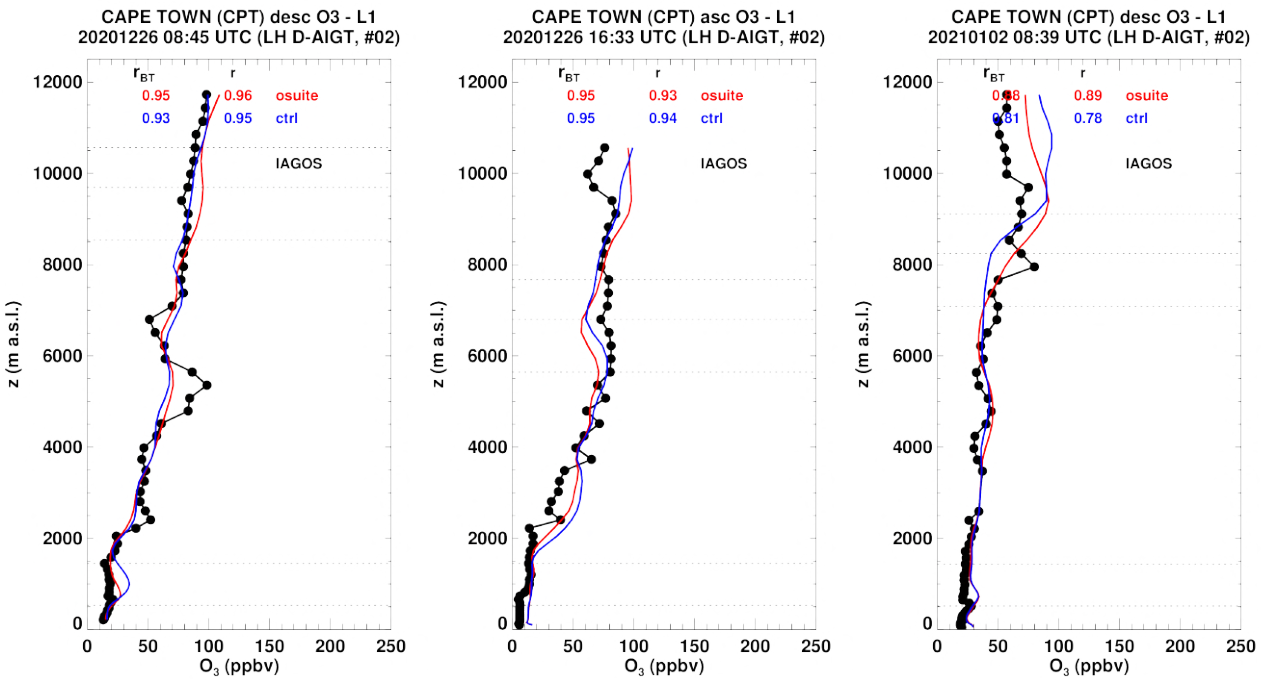


Figure 3.2.11: Selection of individual profiles for ozone from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over South Africa during DJF 2021. Units: ppbv.

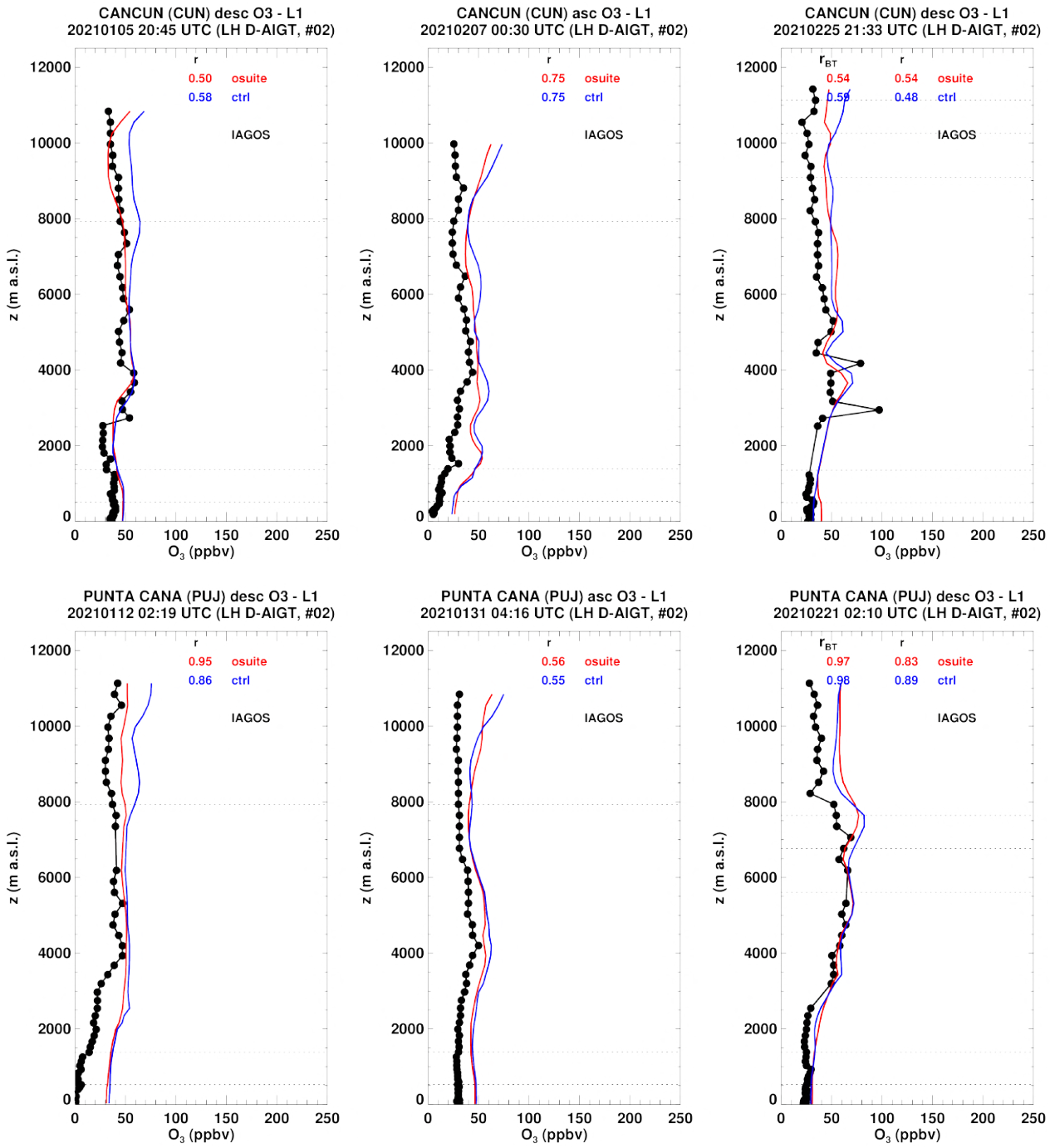


Figure 3.2.12.a: Selection of individual profiles for ozone from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over Central America and the Caribbean during DJF 2021. Units: ppbv.

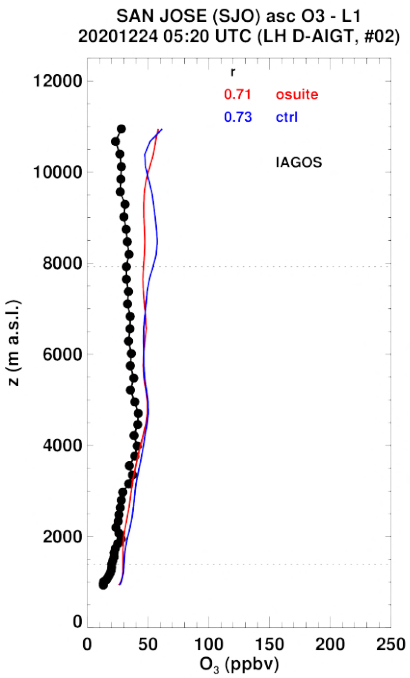


Figure 3.2.12.b: Selection of individual profiles for ozone from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over Central America and the Caribbean during DJF 2021. Units: ppbv.

Central America and the Caribbean

Ozone profiles are available over Central America at the airports of Cancun (Mexico) and San Jose (Costa Rica) and over the Caribbean at Punta Cana (Dominican Republic), with some examples provided in Fig. 3.2.12.a-b. The two runs present similar performance at all these locations. Low ozone values are overestimated in the low to mid-troposphere. The agreement is generally better in the upper part of the free troposphere. In the UTLS ozone is often overestimated by both models, although some the performance can differ between the two runs.

Comparison with cruise level data

Fig. 3.2.13 shows the gridded average maps for the comparisons with IAGOS observations of ozone at flight level. The spatial variability of ozone is well represented by both CAMS runs. The CAMS reanalysis mostly overestimates ozone with a relative bias in the range 50-100% over all regions. The smallest bias is found over western and southern Africa, as well as over India and Eastern Asia, with values within ±25%. The control run also mostly overestimates ozone, but shows regional differences in the bias with larger biases than the o-suite over the Atlantic Ocean, and with the largest values over the Northern part of the Atlantic. Over Africa the results of the control run are similar to those of the o-suite, while over India the o-suite performs better than the control run.

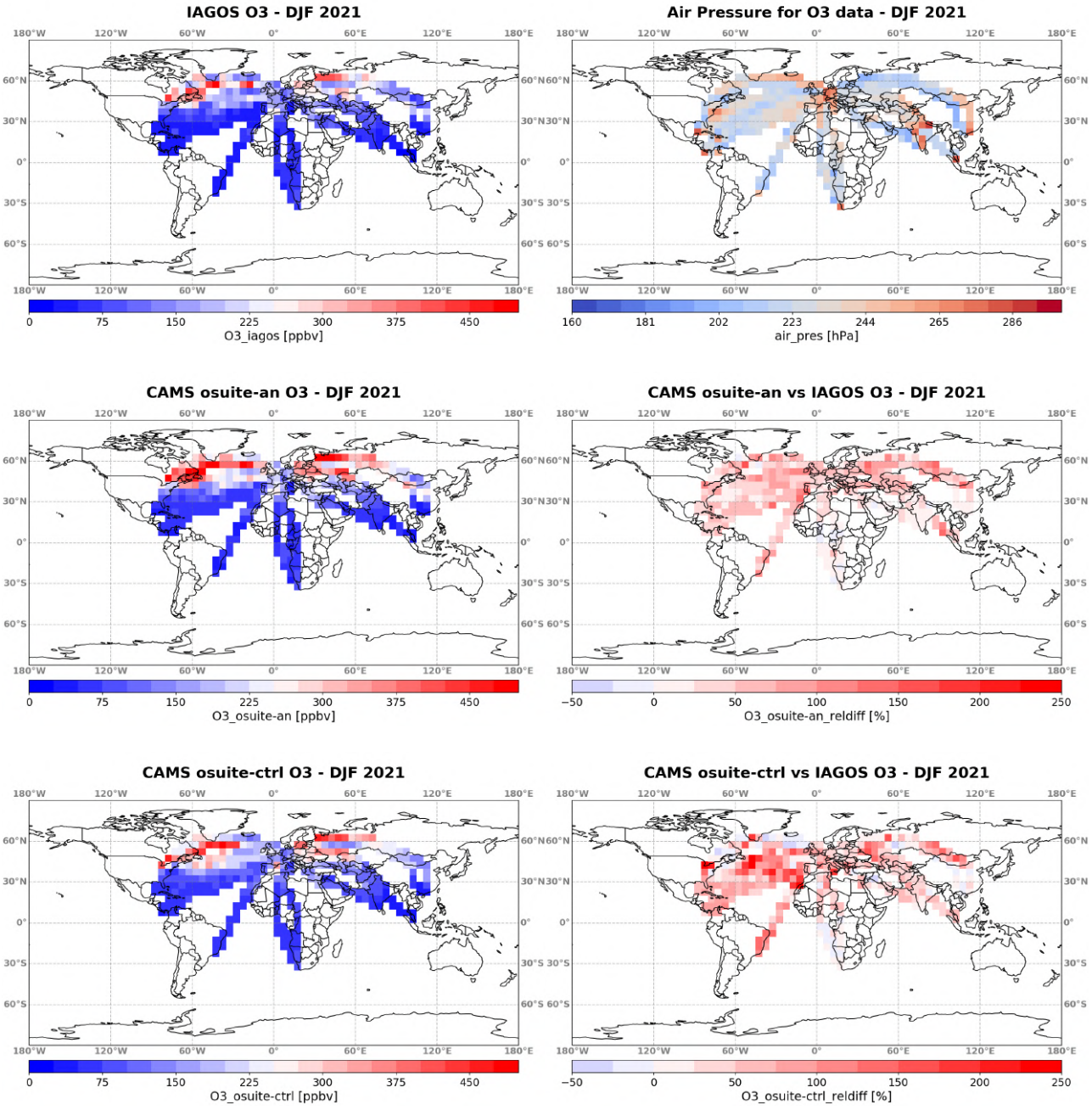


Figure 3.2.13: Global maps of gridded averages (5°x5°) for ozone comparison with IAGOS cruise data during DJF 2021. From left to right, first row: IAGOS ozone (in ppbv) and air pressure (in hPa). Second row: ozone from the analysis of the o-suite (in ppbv) and associated relative differences (in %) with respect to IAGOS. Third row: same as second row for control run.

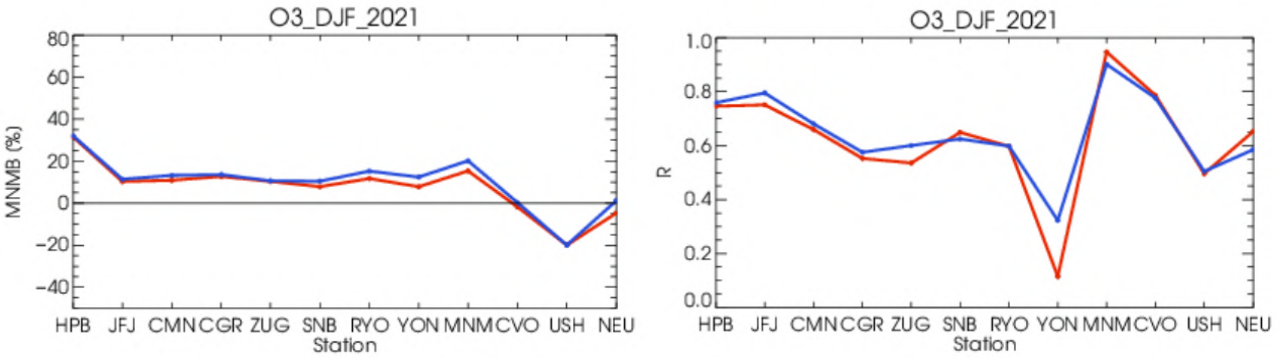


Figure 3.3.1: Modified normalized mean bias in % (left) and correlation coefficient (right) of the NRT forecast runs compared to observational GAW data in the period December 2020 - February 2021 (o-suite: red, control: blue).

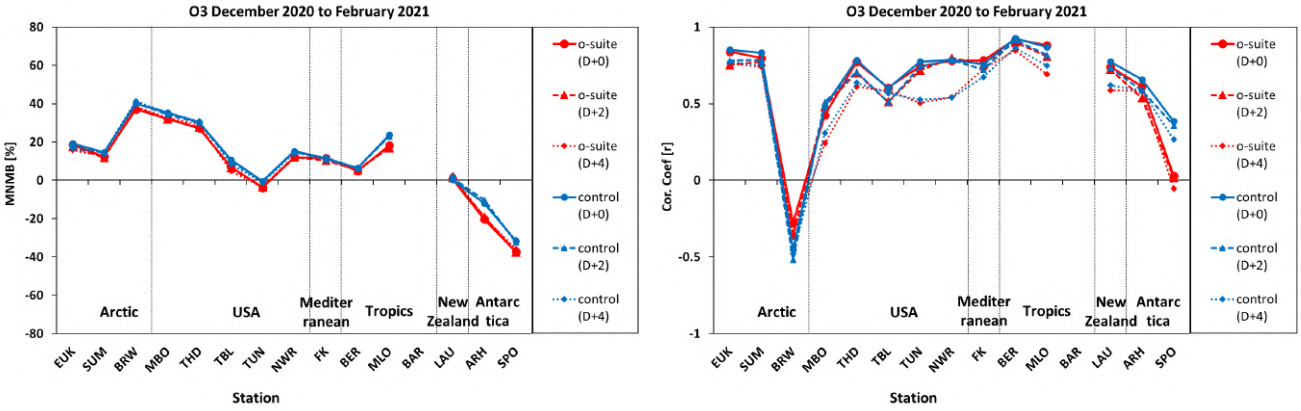


Figure 3.3.2: Modified normalized mean bias in % (left) and correlation coefficient (right) of the NRT forecast runs compared to observational ESRL data in the period December 2020 - January 2021. Circles correspond to D+0, triangles to D+2 and rhombus to D+4 metrics respectively.

3.3 Validation with GAW and ESRL-GMD surface observations

For the Near Real Time (NRT) validation, 13 GAW stations and 13 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 two NRT runs (o-suite and control) has been carried out for the period from December 2020 to February 2021. 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.

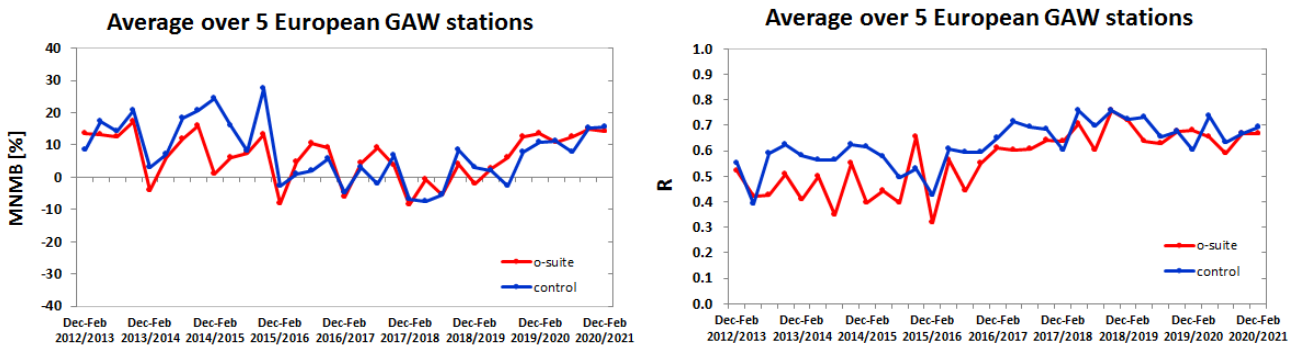


Figure 3.3.3: Long term (December 2012 – February 2021) evolution of seasonal mean MNMB (left) and correlation (right), as averaged over 5 GAW stations in Europe, for o-suite (red) and control (blue).

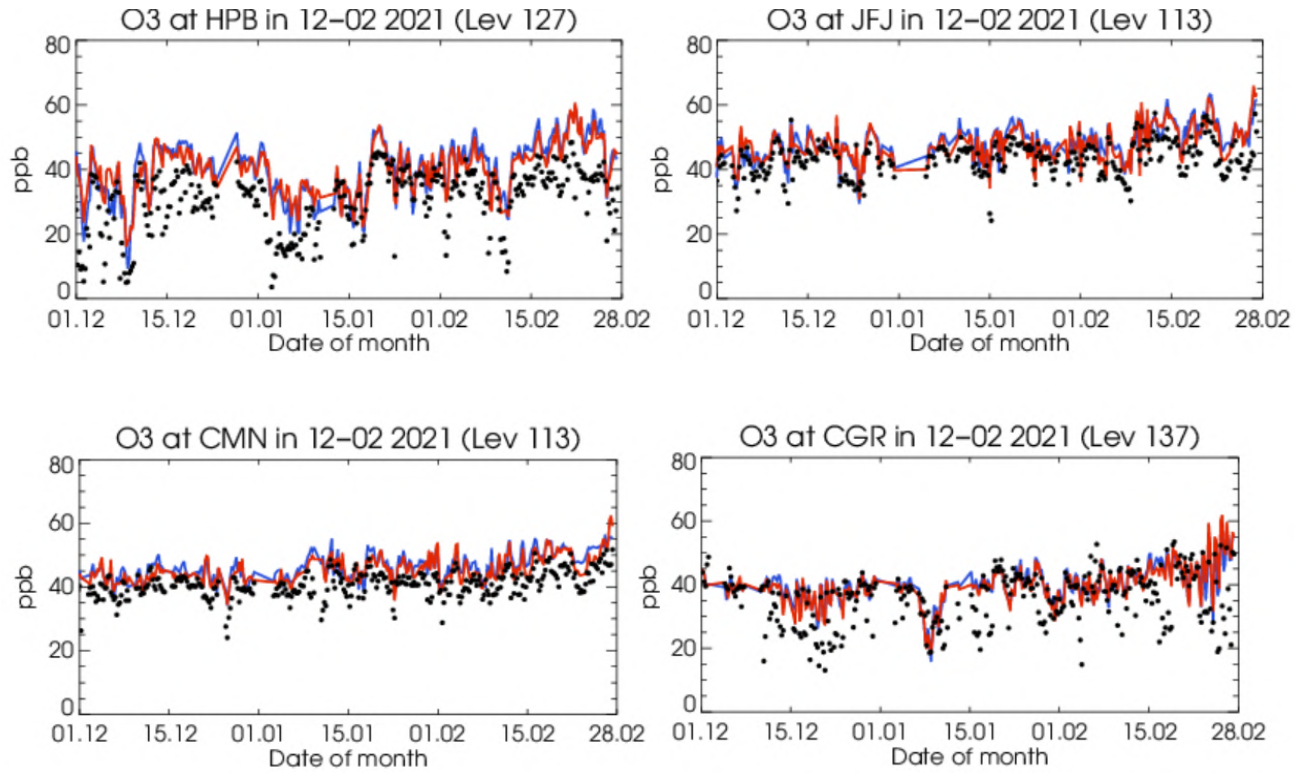


Figure 3.3.4: Time series for the o-suite (red) and control (blue) compared to GAW observations for Hohenpeissenberg (47.8°N, 11.02°E) and Jungfraujoch (46.55°N, 7.99°E) (upper panel), Monte Cimone (44.18°N, 10.70°E) and Capo Granitola (37.67°N, 12.65°E) (lower panel).

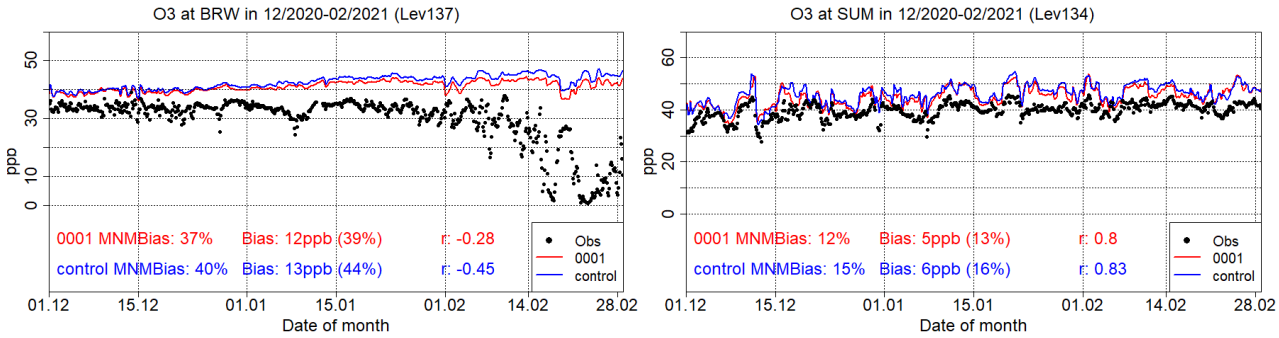


Figure 3.3.5: Time series for the o-suite (red) and control (blue) compared to ESRL observations at Point Barrow Alaska station (71.32°N, 156.47°W, left) and at Summit, Greenland station (72.57°N, 38.48°W, right). Period December 2020 – February 2021.

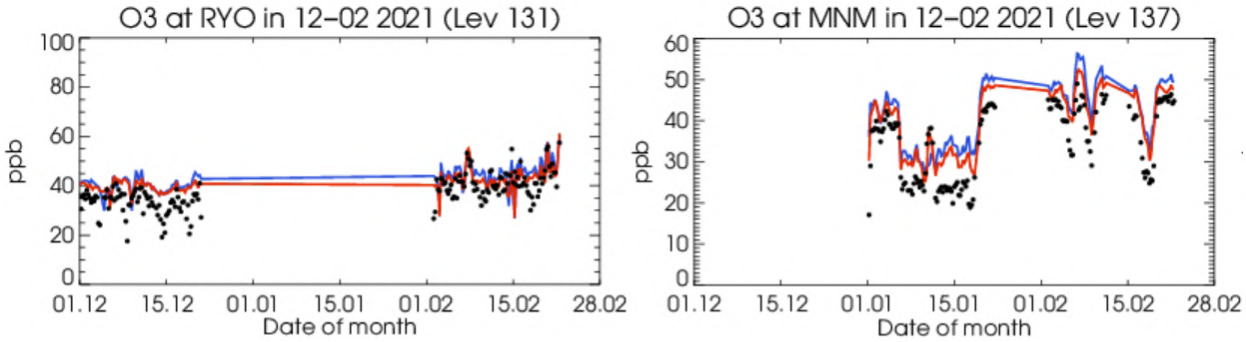


Figure 3.3.6: Time series for the o-suite (red) and control (blue) compared to GAW observations for Ryori (39.03°N, 141.82°E, left panel) and Minamitroishima (24.29°N, 153.98°E, right panel). Period December 2020 – February 2021.

A comparison of the seasonal-mean MNMB over Europe (Fig. 3.3.3) from December 2012 to present shows minimal MNMBs during the winter season and larger biases in other months before 2020. After the year 2020, the biases remain relatively stable, within 12-17 % (with is close to maximum), in all seasons. The MNMB for the o-suite and control runs shows, on average, an improvement over the years (before 2020). The temporal correlation is consistently better for the control run than for the o-suite, but the o-suite shows strong improvements. 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.

Looking at different regions, for European stations (HPB, JFJ, ZUG, SNB, CMN, CGR), observed O_3 surface mixing ratios, are overestimated showing MNMBs within 10%, except for HPB (bias is about 30%). Correlations for European stations are between 0.55 and 0.81 for the o-suite and between 0.58 and 0.82 for the control run, see Fig. 3.3.1.

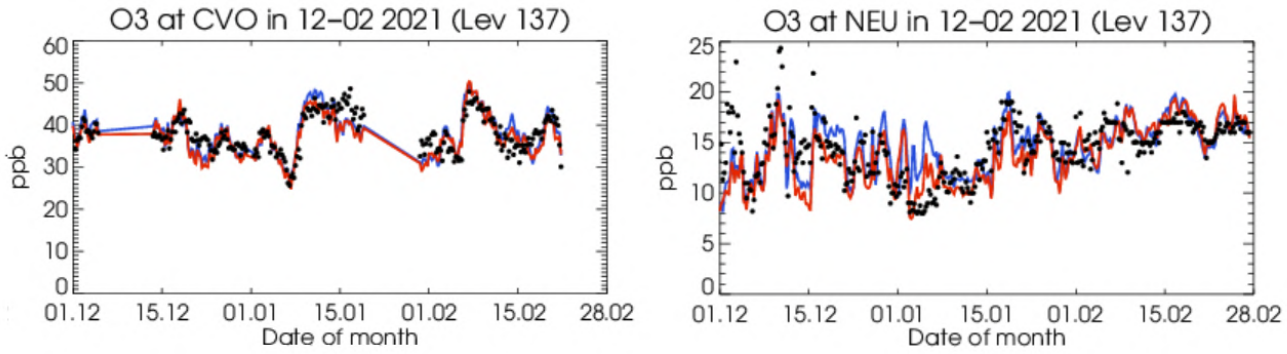


Figure 3.3.7: Time series for the o-suite (red) and control (blue) compared to GAW observations (black dots) at Cape Verde (16.85°N, 24.87°W) and Neumayer (70.65°S, 8.25°W). Period December 2020 – February 2021.

Over Eureka (EUK) and Summit (SUM) Arctic stations, both CAMS NRT runs overestimate surface ozone values by 20% and 15% respectively, while at Point Barrow Alaska station overestimate it more by 40%. Correlations between modelled and observed ozone values are highly significant at EUK and SUM stations, 0.83 and 0.8, respectively. On the contrary at Point Barrow station both CAMS NRT runs cannot reproduce the day-to-day surface ozone variability resulting to poor correlation (r=-0.3).

Concerning USA stations (MBO, THD, TBL, TUN and NWR), the observed ozone mixing ratios are overestimated by the o-suite run by 32% and 27% at MBO and THD respectively, and by 12% and 7% at NWR and TBL respectively, while at TUN station slightly underestimate surface ozone values by -3%. In all USA ESRL stations the control run MNMBs are 3-4% higher than o-suite MNMBs so the o-suite performed slightly better than the control run over all USA stations in terms of biases (exception is the TUN station). Correlations between o-suite and observations is 0.42 at MBO, 0.60, at TBL and 0.74-0.78 at TUN, THD and NWR stations. Correlation between the control run and observations are almost identical with the respective o-suite correlation (exception is MBO station where control run correlation is 0.47).

Modelled O_3 mixing ratios for Asian stations (RYO, MNM, YON) show MNMBs between 6 and 15%. Correlation coefficients for o-suite are high for MNM (0.95), about 0.6 for RYO and low for YON (0.1). (Fig. 3.3.1)

At CVO station, the model corresponds well to the observations with MNMB close to zero and correlation of 0.8 for both runs.

The O_3 mixing ratios of the Southern Hemispheric station (USH) show negative MNMBs of -20% for both, the o-suite and the control runs. Correlation coefficient is 0.5.

The observed ozone mixing ratios are overestimated by both runs at Mauna Loa (MLO) (MNMB o-suite≈18%, MNMB control≈23%) and Bermuda (BMD) (MNMB o-suite≈5%, MNMB control≈7%) stations in the Tropics. Correlations between simulated and observed surface ozone are high for both the o-suite and the control run (r>0.85).

At Lauder (LDR) station in New Zealand both runs have almost zero bias. Correlations between simulated and observed surface ozone values for the o-suite and the control run are 0.74 and 0.77, respectively.

Table 3.4.1: Coordinates, elevation, corresponding model level (level 137 is the surface level), as well as validation scores (MNMBs and correlations for the period DJF 2020-2021) obtained with the two 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	133	16	36.8	37.6	0.79	0.72
Caravaka	ES1882A	-1.87	38.12	1	137	73	NA	NA	NA	NA
Zarra	ES0012R	-1.10	39.08	885	130	70	19.1	21.2	0.78	0.79
VIillar Del Arzobispo	ES1671A	-0.83	39.71	430	137	48	5.1	7.4	0.74	0.75
Cirat	ES1689A	-0.47	40.05	466	137	37	31.4	33.2	0.65	0.65
Bujaraloz	ES1400A	-0.15	41.51	327	137	60	33.5	35.8	0.82	0.80
Morella	ES1441A	-0.09	40.64	1150	128	51	NA	NA	NA	NA
Bc-La Senia	ES1754A	0.29	40.64	428	137	21	-23.1	-20.9	0.79	0.77
Ay-Gandesa	ES1379A	0.44	41.06	368	136	15	24.7	26.7	0.68	0.63
Ak-Pardines	ES1310A	2.21	42.31	1226	135	81	27.7	29.5	0.59	0.57
Hospital Joan March	ES1827A	2.69	39.68	172	133	3	2.3	4.2	0.53	0.49
Al-Agullana	ES1201A	2.84	42.39	214	137	25	-0.4	3.2	0.50	0.47
Pobla	ES0296A	3.02	39.75	7	137	10	51.4	53.3	0.36	0.38
Av-Begur	ES1311A	3.21	41.96	200	132	9	16.3	19.8	0.77	0.73
Plan Aups/Ste Baume	FR03027	5.73	43.34	675	124	21	20.2	21.6	0.63	0.64
Montemonaco	IT1842A	13.34	42.90	1000	127	46	16.3	17.9	0.75	0.79
Gharb	MT00007	14.20	36.07	114	132	31	NA	NA	NA	NA
Aliartos	GR0001R	23.11	38.37	110	137	18	NA	NA	NA	NA
NEO	-	21.67	37.00	50	137	0.5	11.9	13.5	0.68	0.69
Finokalia	GR0002R	25.67	35.32	250	132	4	11.6	11.6	0.78	0.76
Agia Marina	CY0002R	33.06	35.04	532	133	14	3.0	5.2	0.60	0.59

Both CAMS NRT runs underestimate surface ozone values at Arrival Height (ARH) (MNMB o-suite≈-20%, MNMB control≈-12%) and South Pole (SPO) (MNMB o-suite≈-37%, MNMB control≈-32%) Antarctica stations. Correlation coefficients are 0.61 for the o-suite and 0.66 for the control run at ARH, and 0.02 for the o-suite and 0.36 for the control at SPO.

For Neumayer station (NEU) the MNMB is -5% for the o-suite and around 0% for the control run. Correlation coefficient is 0.65 for the o-suite and 0.60 for the control run (Fig. 3.3.1).

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. In addition, 1 station from the Department of Labour Inspection - Ministry of Labour and Social Insurance, of Cyprus () is 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 two 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 o-suite run reproduces slightly better than the control run the surface ozone day-to-day variability over most of the Mediterranean stations (see Table 3.4.1).

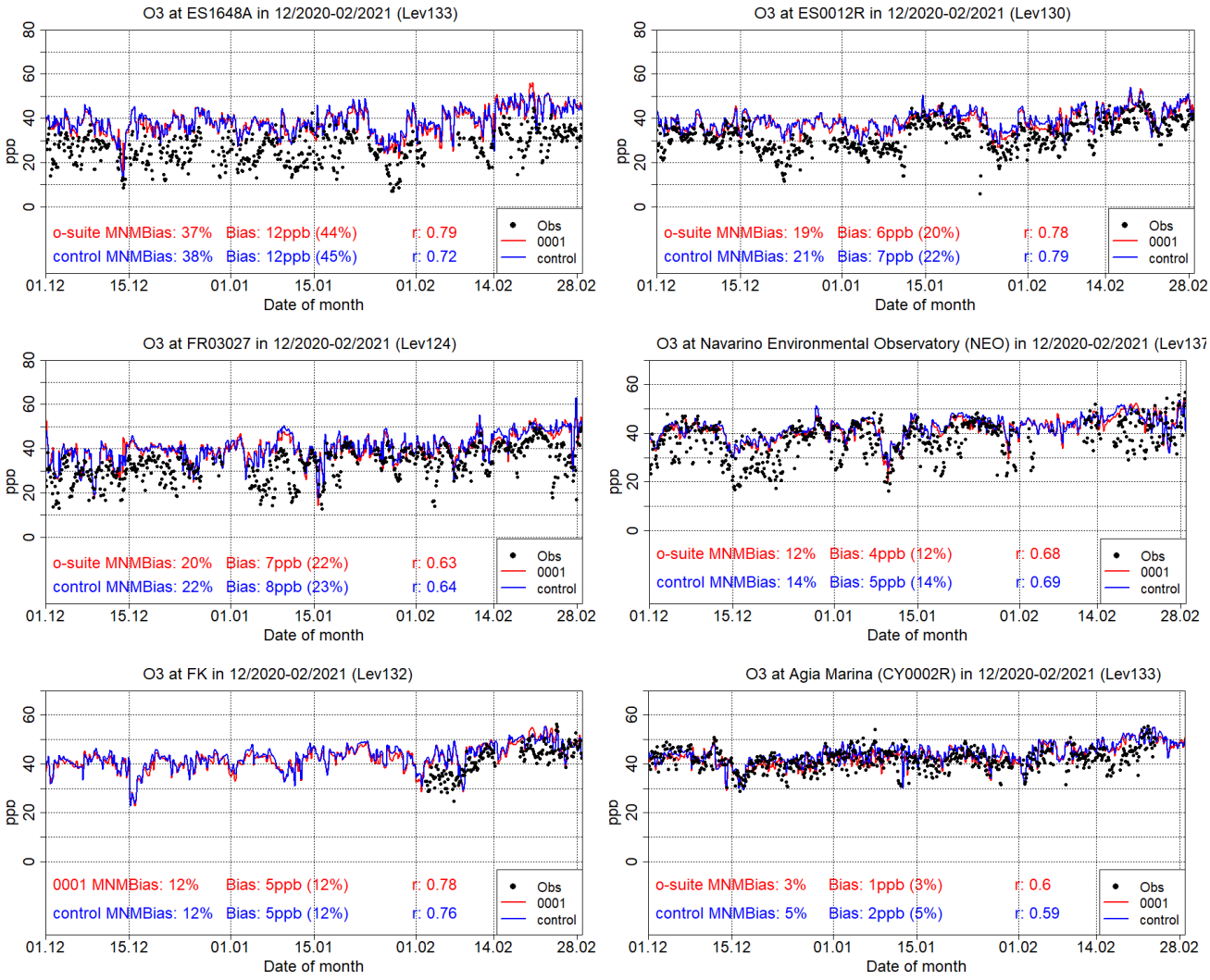


Figure 3.4.1: Time series for the o-suite (red) and Control (blue) compared to Airbase observations at Al Cornocales, Spain station (36.23°N, 5.66°W, top left), at Zarra, Spain station (39.08°N, 1.10°W, top right), at Plan Aups/Ste Baume, France station (43.34°N, 5.73°E, center left), at NEO, Methoni, Greece station (37.00°N, 21.70°E, center 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). Period December 2020 – February 2021.

In terms of bias, the o-suite mostly overestimates surface ozone values (except for Bc-La Senia station) and its MNMBs vary between -23% and 51% depending on the stations over the Mediterranean shore of Spain (average MNMB for the 11 Spain Mediterranean stations is 18.7%). The Control MNMBs are on average 2% higher than those of the o-suite – The o-suite performed slightly better than the control run in terms of biases. Over the stations Plan Aups/Ste Baume in France and Montemonaco in Italy the o-suite overestimates surface ozone concentrations by 20% and 16% respectively. Again, the Control MNMBs are higher by about 1.5% than the o-suite MNMBs. Over NEO station in Methoni Greece the o-suite overestimates surface ozone concentration by 12% and control run overestimates it by about more than 1%.

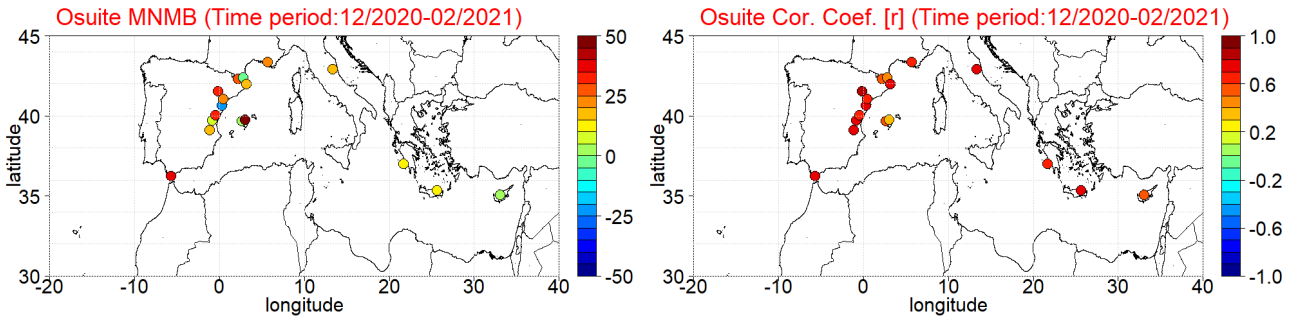


Figure 3.4.2: Spatial distribution of MNMB in % (left) and correlation coefficient (right) of the o-suite run compared to observational data during the period from 1 December 2020 to 28 February 2021.

Finally, over Finokalia station in Crete both runs overestimate surface ozone values by 11.6% and over Agia Marina in Cyprus these two percentages for o-suite and control run are 3% and 5%, respectively. In general o-suite performed slightly better than the control run in terms of biases over the majority of Mediterranean stations

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 run overestimate surface ozone values over the entire Mediterranean from Gibraltar to Cyprus and that CAMS NRT have a better performance over Central and Eastern Mediterranean compared to the Mediterranean shore of Spain in terms of biases.

3.5 Validation with AirBase observations over Europe

The surface ozone validation analysis over Europe is based on an evaluation against background rural Classes 1-2 O_3 Joly-Peuch classification (Joly and Peuch, 2012) station observations from the Airbase Network, .

The spatial distribution of MNMBs and the correlation coefficients of the o-suite over Europe are shown in Fig. 3.5.1, where it is evident that correlations over most European AirBase stations in the entire Europe (with a very few exceptions) are highly significant ($0.6<r<0.95$). CAMS NRT runs mostly overestimate surface ozone values with MNMBs varying from -40% to +60% over Europe depending on the station. More specifically, CAMS NRT surface ozone MNMBs over UK and Spain vary from -15% to +50%, over Great Britain from -15% to +25%. Over Ireland, France, and Italy o-suite overestimate surface ozone values between +5% and +25% while over Belgium mostly underestimate it (MNMBs between -25% and +5%). Over Swiss and Austria o-suite MNMbs vary between -25% and +30% while over Germany and the Czech Republic there is a stronger positive tendency in biases with MNMBs vary depending on the station between -20% and +40% in Germany and between -30% and +45% in the Czech Republic. Over Poland MNMBs vary between -15% and 25%. In Scandinavia o-suite overestimate surface ozone values between 3% and 50% over Norway, 10% and 20% over Sweden and -10% and 15% over Finland. In the Baltics o-suite underestimate surface ozone by -10% over Lithuania and overestimate it by 3%-7% over Estonia and by 10%-25% over Latvia. Finally, the o-suite MBMs are 10% over Serbia and 15% over Romania.

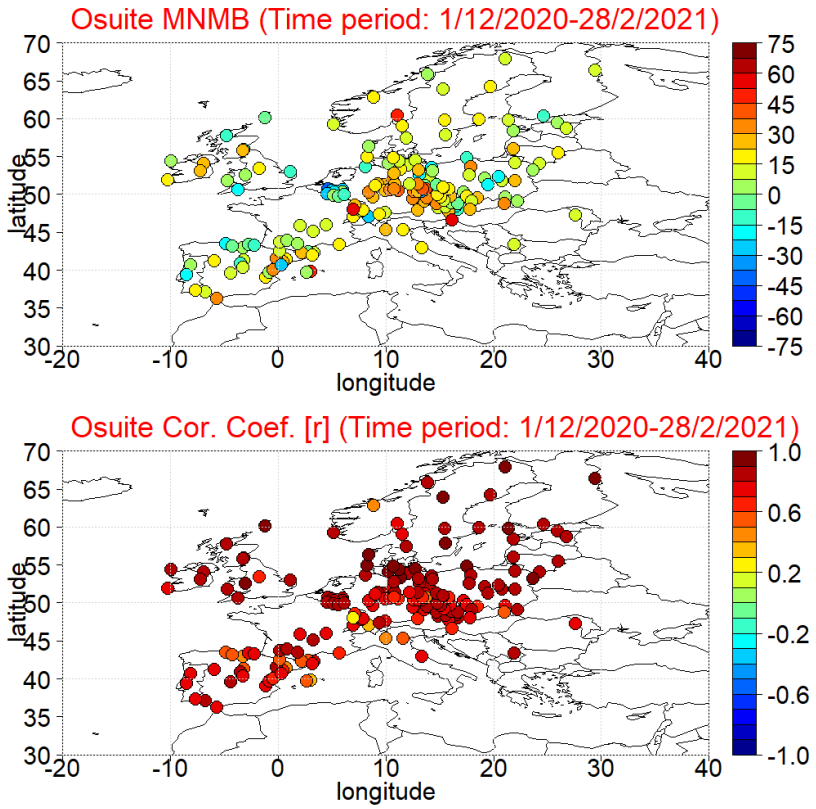


Figure 3.5.1: Spatial distribution of MNMB in % (top) and correlation coefficient (bottom) of the o-suite run compared to observational data during the period from 1 December 2020 to 28 February 2021.

The above-mentioned findings concerning the CAMS biases and correlations are also observed in individual time series at selected stations plotted in Figure 3.5.2. The time series and the plotted validation metrics show that the control run surface ozone mean DJF concentrations are about 1 ppb higher than the o-suite, leading to 1-2% higher MNMBs. The correlation between observed and modelled values for o-suite and the control run are almost identical.

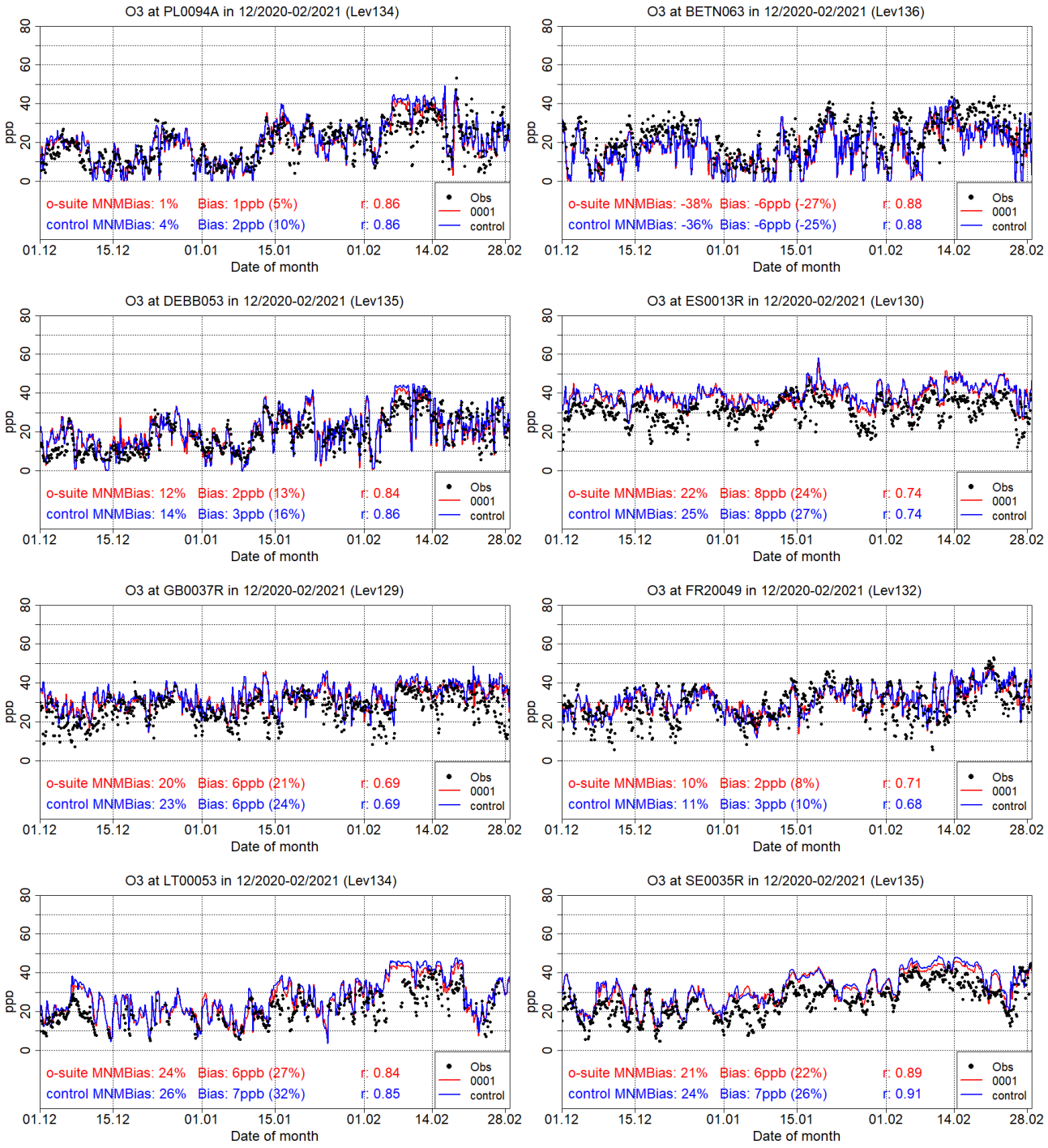


Figure 3.5.2: Time series for the o-suite (red) and Control (blue) compared to Airbase observations at Al Penausende, Spain station (41.24°N, 5.90 °W, 1st row left), at Haut Beaujolais, France station (45.96°N, 4.47°E, 1st row right), at Corroy L.G., Belgium Station (50.67°N, 4.67°E, 2nd row left), at Hasenholz, Germany (52.56°N, 14.02°E, 2nd row right), at Ladybower, Great Britain station (53.40°N, 1.75°W, 3nd row left), at LdGajewWIOSAGajew, Poland station (52.14°N, 19.23°E 3nd row right), at Zemaitija, Lithuania station (56.01°N, 21.89°E, 4nd row left) and at Vindeln, Sweden station (64.25°N, 19.77°E, 4nd row right). Period December 2020 – February 2021.

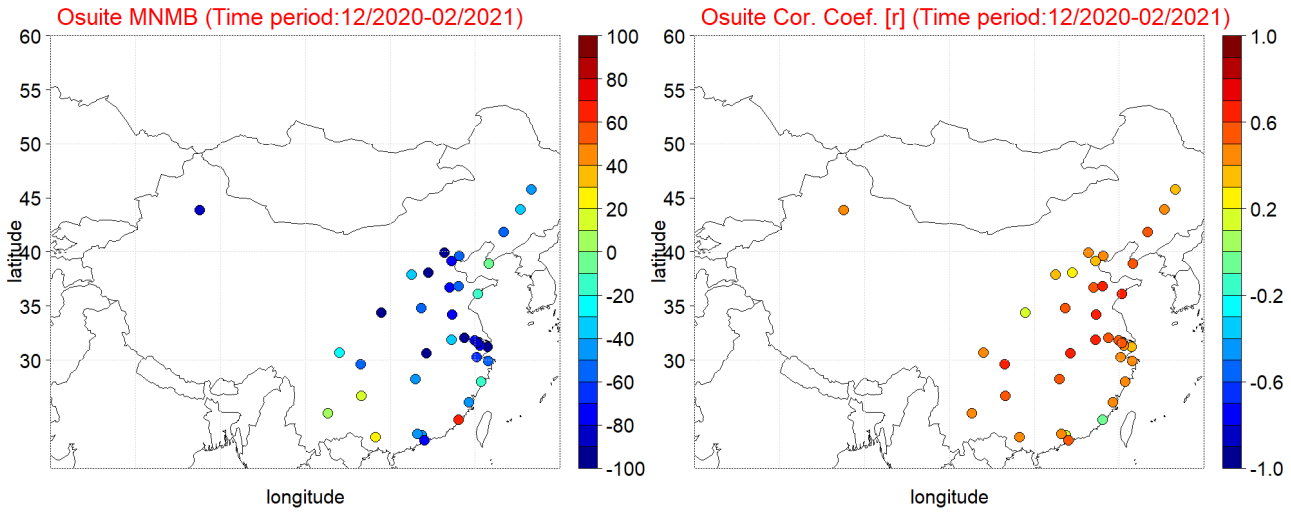


Figure 3.6.1: Spatial distribution of MNMB in % (left) and correlation coefficient (right) of the o-suite run compared to observational data during the period from December 2020 to February 2021.

3.6 Validation with surface ozone observations over China

The surface ozone validation over China is based on station observations from more than 1,500 in situ stations covering all major cities in China, operated by the China National Environmental Monitoring Center, reporting the pollutants PM10, PM2.5, O_3, NO_2, SO_2, and CO (e.g., Bai et al., 2020). The measurements were collected within the EU MarcoPolo and Panda projects. Individual station observations were clustered for 37 megacities (e.g., 10-20 stations per city) and the observed surface ozone values are compared with the model ozone values for the corresponding grid point.

Table 3.6.1 shows the names, coordinates, observed and simulated ozone values as well as validation metrics namely the MNMBs and correlations obtained with the o-suite run. The spatial distribution of MNMBs and the correlation coefficients of the o-suite over China are shown in Fig. 3.6.1, where it is evident that correlations over most stations in the entire China (with the exceptions of Xiamen and Xian Megacities) are highly significant ($0.3<r<0.7$). In terms of biases, the o-suite mostly underestimate surface ozone values. The o-suite underestimates surface ozone values (MNMBs vary between -50% and -35% depending on the Megacity) in North-western China and underestimate it in the Urumqi megacity in the North-East (MNMB is -80%). For stations in the latitudinal belt 30°N-40°N, the o-suite strongly underestimate surface ozone values with MNMBs varying over most megacities between -85% and -20%. There are megacities in this latitudinal belt (namely Beijing, Nanjing, Shanghai, Shijiazhuang, Wuhan, Xi'an) where MNMBs exceed 95% up to 145%. For megacities in the latitudinal belt 20°N-30°N, the o-suite mostly underestimates surface ozone and MNMBs vary between -80% and -15% (exceptions are Kunming, Guiyang, and Nanning where o-suite overestimate surface ozone values; MNMBs are 5%, 15% and 24% respectively). These findings concerning CAMS o-suite biases and correlations are also observed in individual time series at selected cities plotted in Figure 3.6.2. From all the time series it is also evident that the o-suite run reproduced well the observed upward surface ozone trend during DJF 2020-2021 period.

Table 3.6.1: Names, coordinates, observed and simulated ozone values as well as validation metrics (MNMBs and correlations for the period DJF 2020-2021) obtained with the 2 forecast runs (o-suite and control), for each one from 37 China Megacities under study. MNMBs and correlations with blue denote stations where control run performs better while with red are denoted stations where o-suite performs better.

			Ozone (ppb)			MNMB (%)		Cor. Coef.	
MegaCity	Lat	Lon	Observed	o-suite	contol	o-suite	contol	o-suite	contol
Beijing	39.92	116.38	18.7	6.5	6.8	-95.0	-91.2	0.40	0.42
Changchun	43.89	125.33	20.7	14.7	15.7	-35.0	-28.8	0.43	0.44
Changsha	28.20	112.97	16.9	11.5	11.8	-46.4	-45.4	0.55	0.53
Changzhou	31.81	119.97	21.4	9.3	9.8	-77.7	-72.3	0.56	0.58
Chengdu	30.66	104.07	16.6	12.9	14.8	-25.3	-14.1	0.44	0.44
Chongqing	29.56	106.55	12.0	7.9	8.4	-50.1	-44.8	0.69	0.67
Dalian	38.91	121.60	25.4	25.5	26.3	-1.6	0.5	0.54	0.61
Dongguan	23.02	113.75	31.8	17.6	18.0	-46.1	-44.2	0.14	0.13
Fuzhou	26.08	119.31	30.9	20.5	20.8	-45.3	-43.8	0.46	0.45
Guangzhou	23.13	113.25	28.0	16.7	17.0	-43.8	-42.3	0.43	0.40
Guiyang	26.65	106.63	21.9	24.8	26.6	15.3	22.8	0.53	0.45
Hangzhou	30.25	120.17	17.9	8.5	8.9	-67.0	-62.9	0.42	0.44
Harbin	45.75	126.63	25.6	14.6	15.7	-49.7	-43.6	0.36	0.38
Hefei	31.85	117.27	17.1	11.5	11.9	-32.7	-28.7	0.61	0.63
Jinan	36.67	116.98	22.5	8.6	8.9	-79.8	-77.1	0.51	0.48
Kunming	25.04	102.71	25.3	24.2	25.3	4.7	9.1	0.44	0.43
Nanjing	32.05	118.77	21.2	7.3	7.7	-97.5	-91.1	0.56	0.56
Nanning	22.82	108.32	27.2	32.9	33.8	23.6	25.1	0.42	0.42
Ningbo	29.87	121.54	23.3	12.7	13.5	-51.4	-45.9	0.44	0.43
Qingdao	36.07	120.38	23.8	21.0	22.3	-13.7	-8.5	0.64	0.66
Shanghai	31.22	121.47	23.9	4.4	4.6	-136.1	-134.2	0.39	0.42
Shenyang	41.80	123.40	18.0	11.3	12.1	-51.1	-44.6	0.50	0.52
Shenzhen	22.54	114.06	31.0	13.9	14.2	-78.7	-76.5	0.52	0.52
Shijiazhuang	38.04	114.51	20.3	5.5	5.6	-102.9	-103.7	0.27	0.30
Suzhou	31.30	120.60	21.4	8.9	9.5	-75.0	-69.7	0.41	0.42
Taiyuan	37.87	112.55	19.7	12.4	12.5	-37.5	-37.7	0.30	0.34
Tangshan	39.63	118.18	15.2	9.0	9.7	-51.0	-48.0	0.47	0.49
Tianjin	39.13	117.25	16.3	7.0	7.4	-70.1	-67.0	0.36	0.36
Urumqi	43.83	87.62	16.4	7.6	7.9	-80.5	-77.8	0.50	0.52
Wenzhou	27.99	120.70	22.4	17.0	17.7	-17.2	-12.8	0.42	0.45
Wuhan	30.58	114.30	17.7	5.7	5.7	-111.3	-112.1	0.62	0.63
Wuxi	31.57	120.33	20.9	8.0	8.5	-84.5	-78.4	0.54	0.56
Xiamen	24.48	118.09	12.1	23.3	23.5	64.9	65.8	-0.03	-0.03
Xi'an	34.34	108.94	35.6	6.3	6.2	-142.6	-144.0	0.18	0.18
Xuzhou	34.21	117.28	19.6	8.7	9.3	-72.8	-67.5	0.66	0.65
Zhengzhou	34.76	113.65	19.5	10.6	10.6	-55.7	-57.4	0.55	0.62
Zibo	36.78	118.05	21.5	11.2	11.8	-59.8	-56.7	0.62	0.61

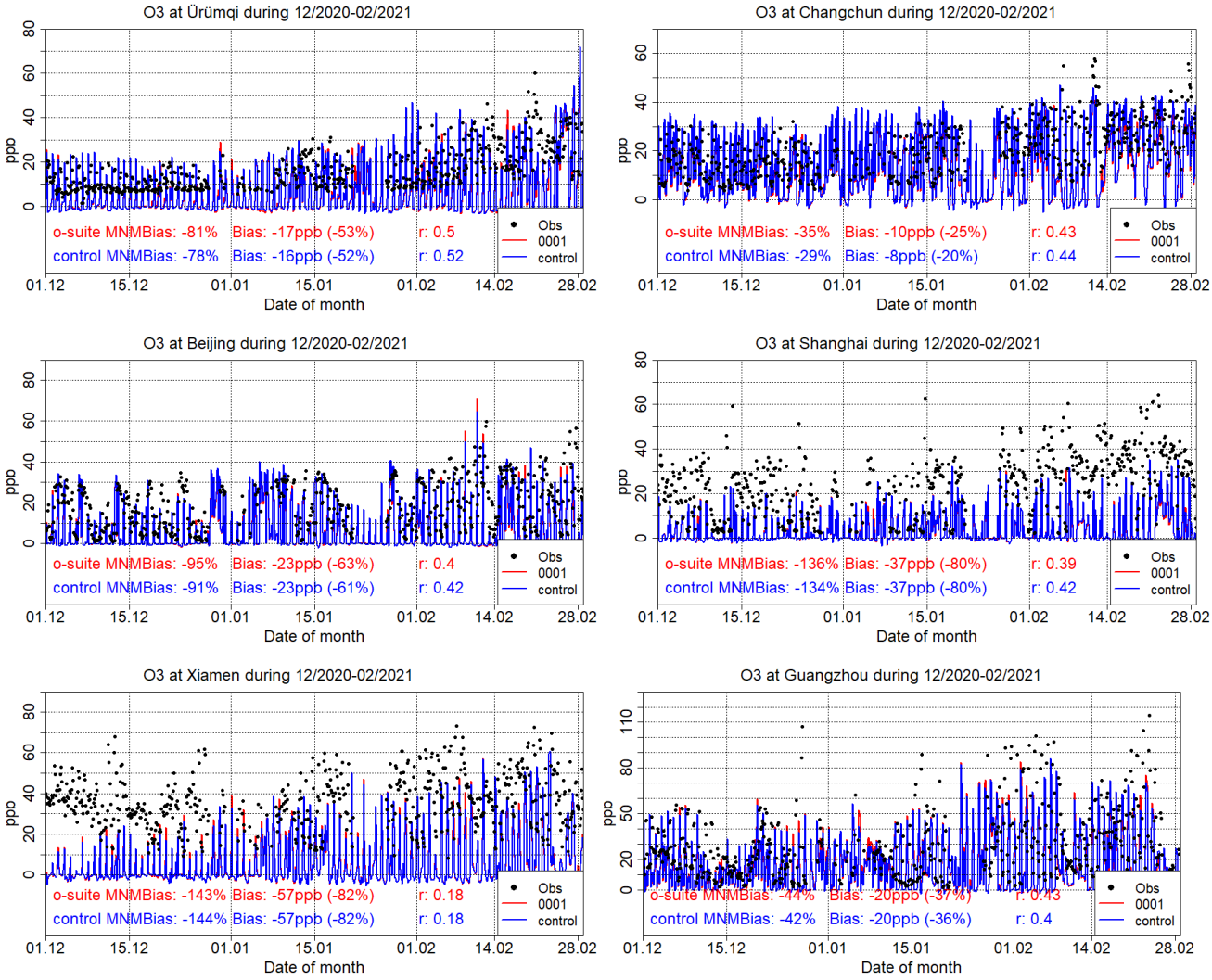


Figure 3.6.2: Surface ozone time series for the o-suite (red) compared to MarcoPolo-Panda project observations at Urumqi (43.83°N, 87.62°E, 1st row left), at Changchun (43.89°N, 125.33°E, 1st row right), at Beijing (39.92°N, 116.38°E, 2nd row left), at Shanghai (31.22°N, 121.47°E, 2nd row right), at Xiamen (24.48°N, 118.09°E, 3rd row left), and at Guangzhou (23.13°N, 113.25°E, 3rd row right). Period December 2020 – February 2021.

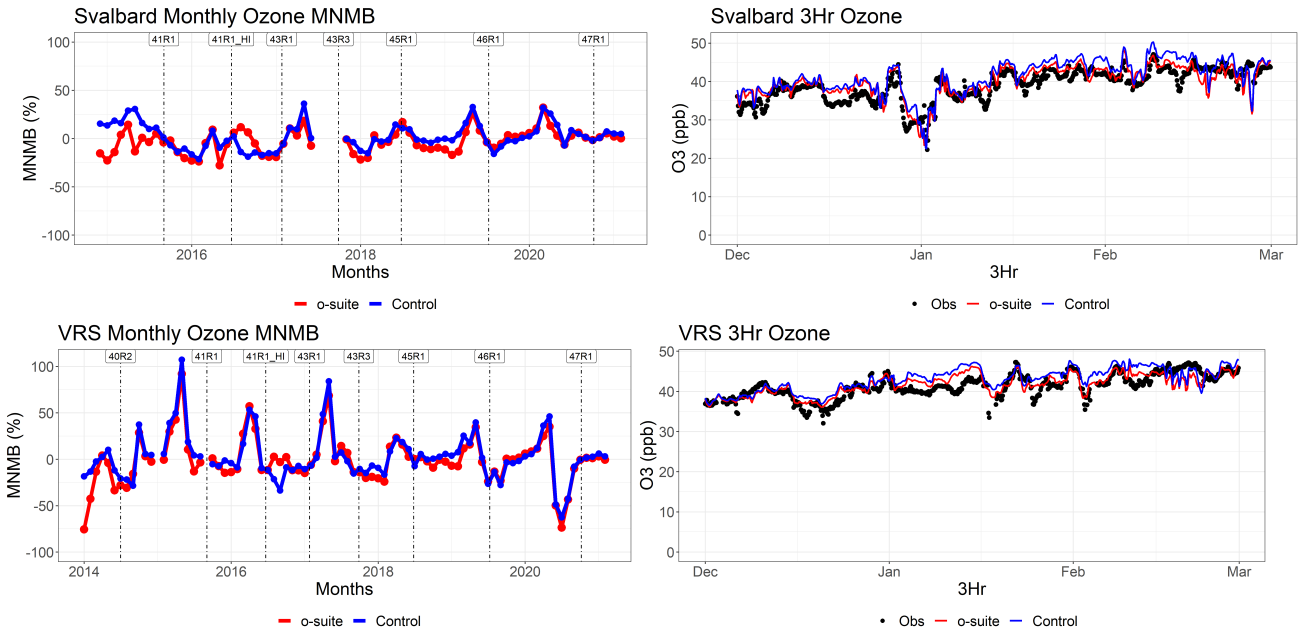


Figure 3.7.1: Time series for o-suite (red) and control (blue) compared to observations (black dots) at Svalbard (top row) and the Villum Research Station, Station Nord, Greenland (bottom row) MNMB for the full period (left, 2014/2015 up to 2021) and concentrations for December 2020 - February 2021 (right).

3.7 Validation with IASOA surface observations

CAMS results were compared to surface O_3 observations from the Villum Research Station, Station Nord in north Greenland (81.6°N 16.7°W) and Zeppelin Mountain, Svalbard (78.9°N 11.9°E) from the IASOA network (Fig. 3.7.1).

The data from Svalbard and VRS are covering the period from December 2014 to February 2021. The CAMS simulations do not capture ozone depletion events in March – June in 2015 – 2020 at any of the sites. These events are related to halogen chemistry reactions that are not represented in the CAMS simulations. The simulations are on average in good agreement with the observations apart from the spring depletion events.

For the period December 2020 – February 2021 the measurements are not quality controlled. The CAMS simulations are capturing the surface O_3 levels and variability very well at Svalbard and VRS with low bias and high correlation (Table 3.7.1).

Table 3.7.1. Modified Normalised Mean Bias (MNMB) and correlation coefficient (r) of the CAMS o-suite and control simulations for the sites Svalbard and Villum Research Station (VRS) for the period December 2020 – February 2021.

		MNMB	R
Svalbard	o-suite	0.03	0.83
	control	0.06	0.85
VRS	o-suite	0.01	0.78
	control	0.04	0.75

* Only data for October-November

3.8 Validation with IASI satellite data

Ozone total columns from the o-suite and control run are compared with IASI Metop-B version V6.5.0 daytime only satellite observations (Clerbaux et al., 2009). For the comparison with the IASI data, the vertically integrated model O_3 data were transformed using IASI averaging kernels (Rodgers, 2000).

The global distribution of the O_3 total column obtained from IASI, as well as the relative difference between the model runs and IASI, are shown in Fig. 3.8.1 for January 2021. Satellite data shows relatively high O_3 over the northern mid- and high latitudes, especially over the east of Russia and Canada and low values over the equatorial area. The o-suite run shows good agreement with the observations with bias within 10% with some exceptions. The underestimation of ozone over cold areas in the Northern Hemisphere, especially over east of Russia, reaches 30% and over Antarctica it is up to 20%. This is probably due to low IASI sensitivity over the cold surfaces. The IASI sensitivity is the lowest over the cold surfaces of Antarctica and Greenland (especially during March-April-May season) where IASI O_3 values are positively biased by up to 20%. The control run shows slightly higher overestimation over the oceans (up to 20%) and slightly smaller underestimation over Antarctica (within 10%). The forecast for day 4 is very similar to the analysis.

Figure 3.8.2. shows data as a function of latitude and time from January 2020 till February 2021. IASI data show record-low ozone values over the Arctic in March and over Antarctic during September-November. During the period from December 2020 – February 2021 the ozone values are high over northern mid- and low-latitudes and relatively low over the equatorial area. The o-suite run shows good agreement with observations with bias within 10% with some regional and temporal exceptions (e.g., bias about 20% over Southern high latitudes in October). The control run has distinct difference before and after model upgrade in October 2020. The control run shows relatively good results with bias within 10%. The overestimation of about 20% can be seen over southern high latitudes in October-December. The forecast day 4 shows smaller negative bias for both runs and growing positive bias over the high southern latitudes in October-November in the control run.

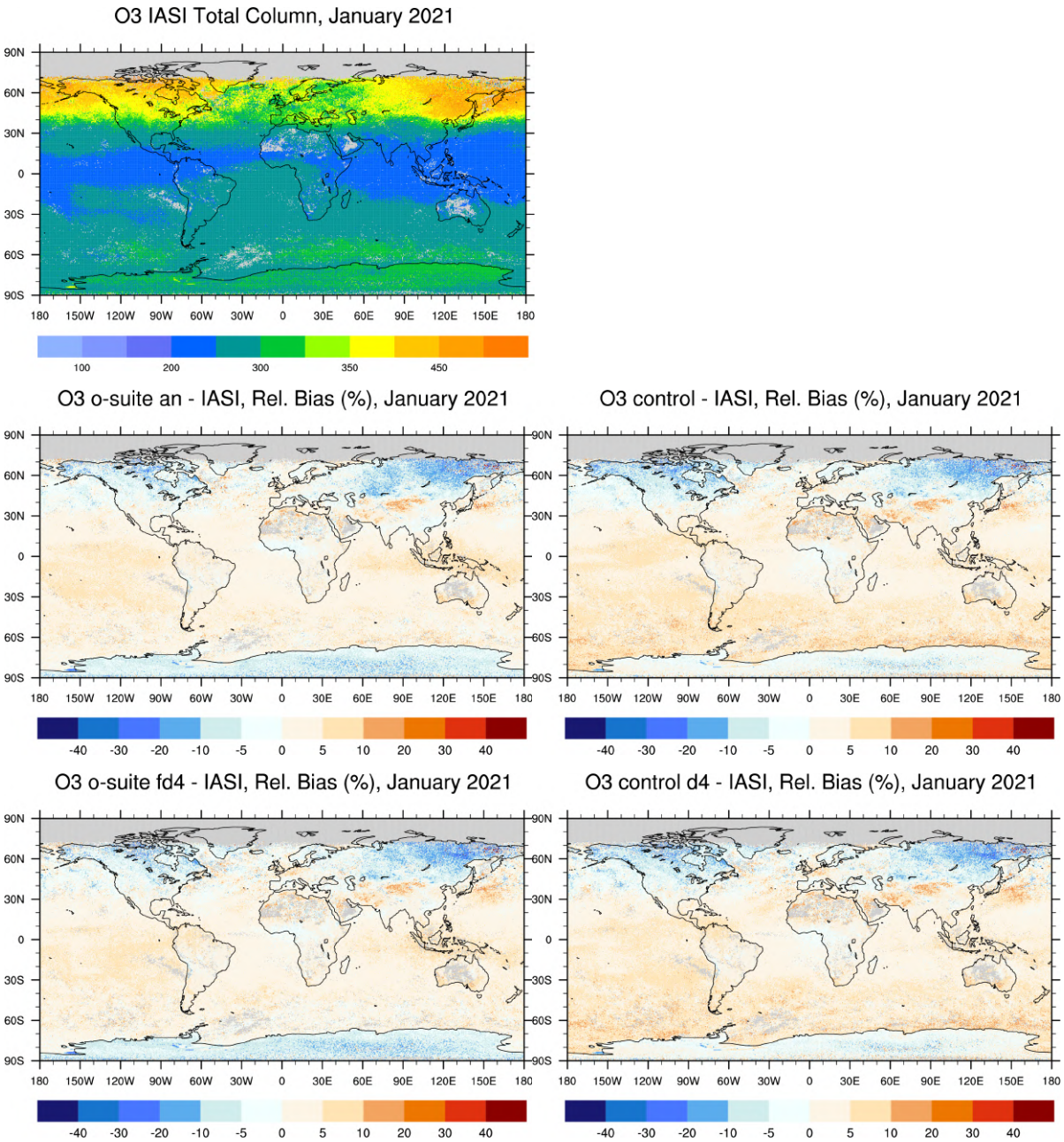


Figure 3.8.1: O_3 total column for IASI satellite observations (top) and relative difference between the model runs and IASI for January 2021: o-suite analysis and forecast day 4 (left), control run and control run forecast day 4 (right). Grey colour indicates missing values.

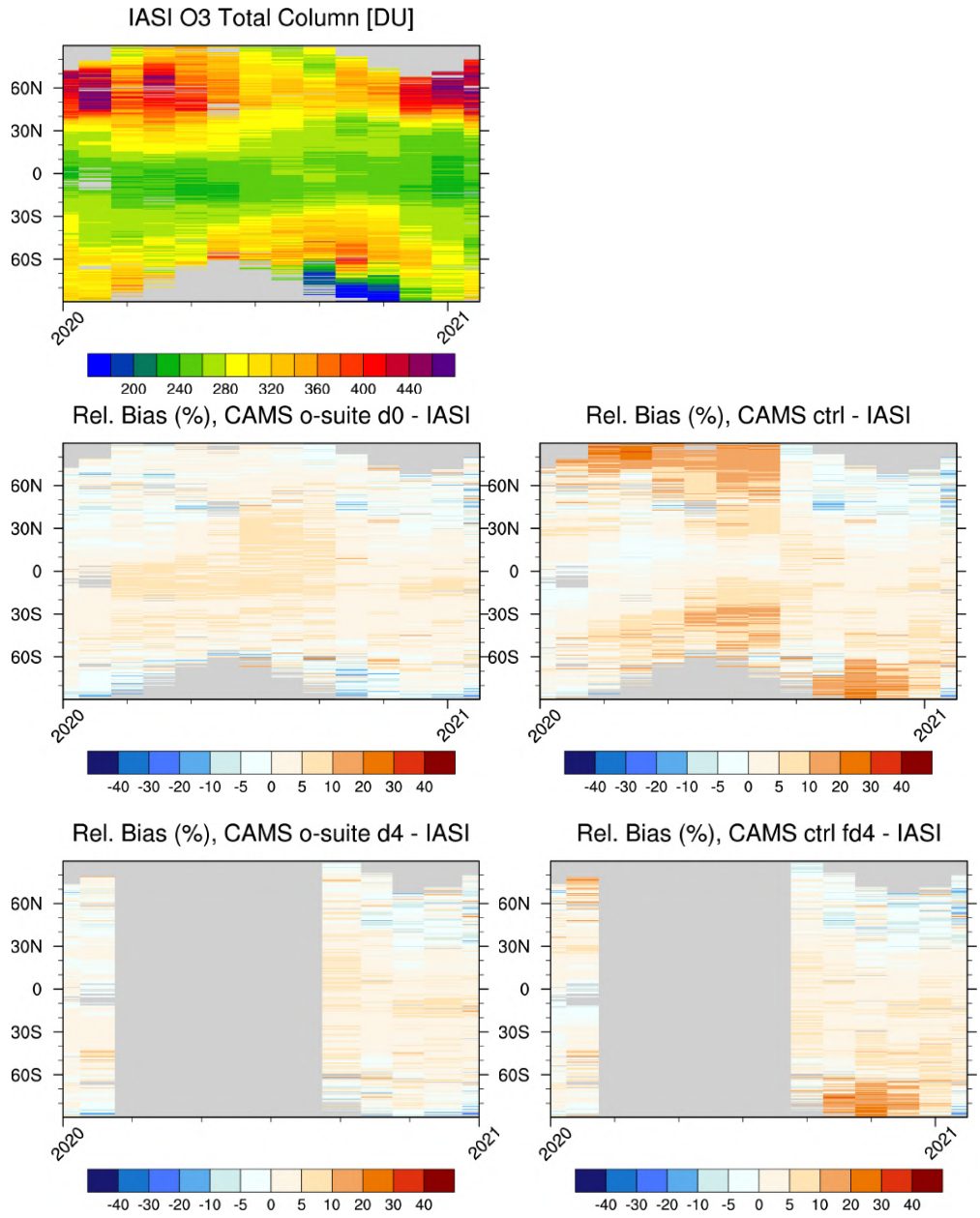


Figure 3.8.2: IASI Metop-B O_3 total column (top) as function of latitude and time from January 2020 to February 2021. Relative difference between the model runs and IASI: o-suite analysis and forecast day 4 (left), control run and control forecast day 4 (right). Grey colour indicates missing values.

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 Eskes et al. (2021) and is used for CAMS model evaluation for December 2020 to February 2021. The latest validation results can be found on the CAMS website, see section 1.

Both runs mostly show negative MNMBs for stations in the Northern Hemisphere (Fig. 4.1.1).

A comparison of the seasonal-mean MNMB for the o-suite over Europe (Fig. 4.1.2) from December 2012 to present shows an improving MNMB from about -20% in 2013 to less than -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.

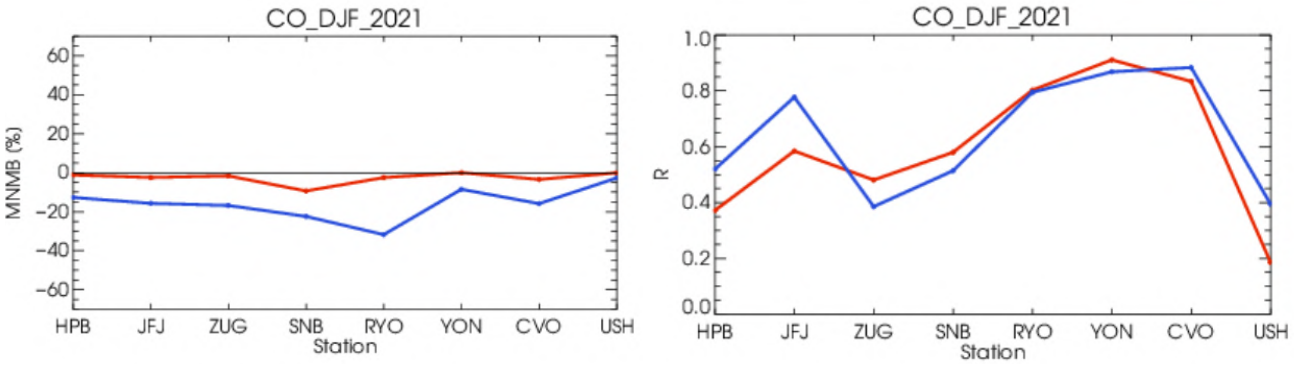


Figure 4.1.1: Modified normalized mean bias in % (left) and correlation coefficient (bottom right) of the NRT model runs compared to observational GAW data in the period December 2020-February 2021 (o-suite: solid red, and control: blue).

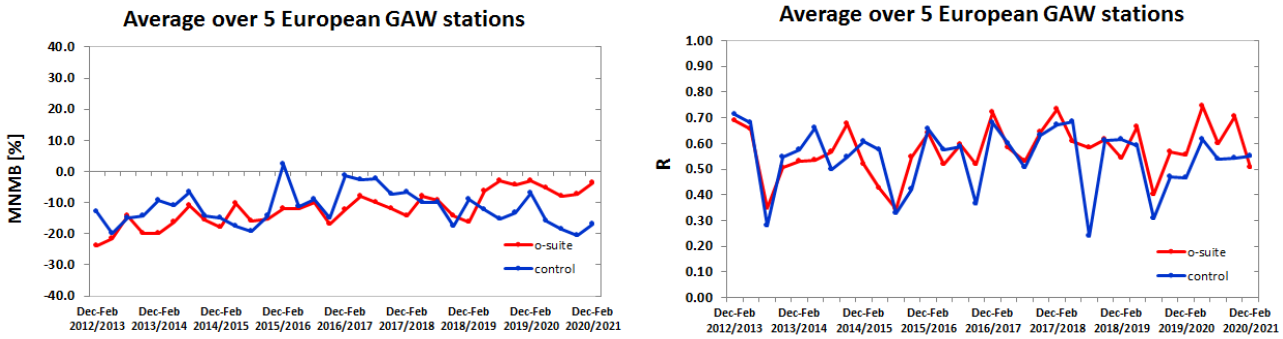


Figure 4.1.2: Long term (December 2012 – February 2021) evolution of seasonal mean MNMB (left) and correlation (right), as averaged over 5 GAW stations in Europe, for o-suite (red) and control (blue).

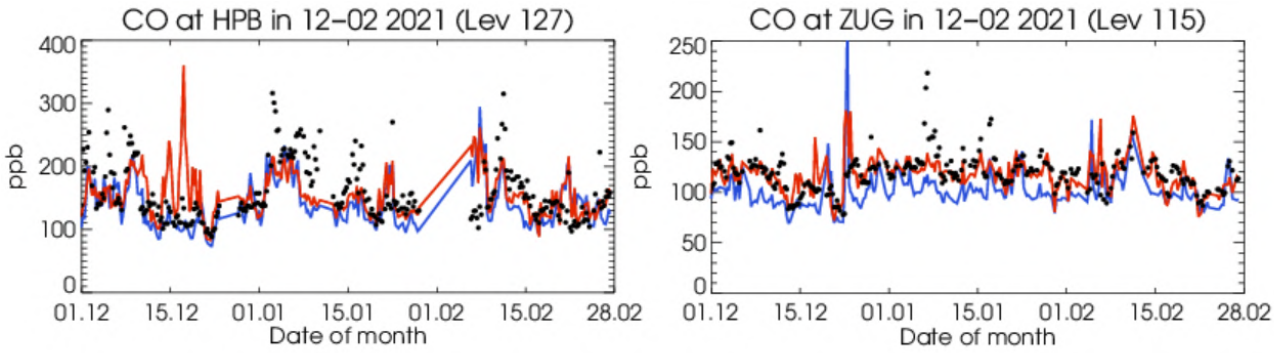


Figure 4.1.3: Time series for the o-suite (red) and control (blue) compared to GAW observations for Hohenpeissenberg (47.8°N, 11.02°E) and Zugspitze (47.4°N, 10.9°E). Period December 2020 – February 2021.

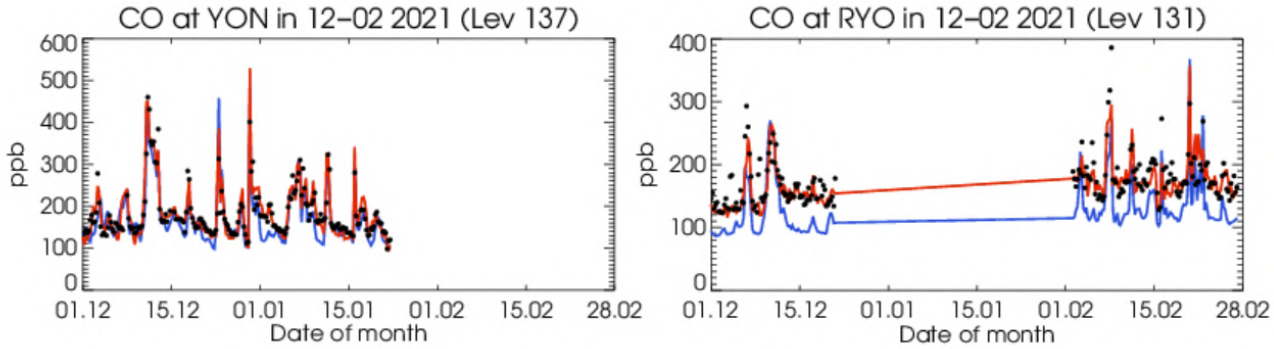


Figure 4.1.4: Time series for the o-suite (red) and control (blue) compared to GAW observations at Yonagunijima (24.47°N, 123.02°E) and Ryori (39.03°N, 141.82°E). Period December 2020 – February 2021.

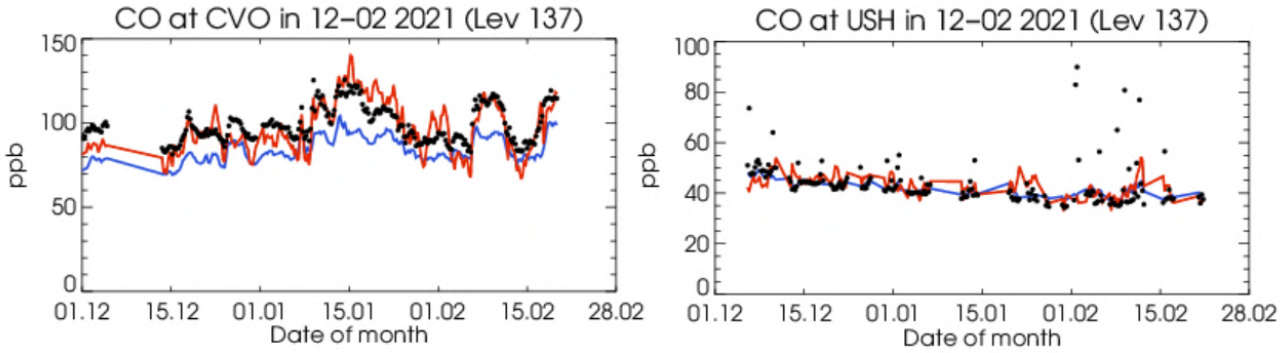


Figure 4.1.5: Time series for the o-suite (red) and control (blue) compared to GAW observations at Cape Verde (16.85°N, 24.87°W) and Ushuaia (54.85°S, 68.32°W). Period December 2020 – February 2021.

For European stations (Fig. 4.1.1. and Fig. 4.1.3), the o-suite shows only a slight underestimation of CO with MNMBs mostly within 5% (except for Sonnblick with bias of -12%), whereas the control run underestimates the observations with MNMBs up to -22%. Correlation coefficients are between 0.38 and 0.59 for the o-suite and between 0.38 and 0.8 for the control run.

For Asian stations, CO mixing ratios are slightly underestimated within 5% for the o-suite and by -32% for the control run. Correlation coefficients range between 0.80 and 0.93 for the o-suite and between 0.80 and 0.87 for the control run (Fig. 4.1.1. and Fig. 4.1.4).

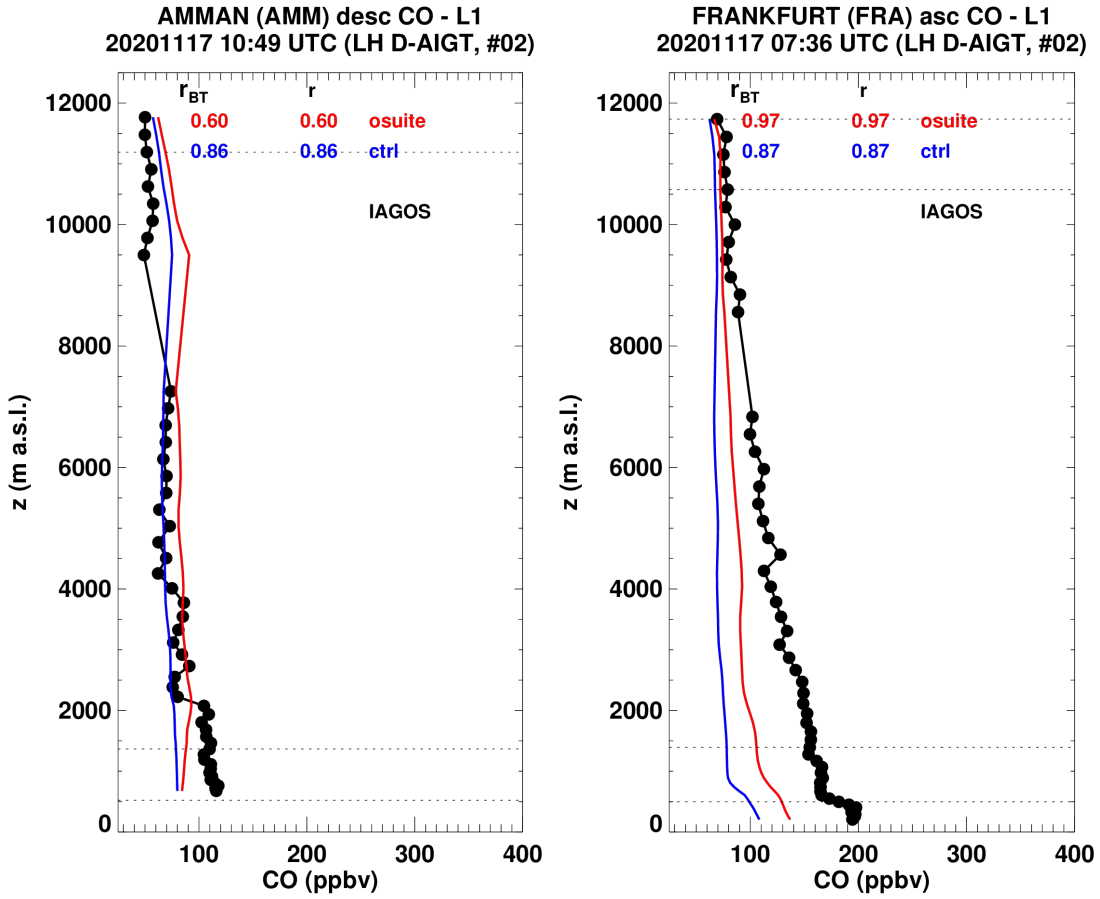


Figure 4.2.1.: Vertical profiles of CO measured with the IAGOS aircraft over Frankfurt and Amman on 17 November 2020, compared to o-suite (red) and control (blue).

For CVO, MNMBs is -2 for the o-suite with a correlation of 0.84 and MNMBs is -18 for control run with a correlation of 0.88 (Fig. 4.1.1 and Fig.4.1.5).

For the station in the Southern mid-latitudes (USH), MNMBs is almost 0% for the o-suite and -5% for the control run. Correlation coefficients are relatively low for both runs (0.19 for the o-suite and 0.39 for the control run).

4.2 Validation with IAGOS Data

Comparison with profile data

Similar to ozone, CO time series at Frankfurt are available starting from the end of December and with some gaps in January and February (Fig. 4.2.1). During this period, CO values observed at Frankfurt in the surface and boundary layer are mostly in between 150 and 200 ppbv with a two major peaks around 4 and 15 January with daily values between 400 and 500 ppbv (Fig. 4.2.1). CO is mostly underestimated by the o-suite, and the largest biases are generally found in the low troposphere, with a slightly better agreement from the o-suite (Fig.4.2.1, 4.2.2 and 4.2.3). On average in the lowest layers an MNMB of about -15% is found for the o-suite against -30% for control run, while correlation values are similar for the two runs with about 60% (Fig. 4.2.3). In the free troposphere the bias is smaller for both runs with an MNMB within 10 % for the o-suite and

between -20% and -5% for control run (Fig. 4.2.3). Correlation results from the two runs are very similar in the mid-troposphere with values between 40 and 70% and in general values slightly higher values for the o-suite than control run (Fig. 4.2.3).

A few individual profiles are shown in Fig. 4.2.4 corresponding to the highest CO values observed during this period and in particular on the profiles of 4 and 15 January with the maximum values observed over the entire period in the surface and boundary layer. All these profiles show clearly that CO is underestimated by both models with the largest bias when CO values are the highest close to the surface (> 400 ppbv). In general, the o-suite performs better than control run. In the low troposphere and in particular in the free troposphere where the bias of the o-suite is the lowest. In the UTLS the results of the models differ less and the agreement with observations is good.

In order to compare the model performance with the previous DJF period at Frankfurt, Fig. 4.2.5 presents the time series of the monthly values for surface and free tropospheric CO together with the associated time series of the MNMB from the two runs for a period starting back from December 2019. In the surface layer, MNMB values for this DJF period are similar to those observed in the previous DJF period for the two runs, although a slight increase in MNMB (absolute) values from control run is found with respect to the same quarter of last year. In the free troposphere, MNMB (absolute) values from the o-suite in DJF 2021 appear smaller compared with that of DJF 2020. Conversely, for control run MNMB (absolute) values are larger.

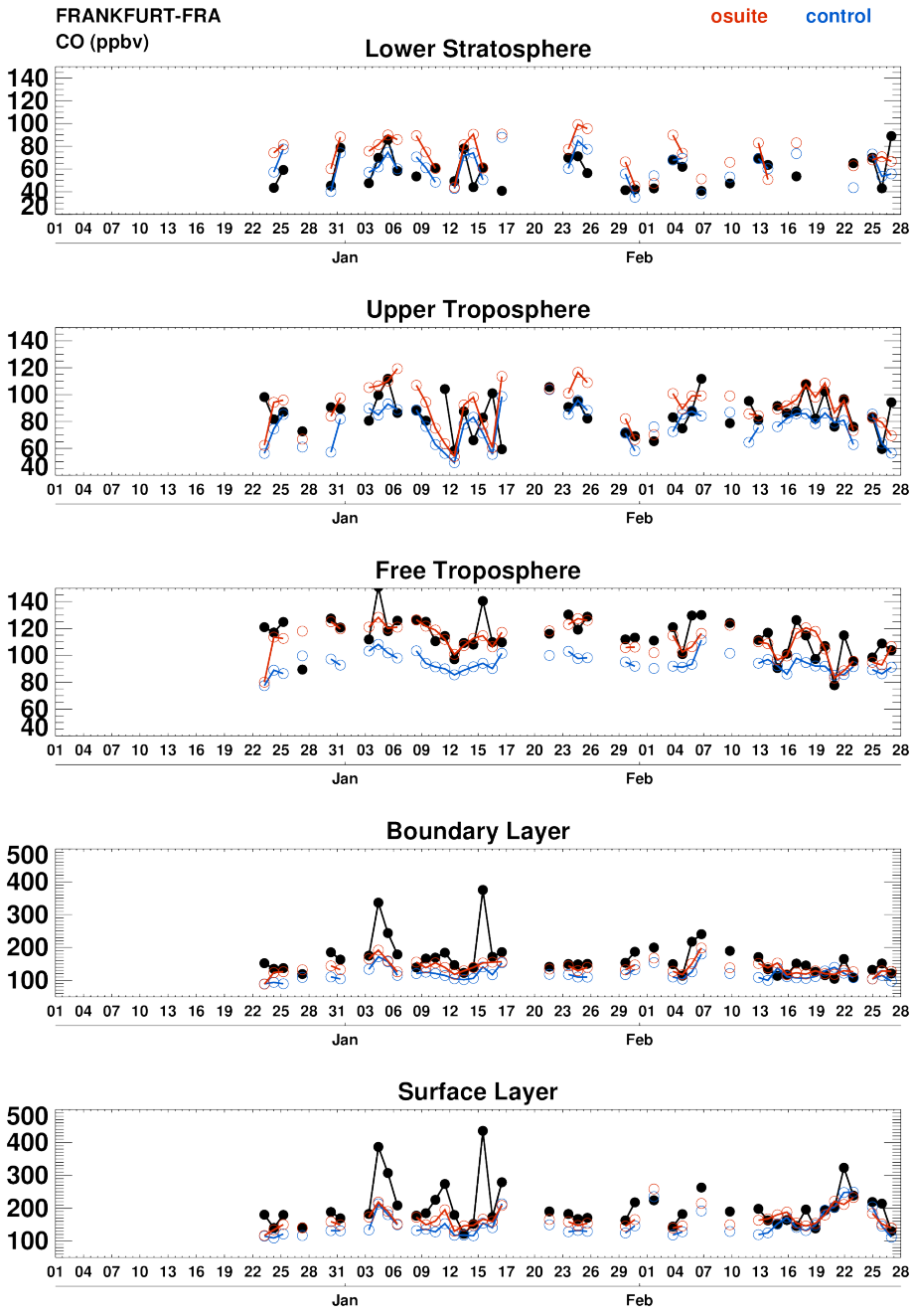


Figure 4.2.1: Time series of daily mean CO over Frankfurt during DJF 2021 for 5 layers: Surface Layer, Boundary Layer, Free Troposphere, Upper Troposphere and Lower Stratosphere. IAGOS is shown in black, the o-suite in red and associated control run in blue. Units: ppbv.

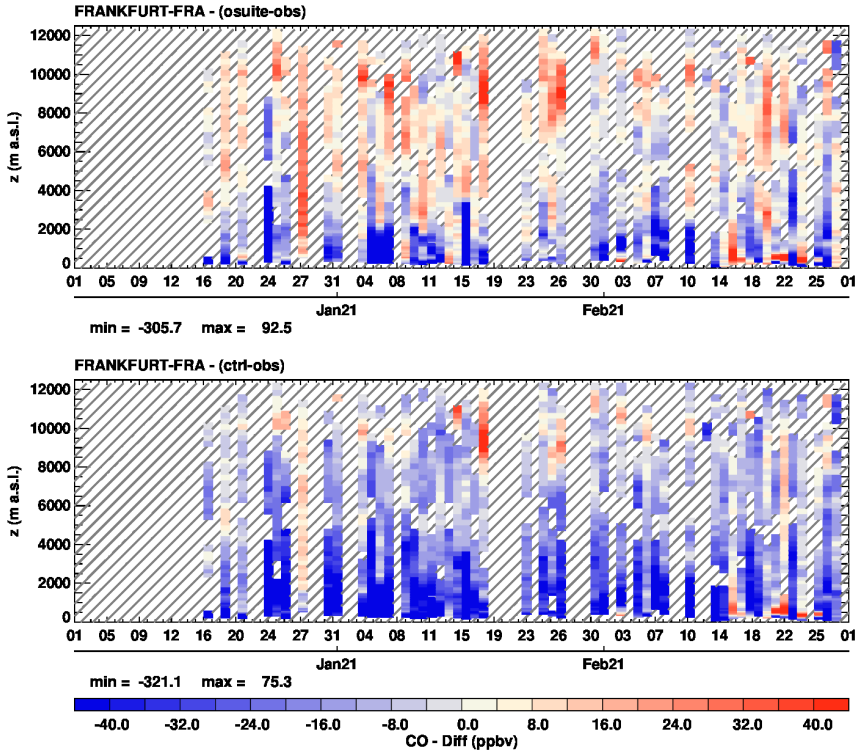


Figure 4.2.2: Time series of the absolute differences (model – IAGOS aircraft observations) in daily profiles for CO over Frankfurt during DJF 2021. The top panel corresponds to o-suite, the bottom panel to control run. Units: ppbv.

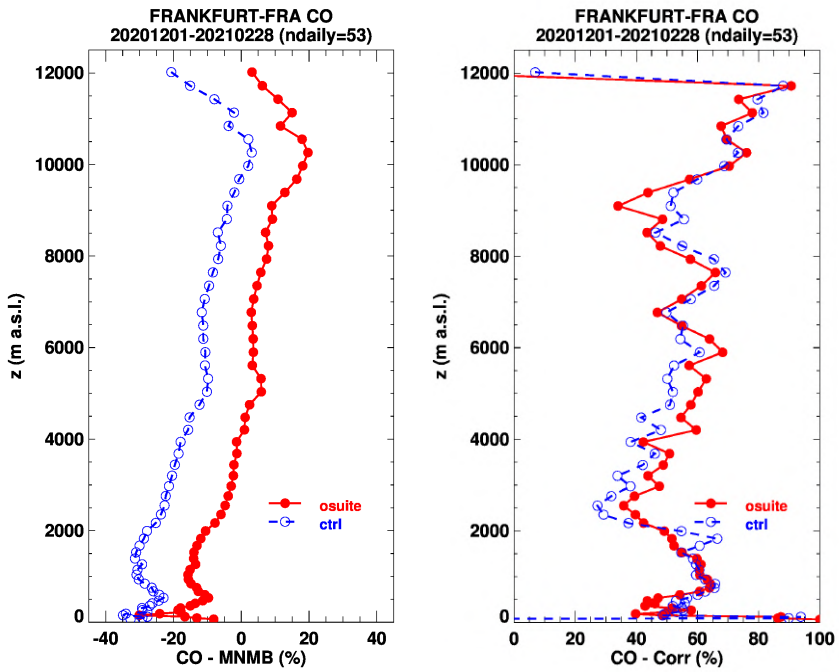


Figure 4.2.3: Model scores (MNMB and Correlation coefficient) for CO at Frankfurt calculated over the period DJF 2021. The left panel corresponds to MNMB and the right panel to Correlation coefficient. The o-suite is shown in red and associated control run in blue. Units: %.

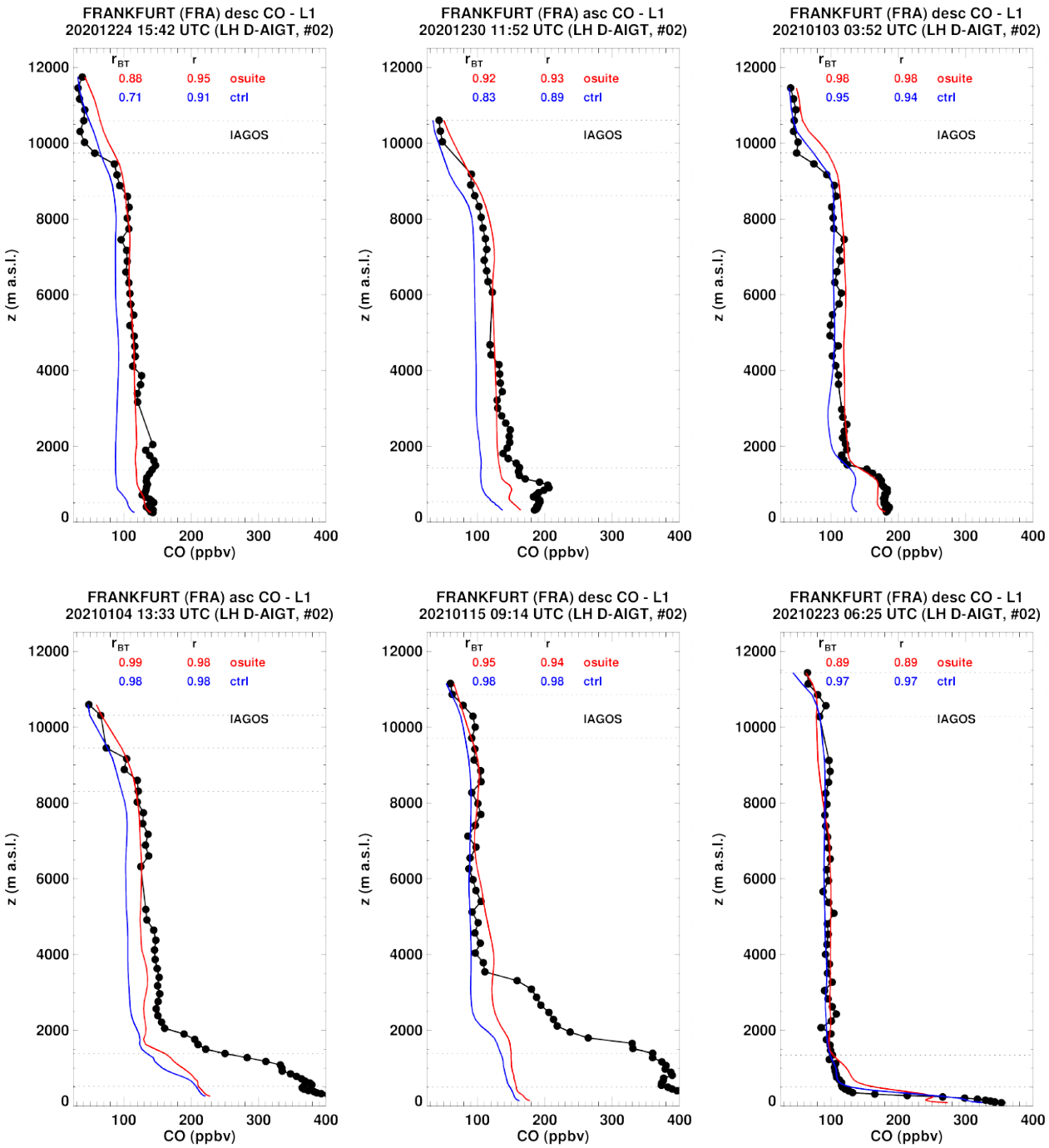


Figure 4.2.4: Selection of individual profiles for CO from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over Frankfurt during DJF 2021. Units: ppbv.

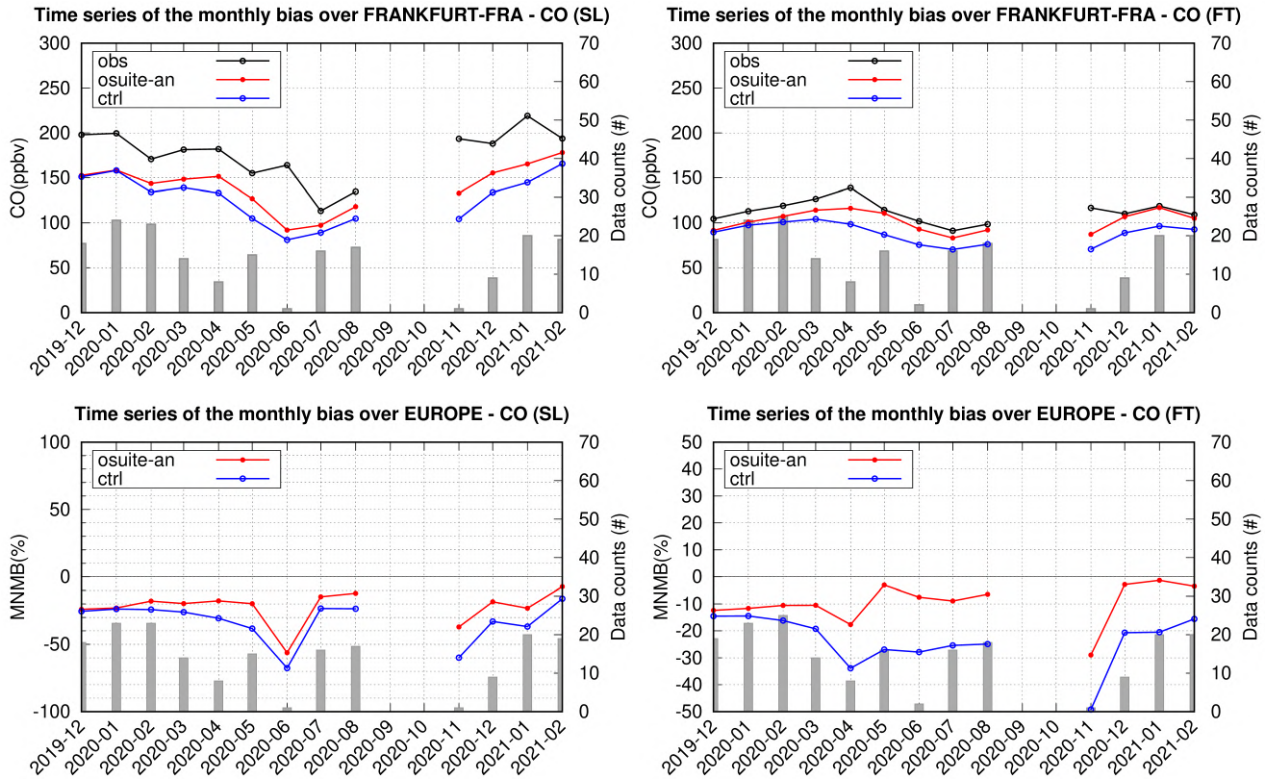


Figure 4.2.5: Top panels: monthly time series for CO from IAGOS from o-suite analysis (red) and control run (blue) at Frankfurt in different atmospheric layers (left: surface layer, right, free troposphere) during the period December 2019 - February 2021. Bottom panels: corresponding MNMB time series for the two runs. The histogram bars indicate the number of available observations.

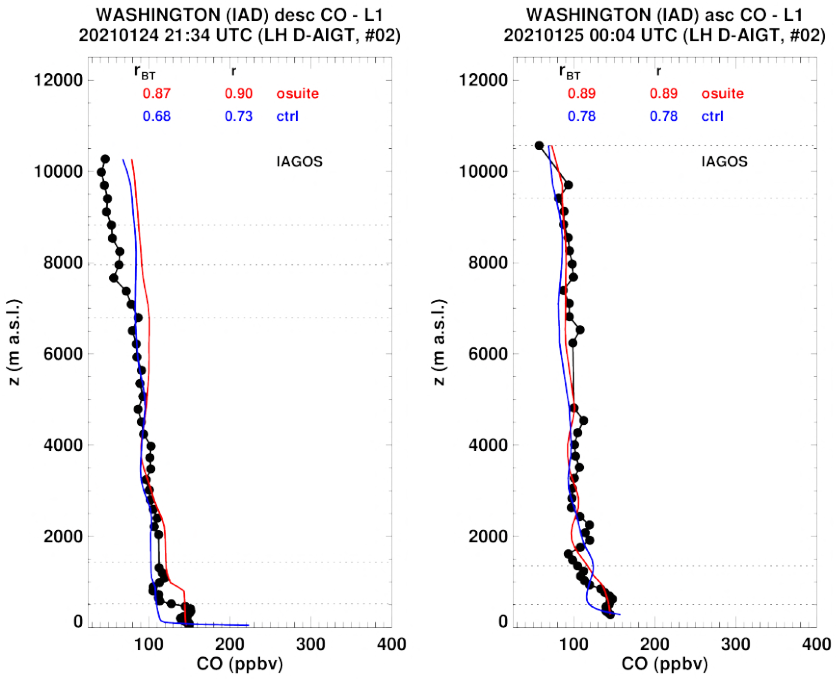


Figure 4.2.6: Selection of individual profiles for CO from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over North America during DJF 2021. Units: ppbv.

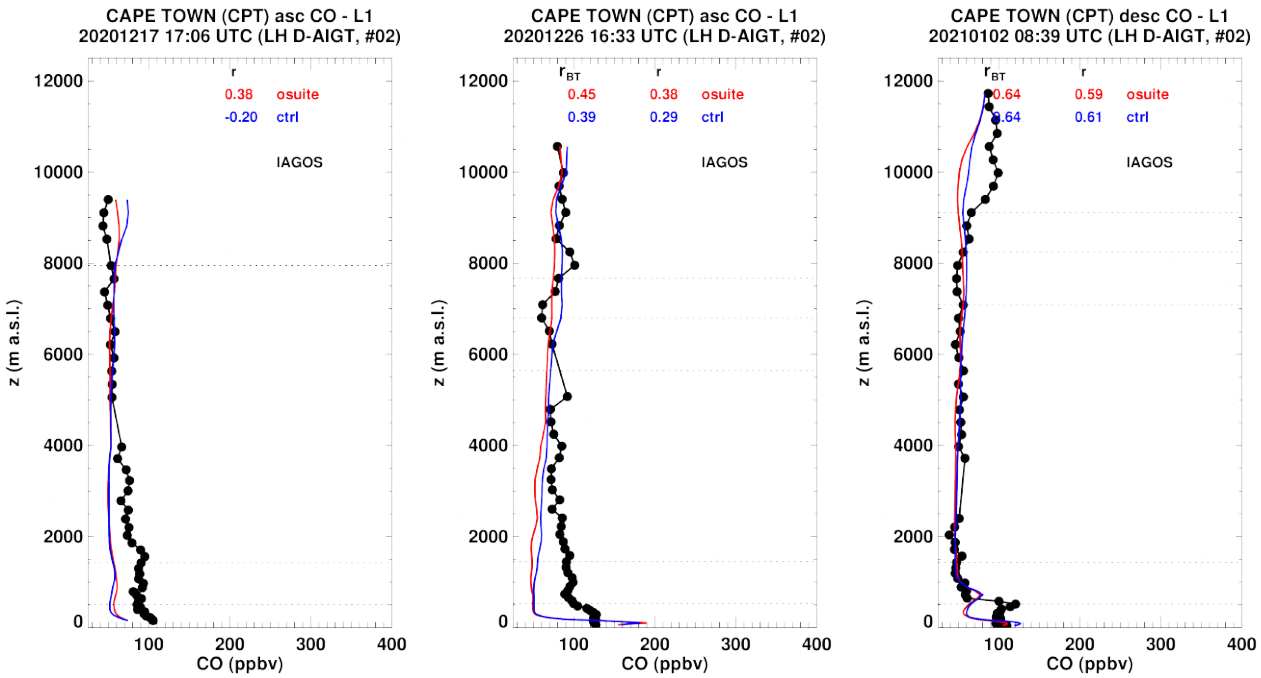


Figure 4.2.7: Selection of individual profiles for CO from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over South Africa during DJF 2021. Units: ppbv.

North America

Only two CO profiles are available over North America and at the airport of Washington during DJF 2021 (Fig. 4.2.6). In these profiles CO surface values are of about 150 ppbv. In the lowest layers the two runs differ slightly with a better performance from the o-suite. In the free troposphere and above the models behave more similarly and agree well with observations although one of the profile present overestimations in the UTLS where low values of CO are observed.

South Africa

Over South Africa a number of CO profiles are available at the airport of Cape Town (Fig. 4.2.7). For these profiles the results from the two runs are rather similar from the surface to the UTLS. In the low to mid-troposphere the bias is large and CO values are underestimated. In the upper layers a better agreement with observations is generally found.

Eastern and South-eastern Asia

Over Eastern and South-eastern Asia, a number of CO profiles are available only at the airport of Hong Kong and two CO profiles at the airport of Singapore (Fig. 4.2.8). At these airports the two runs often behave similarly except for some cases in the low troposphere at Hong Kong. At Hong Kong the agreement of the models with observations is rather good from the low to mid-troposphere, with the exception of large overestimations near the surface in particular for high CO values (>400 ppbv). Like at Hong Kong, one profile at Singapore shows overestimation near the surface, but the models underestimate from the boundary layer to the mid-troposphere. In upper layers the behaviour of the models remains similar, and the agreement is better at both airports.

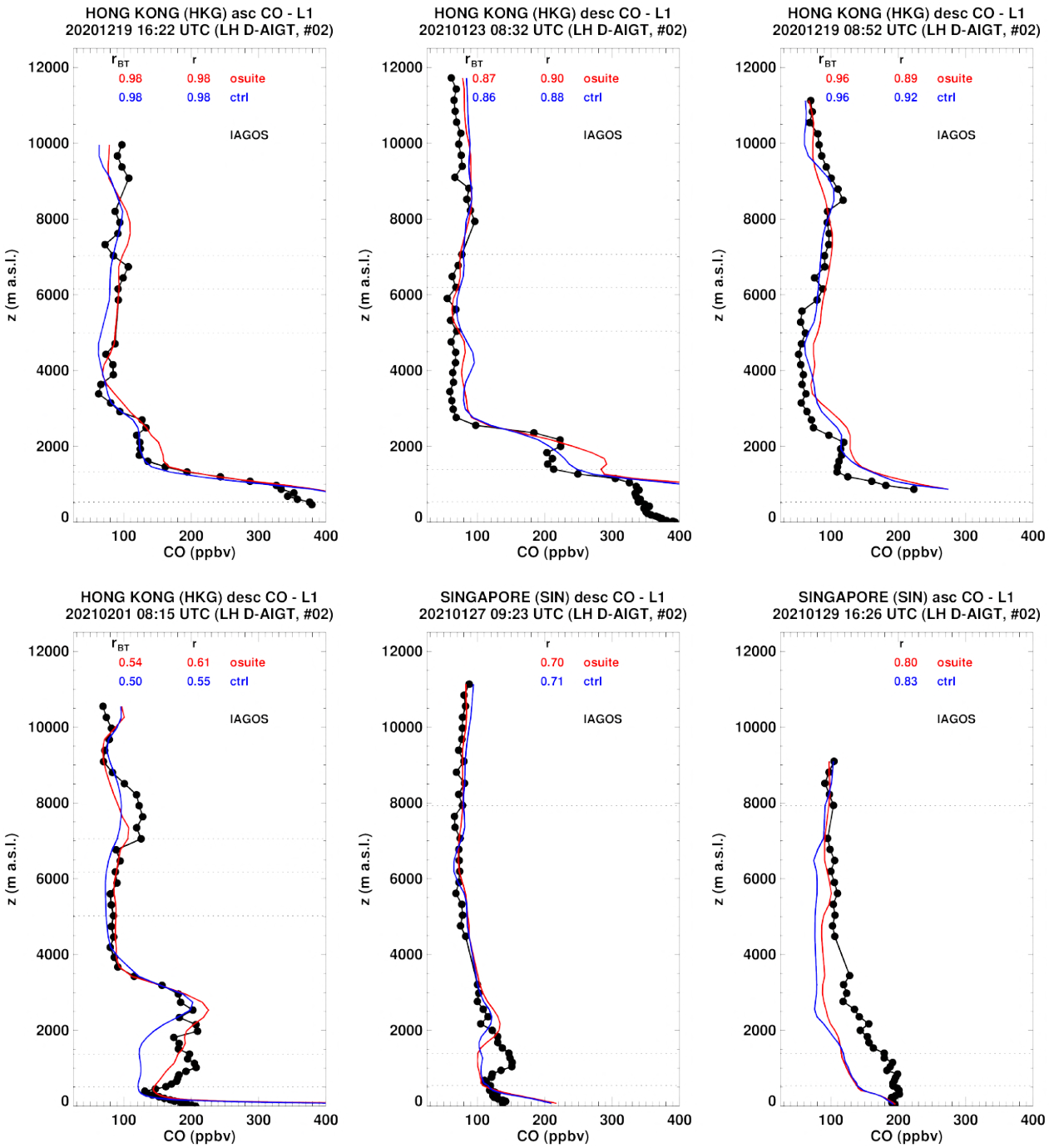


Figure 4.2.8: Selection of individual profiles for CO from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over Eastern and Southeastern Asia during DJF 2021. Units: ppbv.

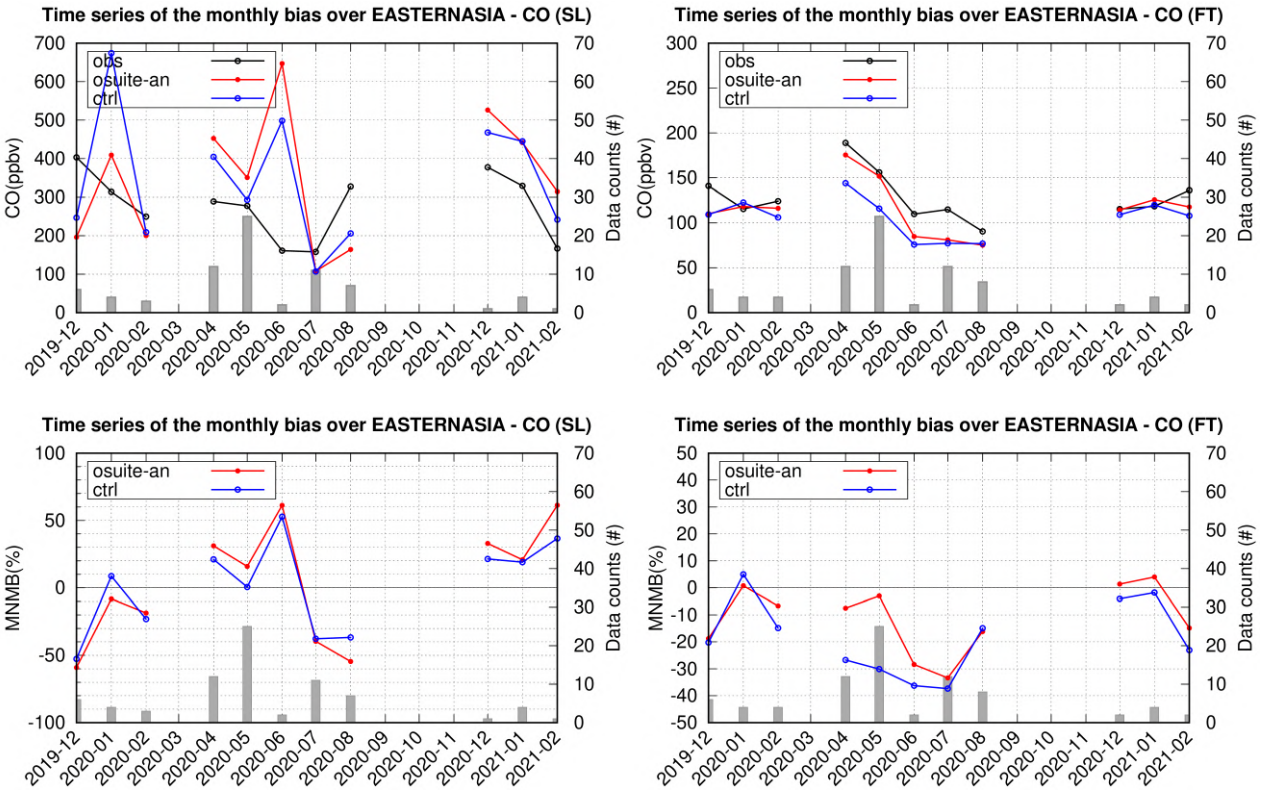


Figure 4.2.9: Top panels: monthly time series for CO from IAGOS from o-suite analysis (red) and control run (blue) over Eastern and Southeastern Asia in different atmospheric layers (left: surface layer, right, free troposphere) during the period December 2019 - February 2021. Bottom panels: corresponding MNMB time series for the two runs. The histogram bars indicate the number of profiles (i.e. layer values) based on available observations.

Eastern and Southeastern Asia

Over Eastern and Southeastern Asia, a number of CO profiles are available only at the airport of Hong Kong and two CO profiles at the airport of Singapore (Fig. 4.2.8). At these airports the two runs often behave similarly except for some cases in the low troposphere at Hong Kong. At Hong Kong the agreement of the models with observations is rather good from the low to mid-troposphere, with the exception of large overestimations near the surface in particular for high CO values (>400 ppbv). Like at Hong Kong, one profile at Singapore shows overestimation near the surface, but the models underestimate from the boundary layer to the mid-troposphere. In upper layers the behaviour of the models remains similar and the agreement is better at both airports.

Like for Frankfurt, Fig. 4.2.9 presents the time series of the monthly values for surface and free tropospheric CO together with the associated time series of the MNMB from the two runs for a period starting back from December 2019. In the surface layer, for both runs MNMB values in DJF 2021 are positive while there were mostly negative in DJF 2020. However, it is not clear whether this could be due to different available airports within the considered region for the two different quarters especially as sampling is poor. In the free troposphere, the MNMB does not present a clear and systematic behaviour for the two DJF quarters, and no notable difference is found for this quarter compared to that of the previous year.

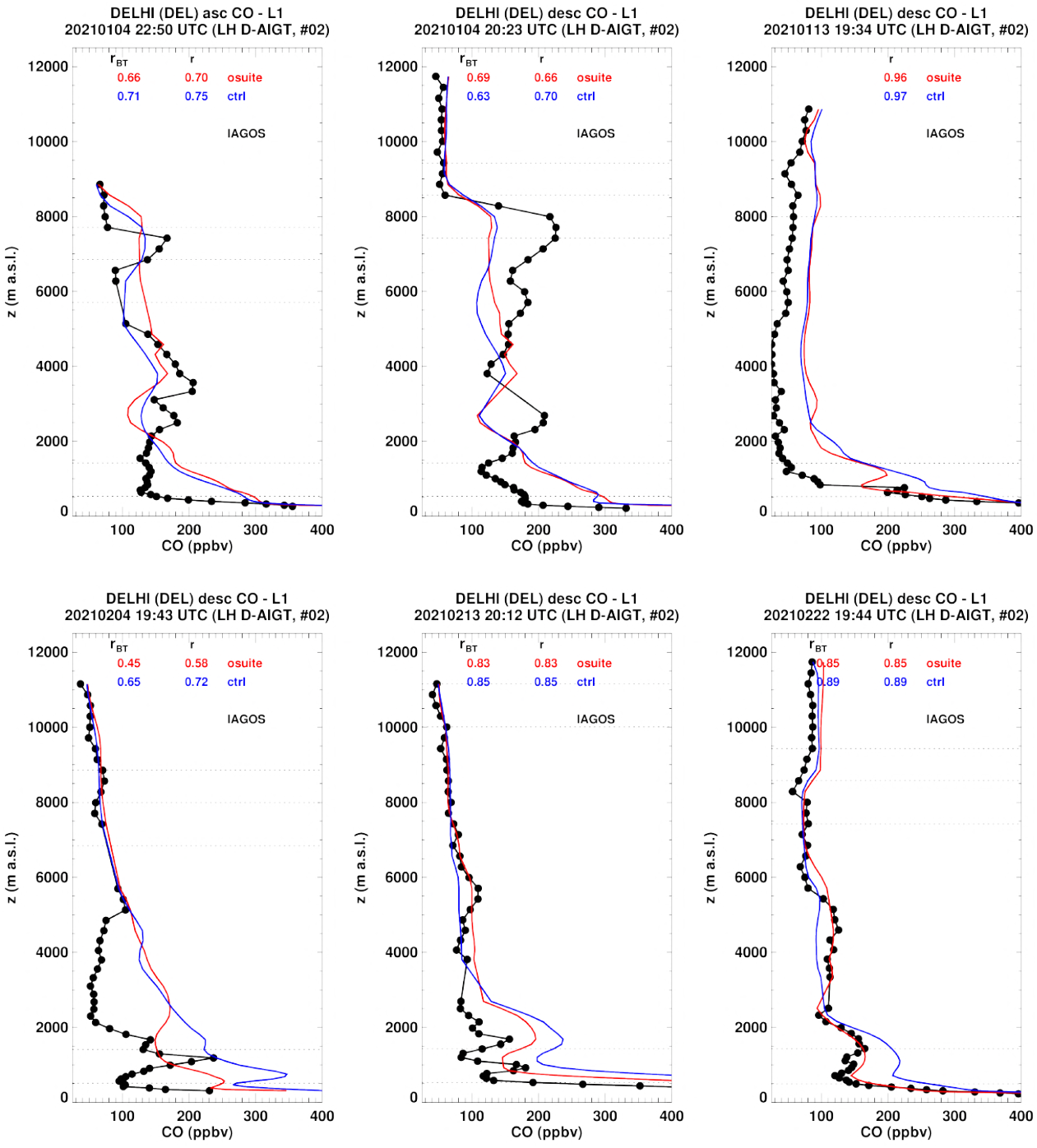


Figure 4.2.10: Selection of individual profiles for CO from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over India during DJF 2021.

India

During DJF 2021, several profiles of CO are available over India at the airport of Delhi (Fig. 4.2.10). The performance of the two runs are mostly similar. CO values are often overestimated in the surface and boundary layer by both runs, with smaller bias from the o-suite. In the mid-troposphere and UTLS the agreement with observations is in general better for both models.

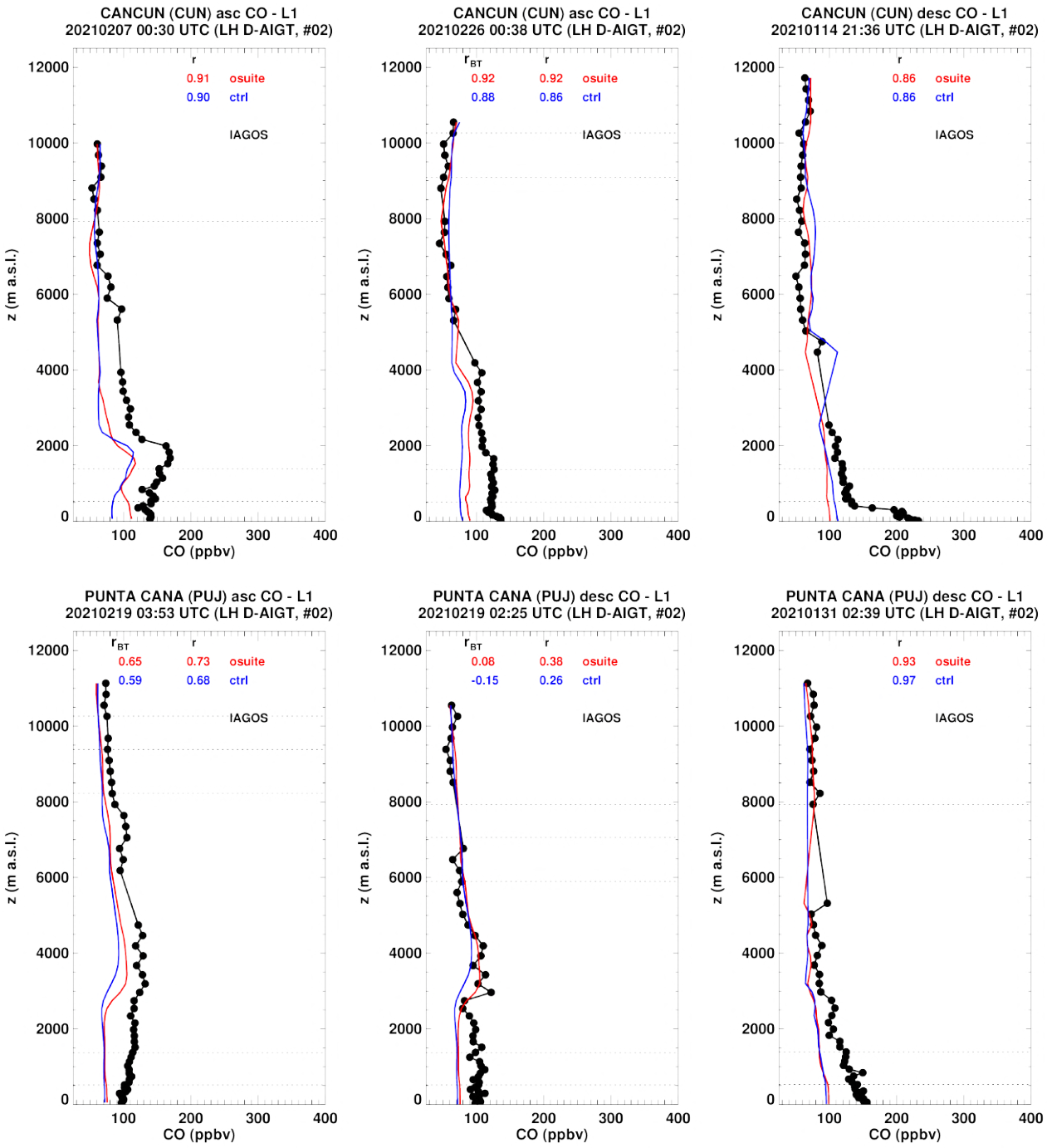


Figure 4.2.11.a: Selection of individual profiles for CO from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over Central America and Caribbean during DJF 2021.

Central America and Caribbean

During DJF 2021, several profiles of CO are available over different locations across Central America and the Caribbean: Cancun, San Jose and Punta Cana (Fig. 4.2.11.a-b). The performance of the two runs is similar. CO values are often underestimated from the surface to the mid-troposphere. In the upper layers the agreement with observations is in general better.

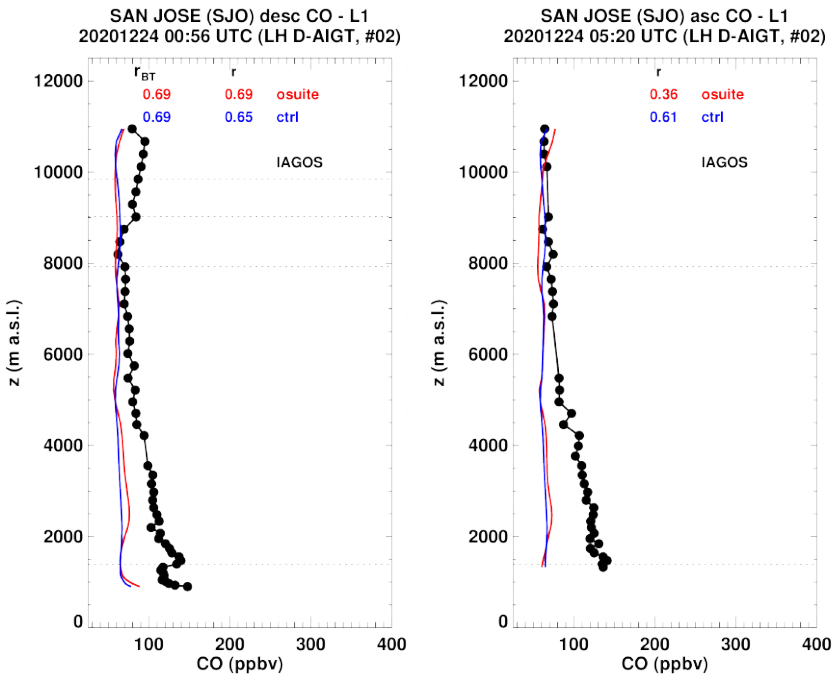


Figure 4.2.11.b: Selection of individual profiles for CO from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over Central America and Caribbean during DJF 2021.

Comparison with cruise level data

Fig. 4.2.12 shows the gridded average maps for the comparisons with IAGOS observations of CO at flight level. The spatial variability of CO is well represented by both CAMS configurations. The o-suite presents a small relative bias within 15% over most regions. However large negative biases are found over the Northwest Atlantic and conversely large positive biases occur over North Eastern Asia and the Mediterranean. Control run mostly underestimates CO and the absolute value of the bias is always larger than that of the o-suite. In particular, large negative biases are found in the northern mid-latitudes.

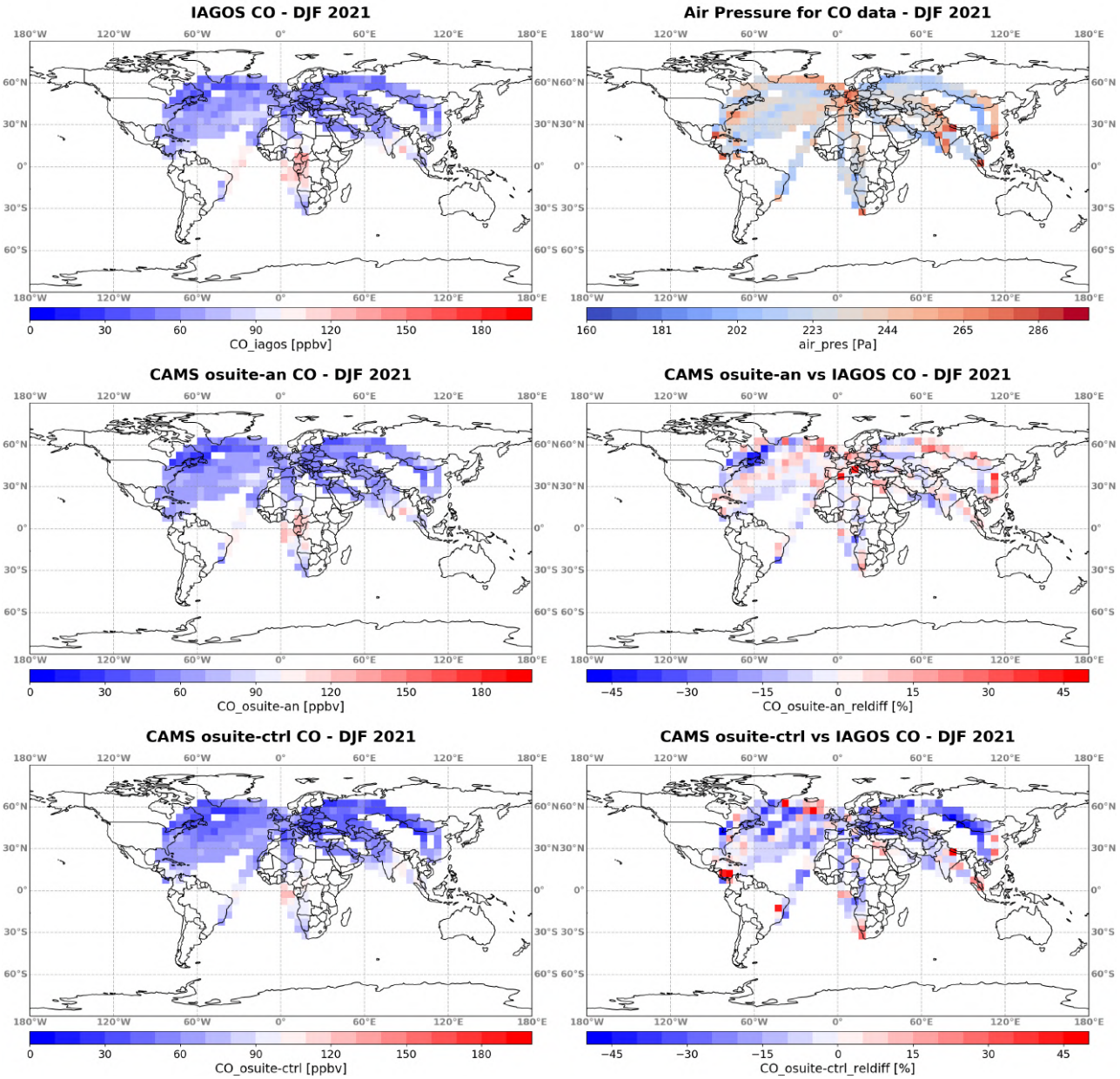


Figure 4.2.12: Global maps of gridded averages (5°x5°) for CO comparison with IAGOS cruise data during DJF 2021. From left to right, first row: IAGOS CO (in ppbv) and air pressure (in hPa). Second row: CO from the analysis of the o-suite (in ppbv) and associated relative differences (in %) with respect to IAGOS. Third row: same as second row for control run.

4.3 Validation against FTIR observations from the NDACC network

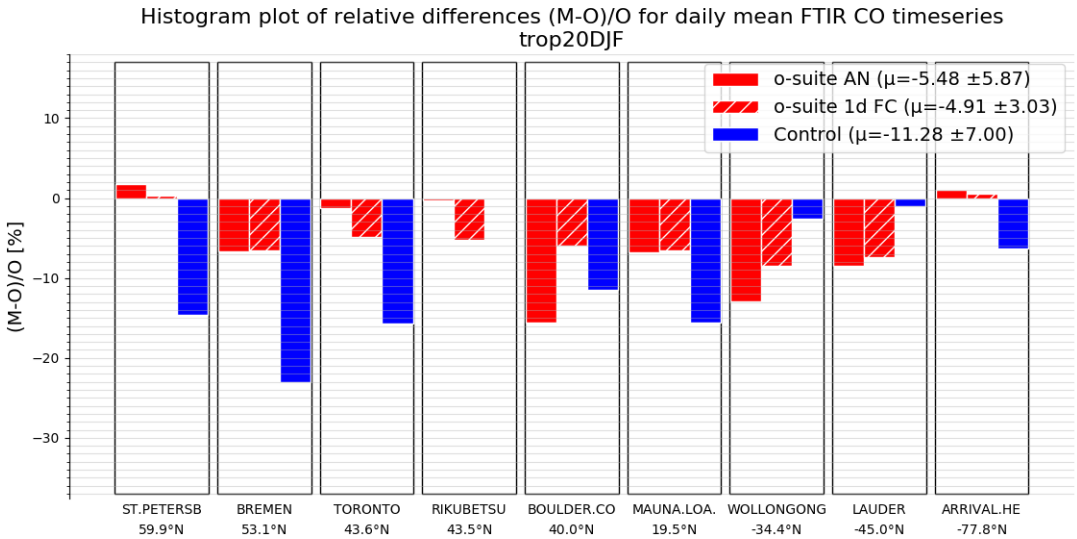
In this section, we compare the CO profiles of the CAMS products with FTIR measurements at 21 FTIR stations within the NDACC network. These ground-based, remote-sensing instruments are sensitive to the CO abundance in the troposphere and lower stratosphere, i.e. between the surface and up to 20 km altitude. Tropospheric and stratospheric CO partial columns are validated. A description of the instruments and applied methodologies can be found at http://nors.aeronomie.be.

Table 4.3.1: Detailed statistics for tropospheric CO column comparisons for FTIR measurements during DJF 2021. Both analysis and 1d FC o-suite behave similar, except at Boulder where the bias is significantly lower for the o-suite AN.

	o-suite AN tropospheric column					o-suite 1d FC tropospheric column					
FTIR site	#	rel. std	corr	rel diff (%)	rel diff std(%)	#	rel. std	corr	rel diff (%)	rel diff std(%)	lat
ST.PETERSBURG	7	0.8	0.98	1.76	2.81	7	0.9	0.95	0.29	3.4	59.9
BREMEN	6	0.7	0.62	-6.68	4.74	6	0.8	0.3	-6.51	6.05	53.1
TORONTO	30	1	0.85	-1.29	4.73	30	1.2	0.73	-4.82	5.65	43.6
RIKUBETSU	5	0.9	0.69	-0.25	2	5	0.7	0.68	-5.25	2.32	43.5
BOULDER.CO	32	1.1	0.74	-15.63	5.62	32	1	0.78	-5.98	5.67	40
MAUNA.LOA.HI	7	1.3	0.69	-6.85	4.11	7	1	0.7	-6.59	4.49	19.5
WOLLONGONG	11	1.1	0.93	-12.92	3.25	11	1.1	0.85	-8.49	4.93	-34.4
LAUDER	41	1.1	0.96	-8.44	3.4	41	1.1	0.97	-7.36	3.43	-45
ARRIVAL.HEIGHTS	21	0.9	0.97	1.02	4.41	21	0.9	0.97	0.56	4.59	-77.8
		1	0.83	-5.48	3.9		1	0.77	-4.91	4.5	

Figure 4.3.1 show that the o-suite tropospheric columns of CO agree well. The model upgrade (60 to 137 levels) implemented in July 2019 changes the overall biases in both the troposphere and stratosphere. The negative bias for the tropospheric columns increased from -2% before July 2019 to -7% in the most recent quarterly and is larger than the reported 3% measurement uncertainty. The stratospheric column bias also changed to -7% in DJF 2021 compared to values well above +10% before July 2019. The current stratospheric bias is below the reported 10% measurement uncertainty. There are no large differences between the o-suite AN and 1d FC, except at Boulder and Jungfraujoch where the o-suite AN performs significantly worse than the o-suite 1d forecast (see Table 4.3.1).

Figure 4.3.2 shows a negative trend in the tropospheric CO column at Jungfraujoch (4km – TP) of about 2% per year. A similar trend is observed at Zugspitze (3km above sea level), but not at other non-mountain sites like St Petersburg. The trend at the o-suite 1dFC at both mountain stations is much lower (around -0.5%/y), which suggests the trend is located in the upper tropospheric column and is related to the assimilation. The negative trend seems to have stopped with the June 2019 upgrade.



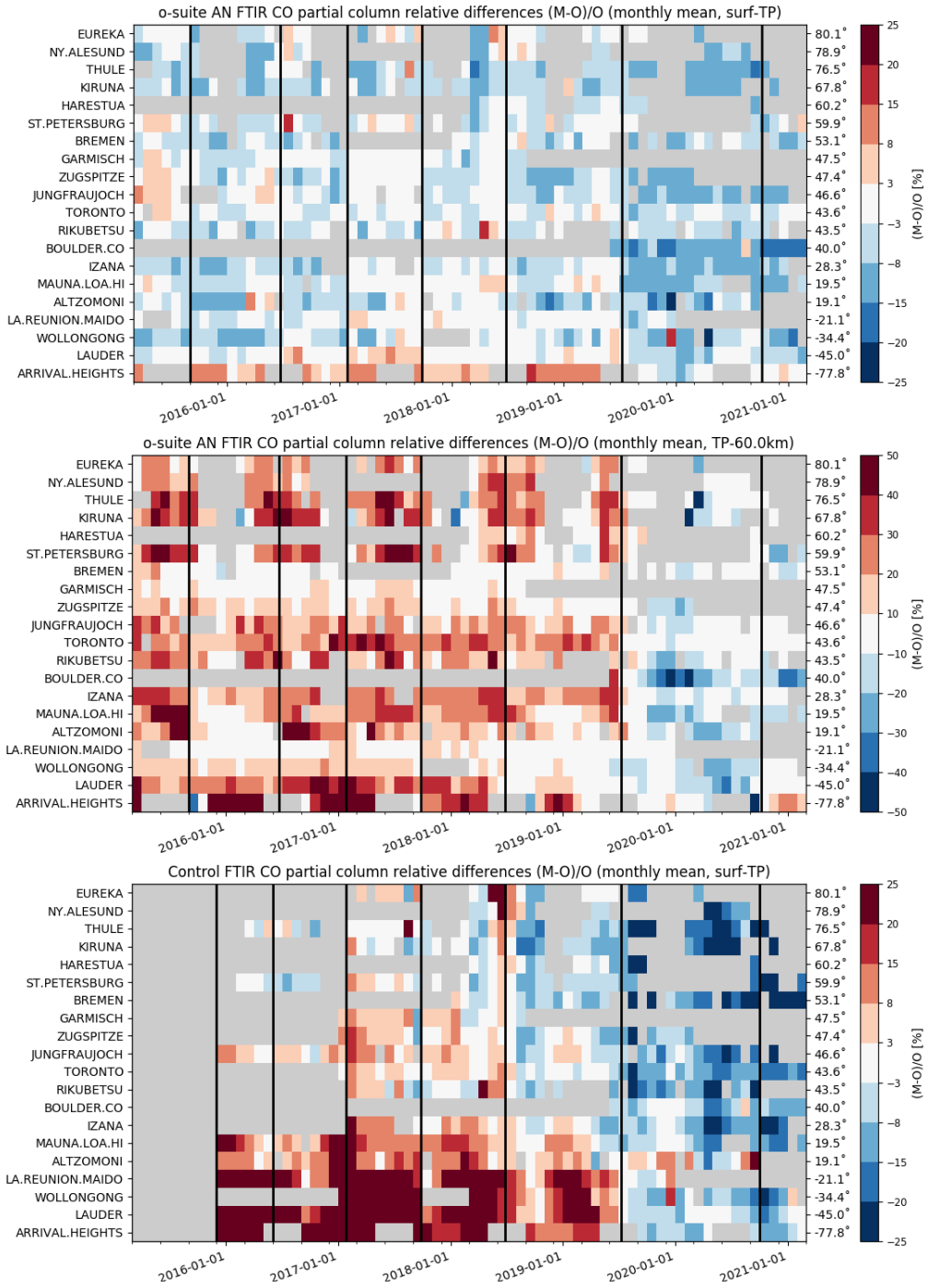


Figure 4.3.1: Monthly mean relative mean bias for tropospheric CO columns (MB, %) for the considered period DJF 2021 (top) and monthly mean biases for a longer period for the tropospheric CO columns (second row) and stratospheric CO columns (third row) (o-suite upgrades are indicated in black vertical lines, stations are sorted by latitude). The overall uncertainty for the CO measurements is approximately 3% on the tropospheric columns and 10% for the stratospheric columns. The o-suite analysis averaged bias in tropospheric columns increased to -6% for SON/DJF compared to -2% bias before the model update in July 2019. The bias in the stratosphere reduced to -8.5% compare to +18% before July 2019 and is comparable to the measurement's uncertainty. The underestimation of the control run (bottom) since July 2019 increases further to values reaching -20% in the NH.

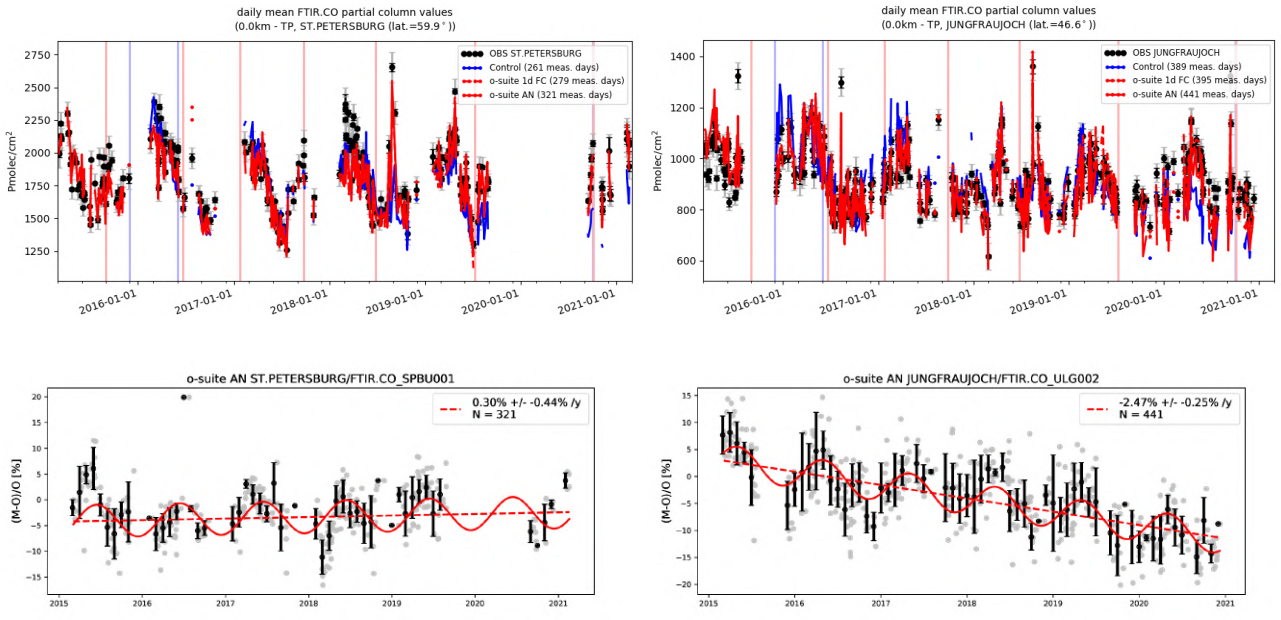


Figure 4.3.2: Top: daily mean values of tropospheric CO columns by the o-suite (AN and 1d FC, red) and the control run (blue) compared to NDACC FTIR data at St Petersburg and Jungfraujoch for the period March 2015-February 2021. During March 2018 the o-suite underestimated the CO columns at St. Petersburg. Bottom row contains a linear fit and seasonal cycle fit through the relative differences for the o-suite AN. An underestimation is observed during the local autumn/winter months. The negative trend at Jungfraujoch is -1%/y in the o-suite 1dFC and 2.5% in the o-suite AN (at first sight, the upgrade in June 2019 seems to have flattened the trend).

The Taylor diagrams in Figure 4.3.3 provide information on the correlation of all three CAMS products under consideration with the FTIR time series. Leaving out the sites with few measurements, the assimilation has a positive effect on the correlation coefficient. Looking at the correlation values for the period DJF 2021, the o-suite 1d FC (averaged correlation for all sites is 0.77) is slightly worse to the o-suite AN (averaged correlation for all sites is 0.83).

4.4 Validation against FTIR observations from the TCCON network

CO column averaged mole fractions of the CAMS models are compared with data from the Total Carbon Column Observing Network (TCCON). Column averaged mole fractions provide different information content than the in-situ measurements and are therefore complementary to the in-situ data. In this section, we compare column averaged mole fractions of CO of the CAMS models with TCCON retrievals. Data from the following TCCON sites has been used:
Izana (Blumenstock et al., 2017), Reunion (De Mazière et al., 2017), Bialystok (Deutscher et al., 2019), Manaus (Dubey et al., 2017), Four Corners (Dubey et al., 2017), Ascension (Feist et al., 2014), Anmeyondo (Goo et al., 2017), Darwin (Griffith et al., 2017), Wollongong (Griffith et al., 2017), Karlsruhe (Hase et al., 2017), Edwards (Iraci et al., 2017), Indianapolis (Iraci et al., 2017), Saga (Kawakami et al., 2017), Sodankyla (Kivi et al., 2017), Hefei (Liu et al., 2018), Tsukuba (Morino et al., 2017), Burgos (Morino et al., 2018), Rikubetsu (Morino et al., 2017), Bremen (Notholt et al., 2017),

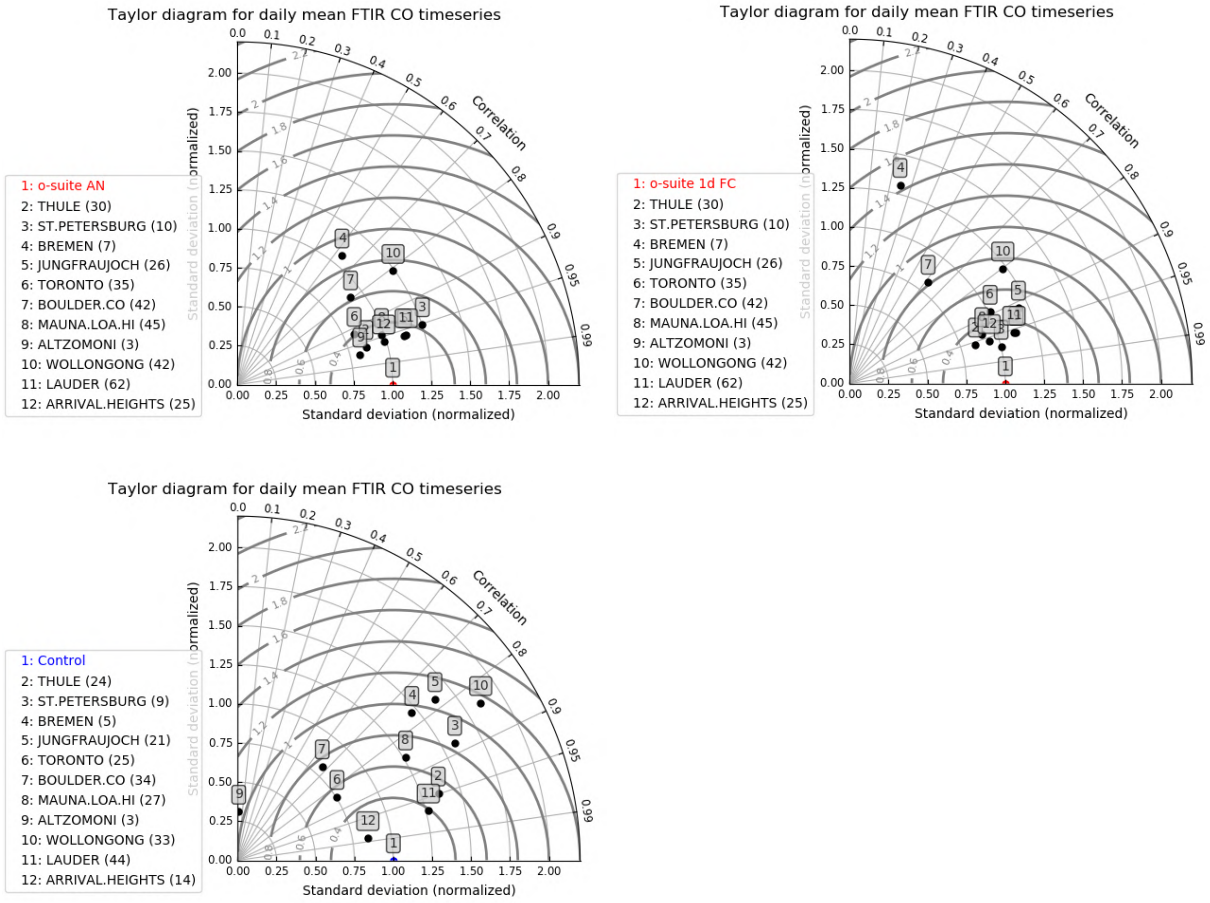


Figure 4.3.3: Taylor diagrams relating the standard deviations for the model /GB time series of tropospheric CO column data and their correlation during DJF 2021. All time-series are normalized such that the std of the model is 1. Bremen has low correlation in this quarter because the number of measurements is low.

Spitsbergen (Notholt et al., 2017), Lauder (Sherlock et al., 2017, Pollard et al., 2019), Eureka (Strong et al., 2018), Garmisch (Sussmann et al., 2017), Zugspitze (Sussmann et al., 2018), Paris (Te et al., 2017), Orleans (Warneke et al., 2017), Park Falls (Wennberg et al., 2017), Caltech (Wennberg et al., 2017), Lamont (Wennberg et al., 2017), Jet Propulsion Laboratory (Wennberg et al., 2017), East Trout Lake (Wunch et al., 2017), Nicosia (Petri et al., 2020)

For the validation of the models in December, January, and February the only site that made data available for the whole comparison period was Nicosia.

For the comparison period the o-suite analysis and the o-suite forecast simulations compare very well with the measurements (Fig 4.4.1). The control model underestimates the CO (Fig. 4.4.1 and Fig. 4.4.3) and does not capture the seasonality.

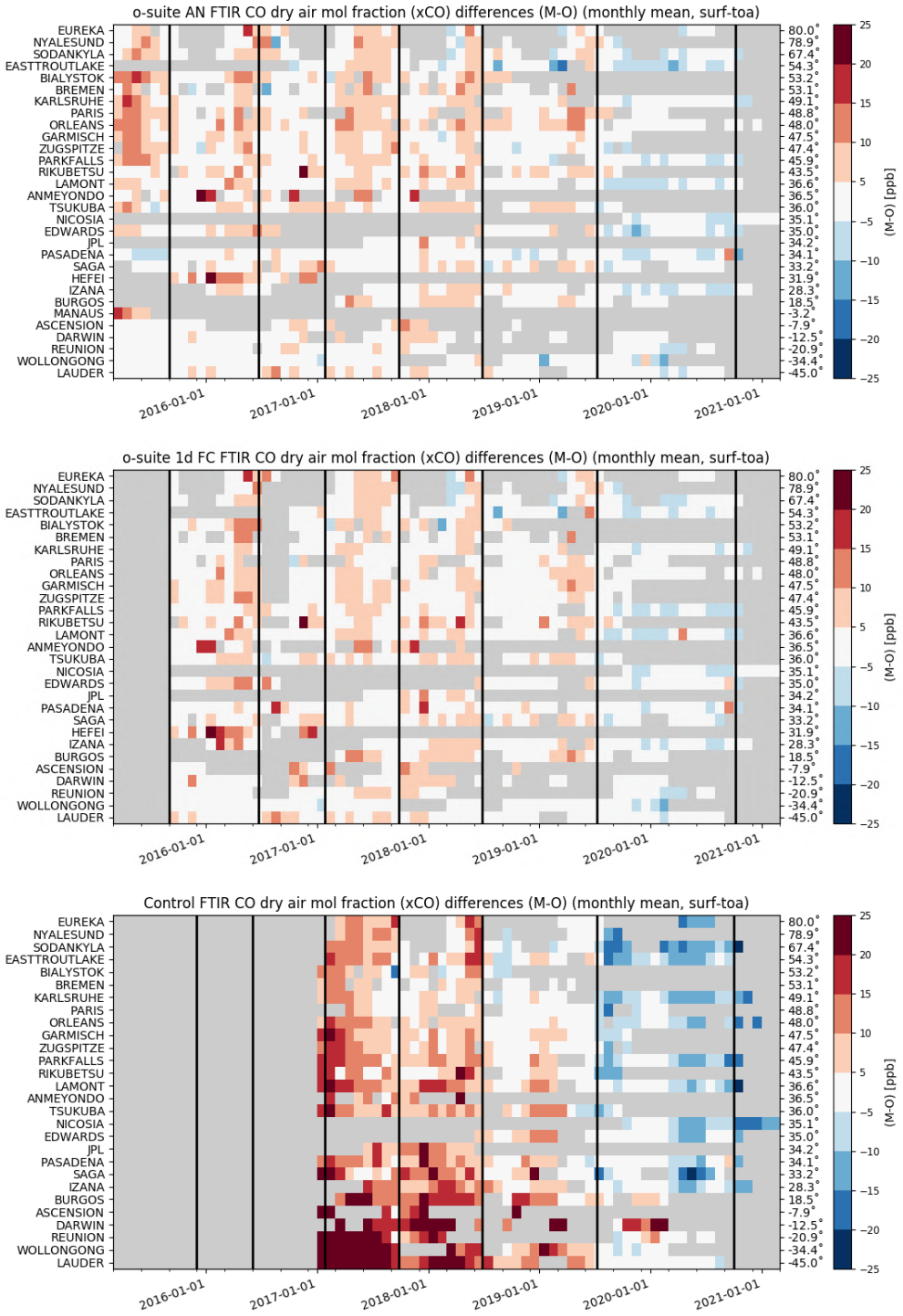


Figure 4.4.1: Monthly CO differences since March 2017, for the o-suite analysis (top), o-suite 1-day forecast (middle) and control run (bottom) compared to TCCON observations. The stations are sorted by latitude (northern to southern hemisphere).

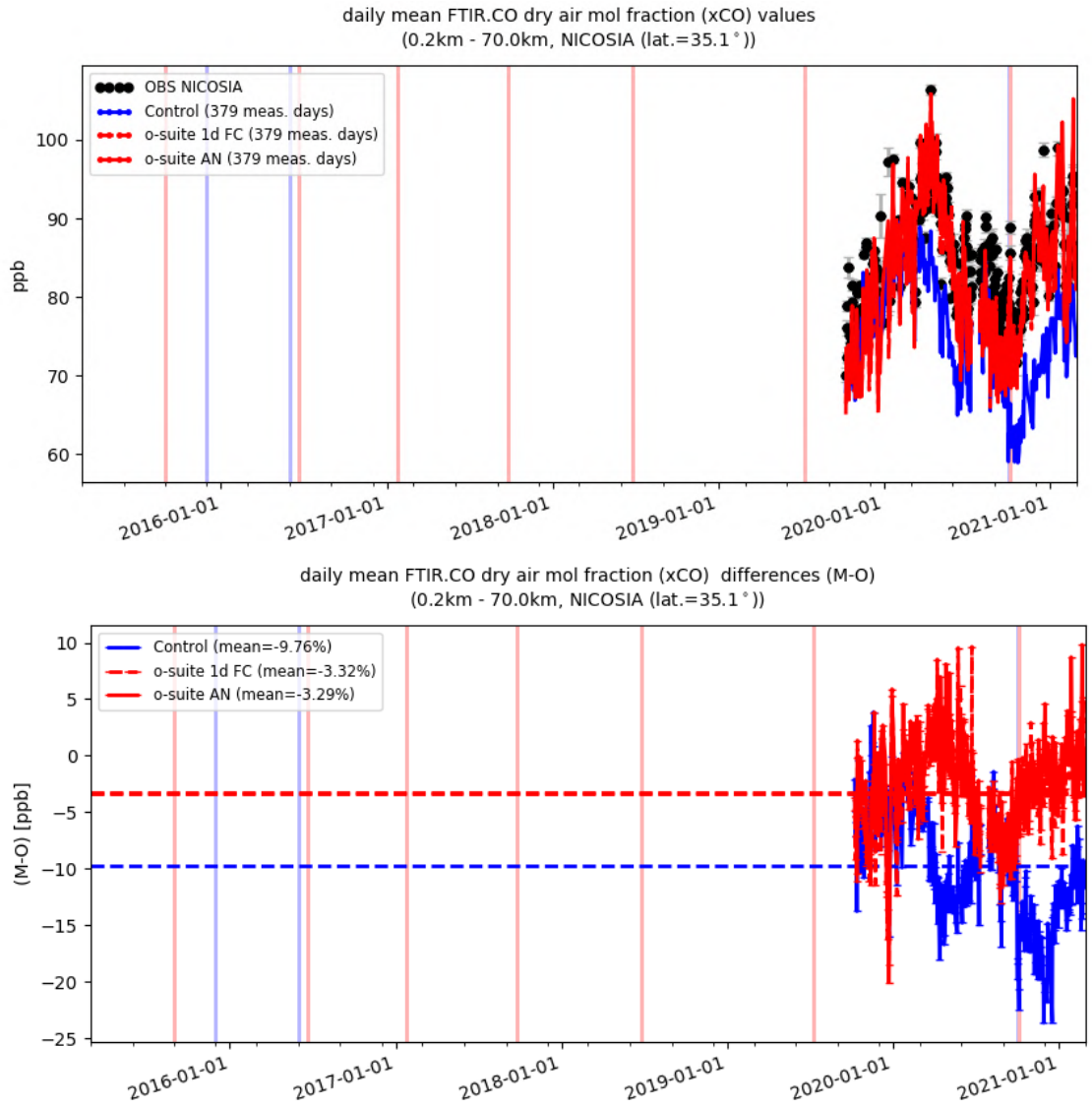


Figure 4.4.2: Comparison of the CO CAMS data with TCCON CO at Nicosia.

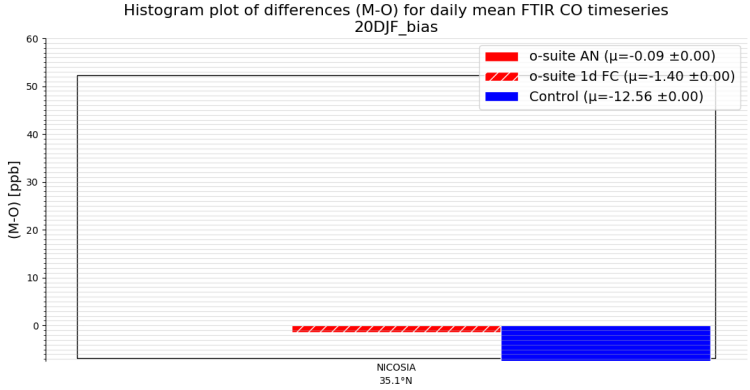


Figure 4.4.3: Differences during the reporting period for the TCCON site Nicosia.

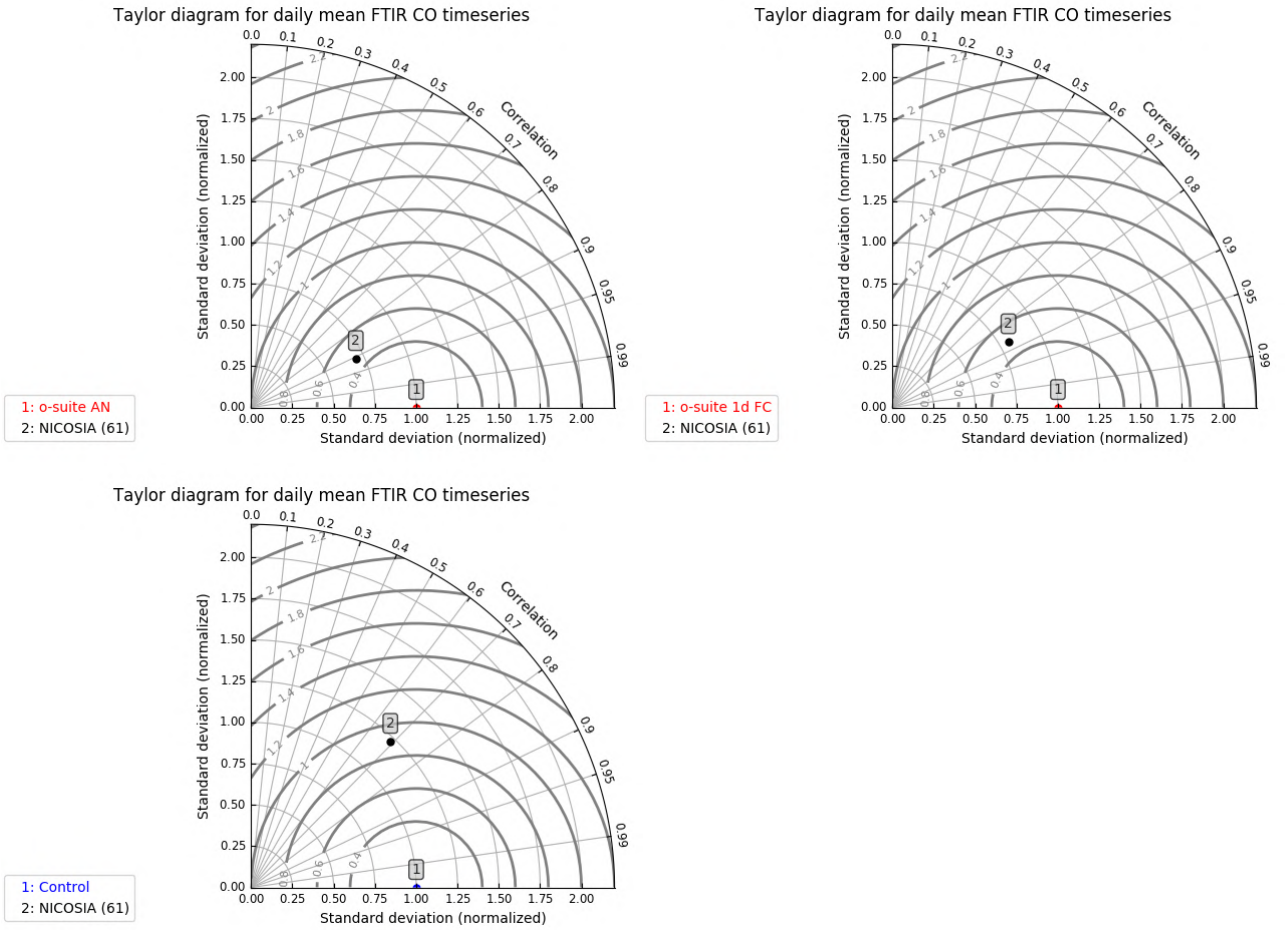


Figure 4.4.4: Taylor diagrams for the comparison period, December 2020 – February 2021..

4.5 Evaluation with MOPITT and IASI data

In this section, modelled CO total columns are compared to MOPITT version 8 (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 V8 (top left) and IASI (top right) and the relative bias of the model runs with respect to MOPITT V8 (analysis and forecast day4) (middle) for January 2021. MOPITT shows high values over the biomass burning area in Africa and over East Asia. IASI shows higher values over the above-mentioned regions.

The modelled CO geographical distribution and magnitude of values show that the o-suite performs well. The relative difference between the model runs and MOPITT shows that the o-suite performs better than the control run without data assimilation. The o-suite mainly underestimates the satellite data by about 10% with some regional exceptions where the negative bias reaches 20% (e.g., over Australia, south of Southern America). The model shows overestimation over the areas with high CO values mentioned above (up to 20%). The control run underestimates/overestimates ozone over the Northern/Southern Hemisphere up to 30%. The o-suite run shows a growing negative bias on the 4th forecast day mainly over Northern Africa and south of Southern America and a growing positive bias over Indonesia, North of China, biomass burning area in South America and other areas.

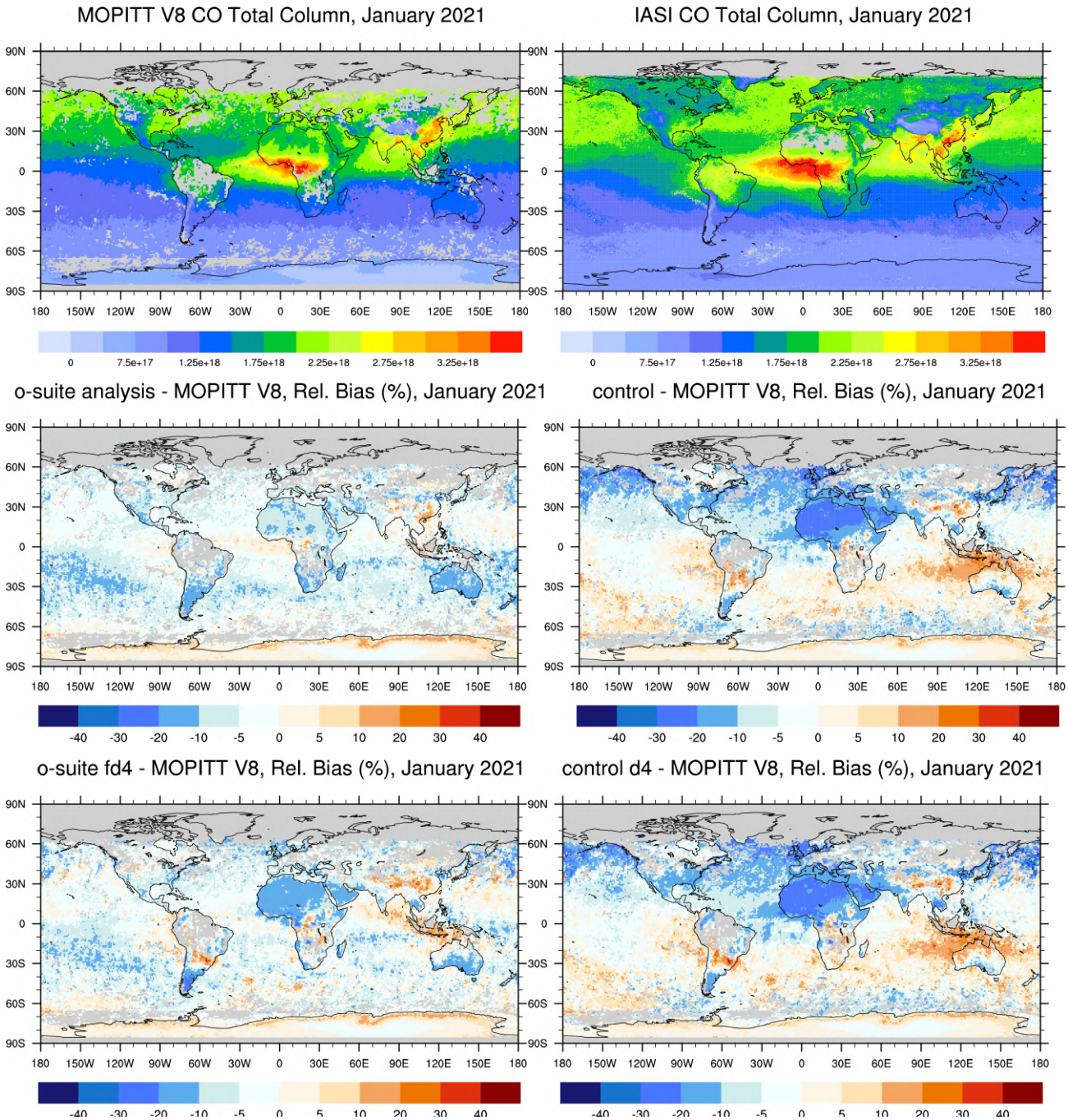


Fig. 4.5.1: CO total columns for MOPITT V8 (top left) and IASI (top right) satellite retrievals and relative difference between the model runs and MOPITT for January 2021: o-suite analysis (middle left), control run (middle right), o-suite 4th forecast day (middle left), o-suite 4th forecast day (middle right). Grey colour indicates missing values.

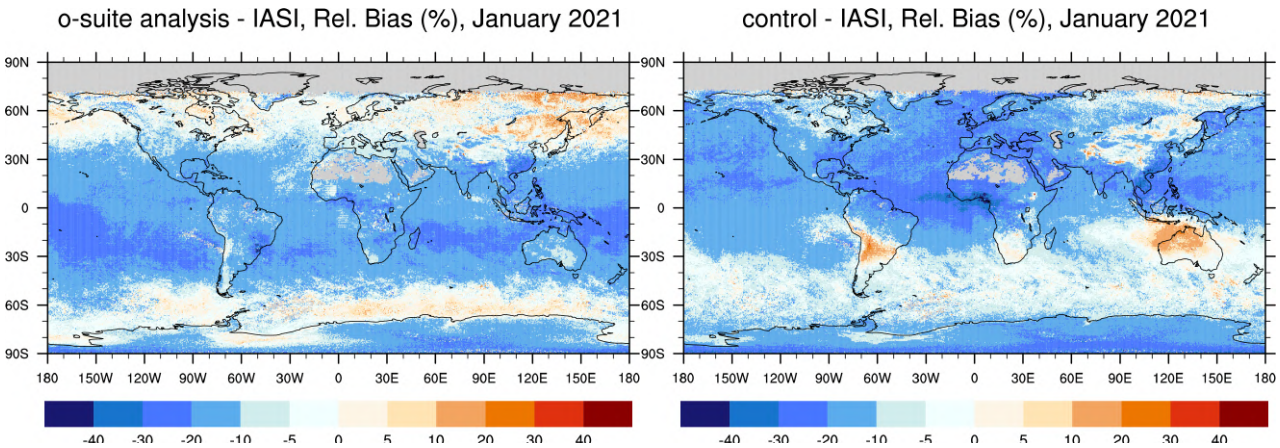


Fig. 4.5.2: Relative difference between the model runs and IASI for January 2021: o-suite analysis (bottom left), control run (bottom right). Grey colour indicates missing values.

In comparison to IASI data (Fig. 4.5.2), the o-suite run shows underestimation over the low latitudes and southern mid-latitudes up to 30% and overestimation over east of Russia and over Southern Ocean up to 30%. The control run mainly underestimates the observations up to 30% with exceptions of Australia and central part of South America up to 20%.

Figure 4.5.3 shows time series of CO total column for MOPITT V8, IASI and the model runs over the eight selected regions. For the comparison with MOPITT, the modelled CO concentrations were transformed using MOPITT V8 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 until 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 being 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 run shows that the assimilation of satellite CO has a 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 biases less than 5%. In late summer and early autumn of 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.

A general reduction of CO values from the year 2015 to the year 2018 can be seen over Europe, the US, and East Asian regions. The South African region shows a slight increase of the seasonal minimum compared to previous springs. Summer 2019 was characterised by a strong fire events in Siberia. This can be seen in IASI data (peak in August), but it is not reflected in the MOPITT data partly due to only few days of observations available in August.

In December 2020 – February 2021, the CO concentrations over Europe, East and South Asia and Siberian and Alaskan fire regions are increased compared to previous year.

The modified normalized mean bias (MNMB) of the model runs compared to MOPITT V8 (Fig. 4.5.4) allows quantifying the impact of the assimilation on the model performance. In the end of 2020, beginning 2021, the o-suite model run shows negative biases over all selected regions within 6%. The better agreement can be seen over Siberian fire region with bias within 2%. The growing negative bias can be seen over North Africa in February 2021, which reached 15%.

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.

Starting from the second half of the year 2019, the negative biases over Europe and US increase for both runs (from about 5% to about 10% for o-suite). The o-suite results over the Asian regions improved and show better agreement with the observations. The control run shows reduction of biases over South Asia. A change of bias sign from positive to negative and/or increase of the negative bias can be seen over all selected regions for both runs. For the control run, the strong increase of the negative bias in September-October 2020 can be seen over Europe, US, Siberian and Alaskan fire regions and over South African region and over North Africa in February 2021.

In general, the increase of an underestimation in both runs can be seen over the selected regions.

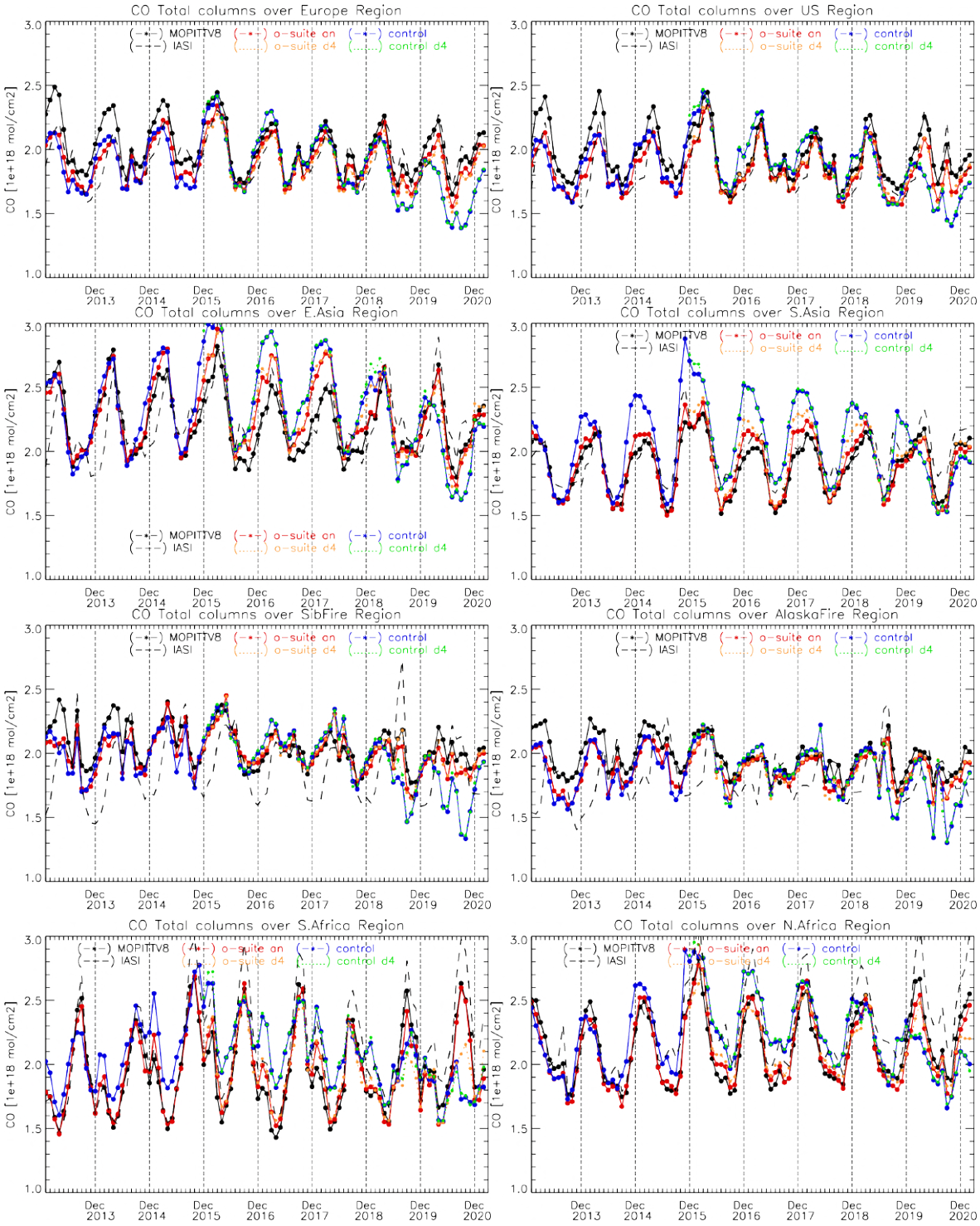


Fig. 4.5.3: Time series of CO total columns for satellite retrievals MOPITT V8, IASI (black) and the model runs over the selected regions: o-suite analysis (red, solid), control (blue, solid), o-suite 4th forecast day (orange, dotted), control 4th forecast day (green, dotted). Period: January 2013 to February 2021.

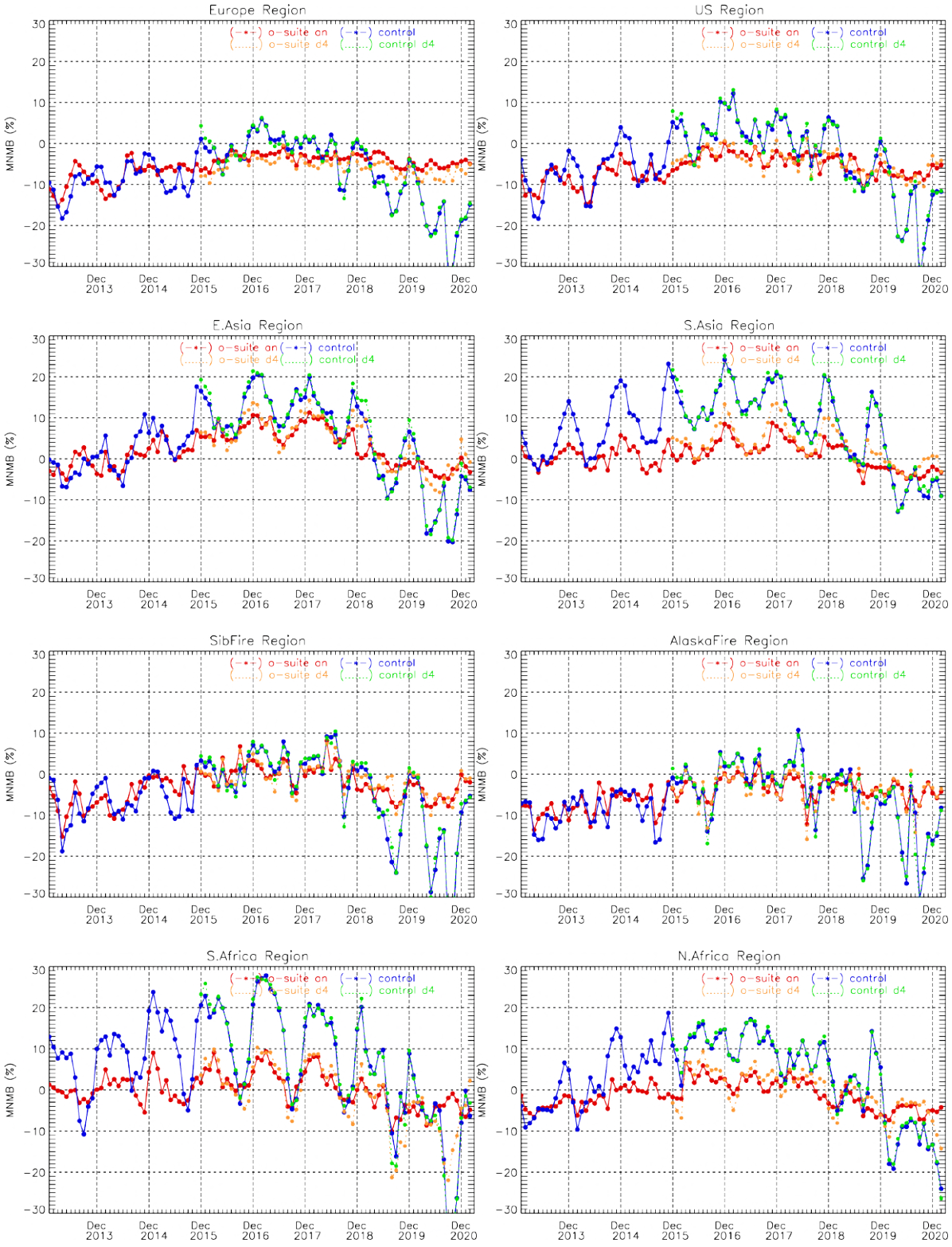


Fig. 4.5.4: Timeseries of modified normalized mean bias (%) for CO total columns from the model simulations vs MOPITT V8 retrievals over selected regions. O-suite analysis (red, solid), control run (blue, solid), o-suite 4th forecast day (orange, dotted), control 4th forecast day (green, dotted). Period: January 2013 to February 2021.

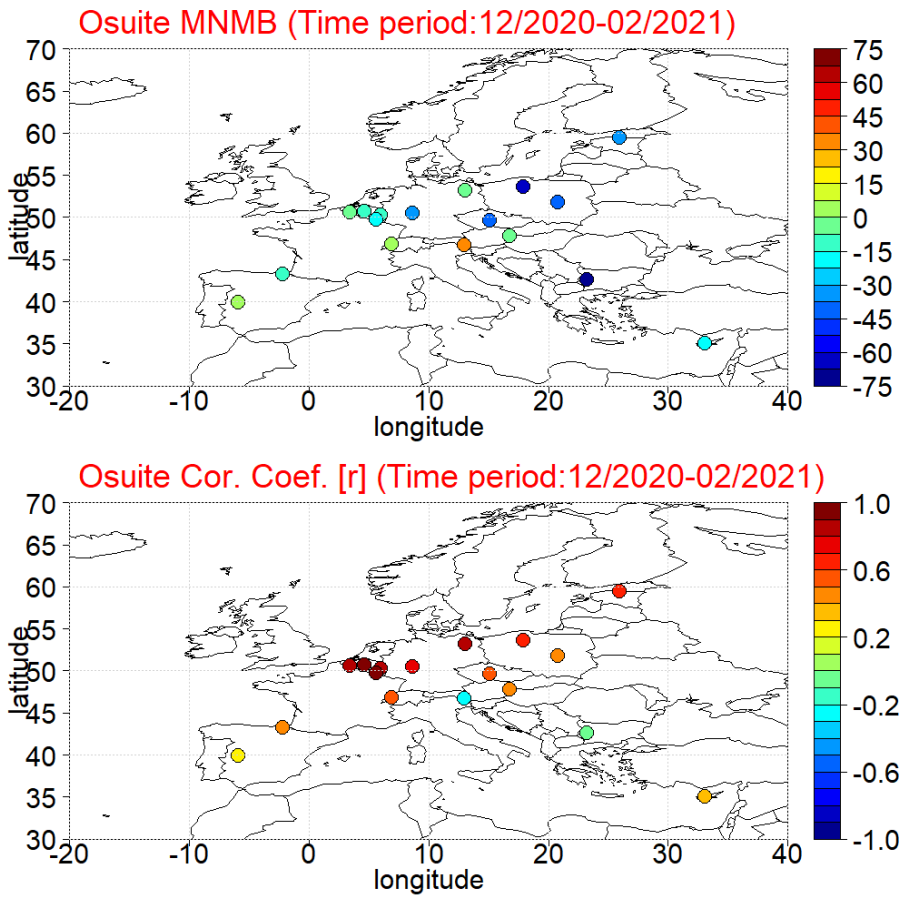


Figure 4.6.1: Spatial distribution of MNMB in % (left) and correlation coefficient (right) of the o-suite run compared to observational data during the period from 1 December 2020 to 28 February 2021.

4.6 Evaluation with CO surface observations over Europe

The surface carbon monoxide validation analysis over Europe is based on an evaluation against background rural classes 1 to 5 of the Joly-Peuch classification (Joly and Peuch, 2012). The station observations are taken from the Airbase Network database at EEA (http://acm.eionet.europa.eu/databases/airbase/). In addition, 1 station from the Department of Labour Inspection - Ministry of Labour and Social Insurance, of Cyprus (http://www.airquality.dli.mlsi.gov.cy/) is used in the validation analysis.

The spatial distribution of bias and correlation coefficients of the o-suite over Europe are shown in Fig. 4.6.1. The results show that correlations over almost all CO European AirBase stations are highly significant (0.3<r<0.9) with the exceptions of Byana (Bulgaria) and Kärnten (Austria) where correlations drop close to zero. Concerning biases CAMS o-suite underestimate surface carbon monoxide mean concentrations over Belgium (depending on the station MNMBs vary between -5% and -20%) and Germany (depending on the station MNMBs vary from -2% to -30%), and underestimate more over Estonia (MNMB=-35%), Czech Republic (MNMB=-40%), Poland (depending on the station MNMBs vary between -40% and -60%) and Bulgaria (MNMB=-95%). In Austria and Switzerland stations o-suite underestimate surface carbon monoxide mean

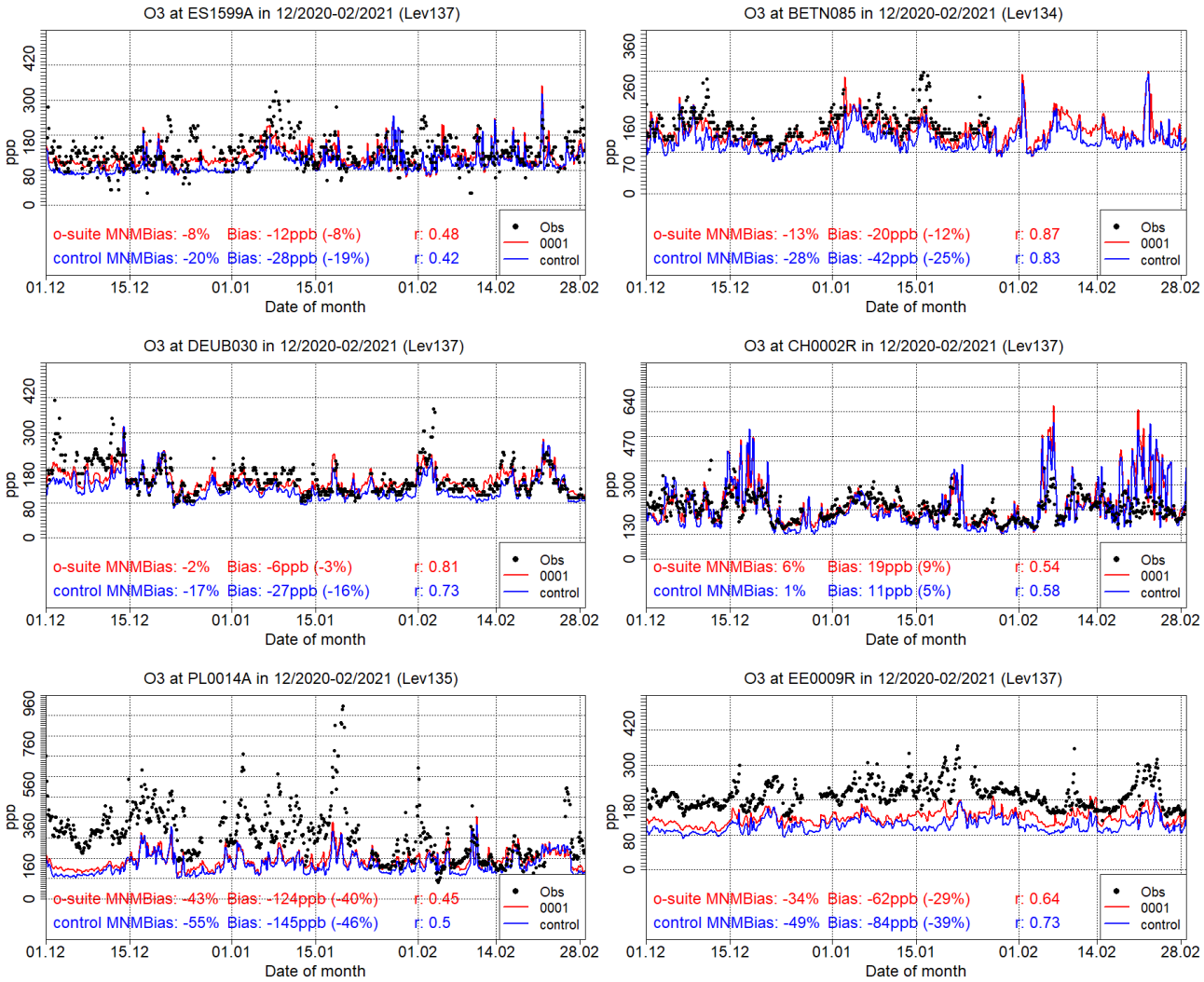


Figure 4.6.2: Time series for the o-suite (red) and control run (blue) compared to Airbase observations at Pagoeta, Spain station (43.25°N, 2.15 °W, 1st row left), at Vielsalm, Belgium station (50.30°N, 6.02°E, 1st row right), at Neuglobsow, Germany (53.14°N, 14.02°E, 2nd row left), at Payerne, Swiss station (46.81°N, 6.94°E, 2nd row right), at Belsk, Poland station (51.84°N, 20.79°E 3rd row left), and at Lahemaa, Esthonia (59.49°N, 25.93°E, 3rd row right)

concentrations over Illmitz in Austria (MNMB=-5%) and overestimate it Rigi in Switzerland (MNMB=6%) and over Kärnten in Austria (MNMB=32%). O-suite MNMBs vary between -8% and +6% over Spain while over Cyprus o-suite underestimate surface carbon monoxide mean concentrations by -22%. These findings are further illustrated by the timeseries at selected stations plotted in Figure 4.6.2. The control run surface carbon monoxide concentrations are 10-20 ppb (5%-15%) lower than the o-suite, resulting in a stronger negative bias in all stations (exceptions are Kärnten station in Austria and Rigi station in Switzerland).

5. Tropospheric nitrogen dioxide

5.1 Evaluation against GOME-2 and TROPOMI retrievals

In this section, model columns of tropospheric NO_2 are compared to TROPOMI/Sentinel-5P data (IUP-UB v0.9, preliminary) and to GOME-2/MetOp-C (IUP v0.9, preliminary), using the CAMS o-suite as a-priori in the retrievals. The 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 09:30 LT for the GOME-2 instruments and 13:30 LT for TROPOMI and at clear sky only. Therefore, model data are vertically integrated, interpolated in time and then sampled to match the satellite data. The satellite data were gridded to model resolution (currently 0.4° x 0.4° degree). For the comparisons to TROPOMI and GOME-2C satellite data, the stratospheric contribution has been removed from the measurements using STREAM-B which is an IUP-Bremen version of the STREAM algorithm by Beirle et al. (2016). In the current version of STREAM-B, the free tropospheric contribution is not yet well accounted for, which leads to a negative offset in the current preliminary TROPOMI and GOME-2C data versions and will be improved by the addition of tropospheric background values in the near future. 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. Systematic uncertainties in regions with significant pollution are on the order of 20% – 30%.

Figure 5.1.1 shows global maps of monthly mean tropospheric NO_2 columns from TROPOMI and GOME-2C and the model runs as well as differences between retrievals and simulations for February 2021 as an example for the last winter. The overall spatial distribution and magnitude of tropospheric NO_2 is reproduced by both CAMS runs, indicating that emission patterns and NO_x photochemistry are reasonably well represented. However, a systematic overestimation over anthropogenic pollution is visible in the TROPOMI and GOME-2C based map comparisons. The shipping signals simulated by the models are generally larger than the observed ones. For example, shipping signals are more pronounced in model simulations to the south of India.

The TROPOMI IUP Bremen data product shows lower background values compared to the operational offline product, this is expected to be reduced with the next data version of the Bremen product (see data description above). Differences in comparison results between the sensors are in principle due to differences in observation time and the retrieval products.

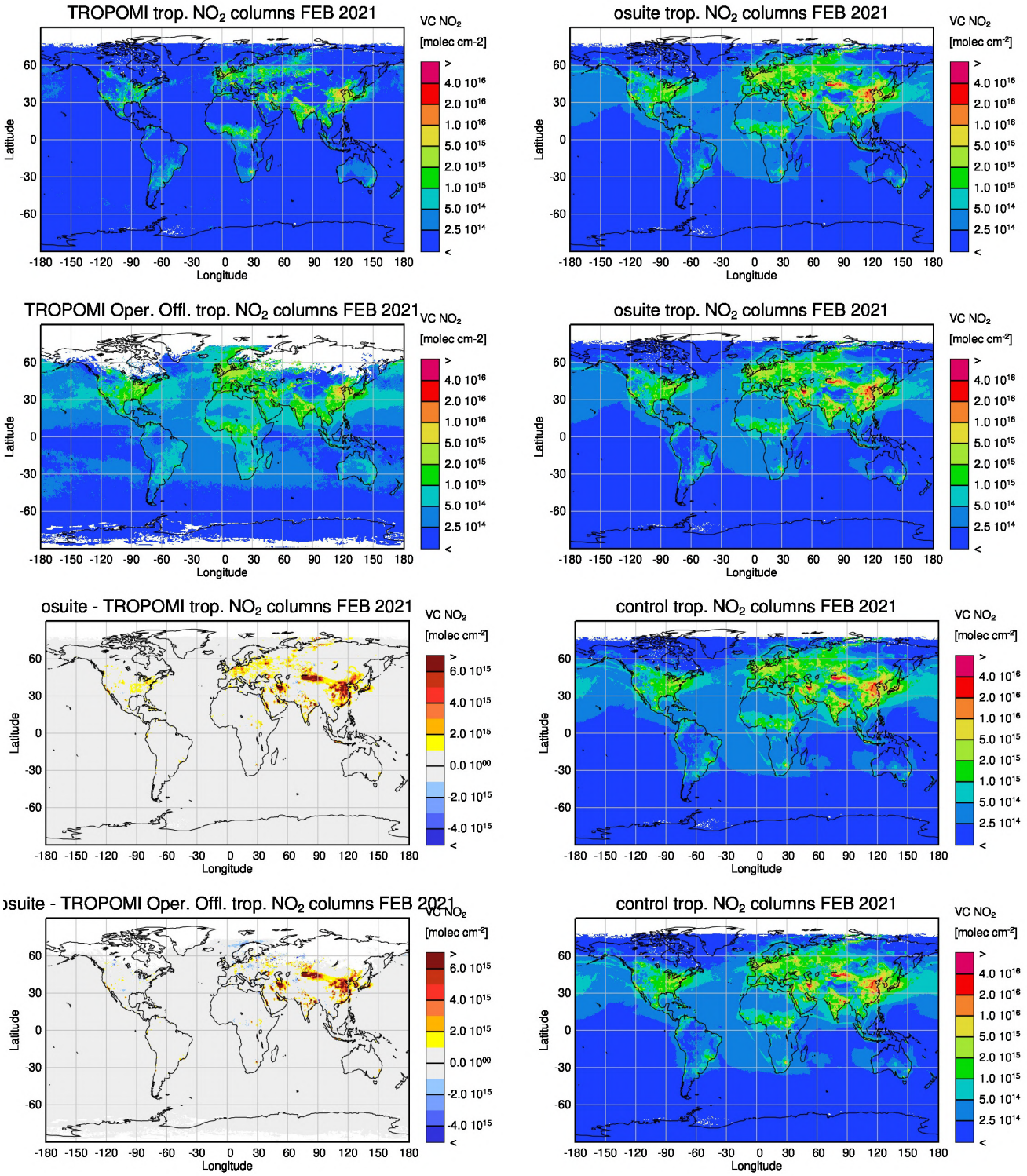


Figure 5.1.1a: Global map comparisons of satellite-retrieved and model simulated tropospheric NO_2 columns [molecules cm^{-2}] for February 2021 based on TROPOMI Bremen (first, third row) and TROPOMI operational offline (second, fourth row). The top rows show the satellite observations (left) and o-suite (right), the bottom rows the difference between o-suite and satellite observations (left) , and the control run (right). The satellite data were gridded to model resolution (i.e., 0.4° x 0.4° degree) and the CAMS o-suite was used as a-priori in the TROPOMI IUP-UB retrievals only.

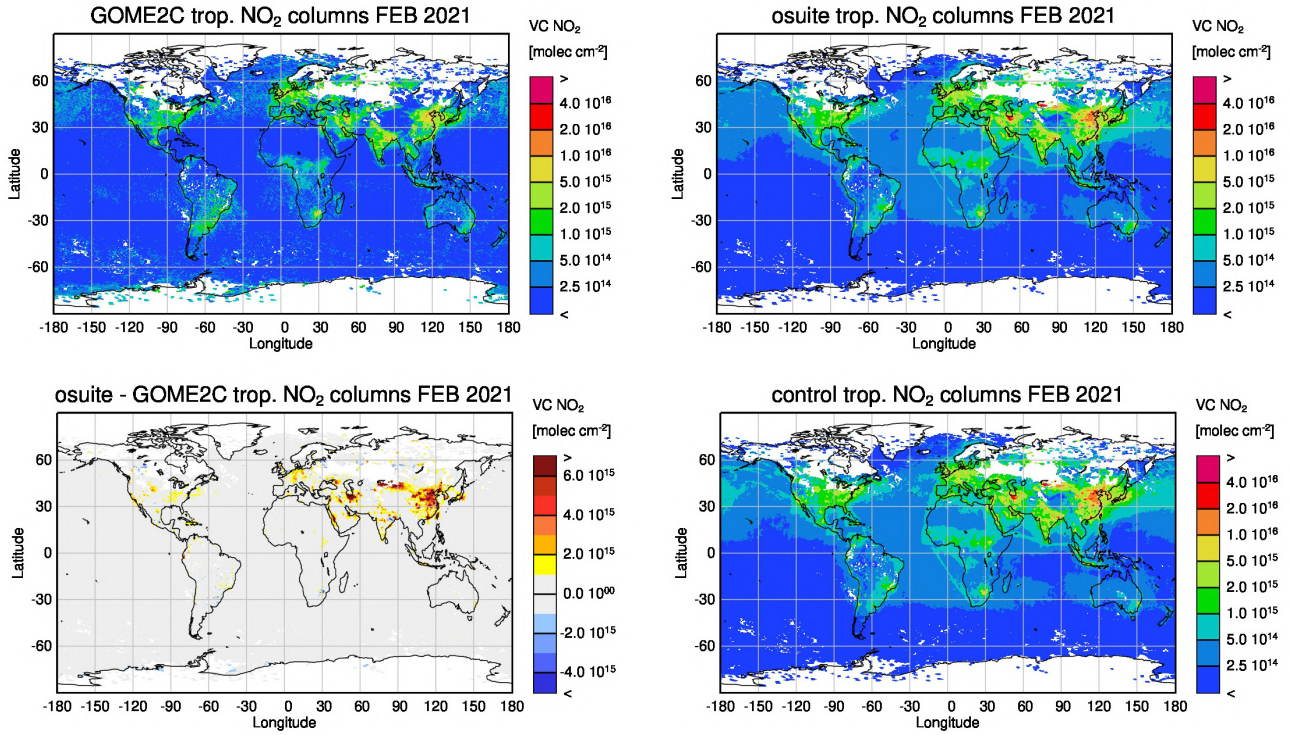


Figure 5.1.1b: Global map comparisons of satellite-retrieved and model simulated tropospheric NO_2 columns [molecules cm^{-2}] for February 2021 based on GOME-2C. The columns show (from top left to bottom right) satellite observations, o-suite, the difference between o-suite and satellite observations, control run. The satellite data were gridded to model resolution (i.e., 0.4° x 0.4° degree) and the CAMS o-suite was used as a-priori in the GOME-2C retrievals.

Time series comparisons between the o-suite and TROPOMI as well as GOME-2C are shown in Figure 5.1.2 for data since January 2019. The model runs are in general positively biased compared to the TROPOMI and GOME-2C retrievals. Only over North- and South-Africa the simulations show smaller values than the TROPOMI operational offline product (but not compared to the IUP-UB TROPOMI and GOME-2C retrievals). For 2020, the TROPOMI observations show the peak in the time series over South-Africa for September, while the o-suite and control run simulate the peak for July/August. Apart from this, the occurrence of maxima and minima due to seasonality is reproduced by both model runs, and for both the regions dominated by anthropogenic emissions and those dominated by biomass burning. However, the magnitude of seasonality is largely overestimated over East-Asia and Eastern-US, with the o-suite showing better results than the control here. As described above, differences in comparison results are in principle due to differences in observation time or differences in the retrieval products.

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

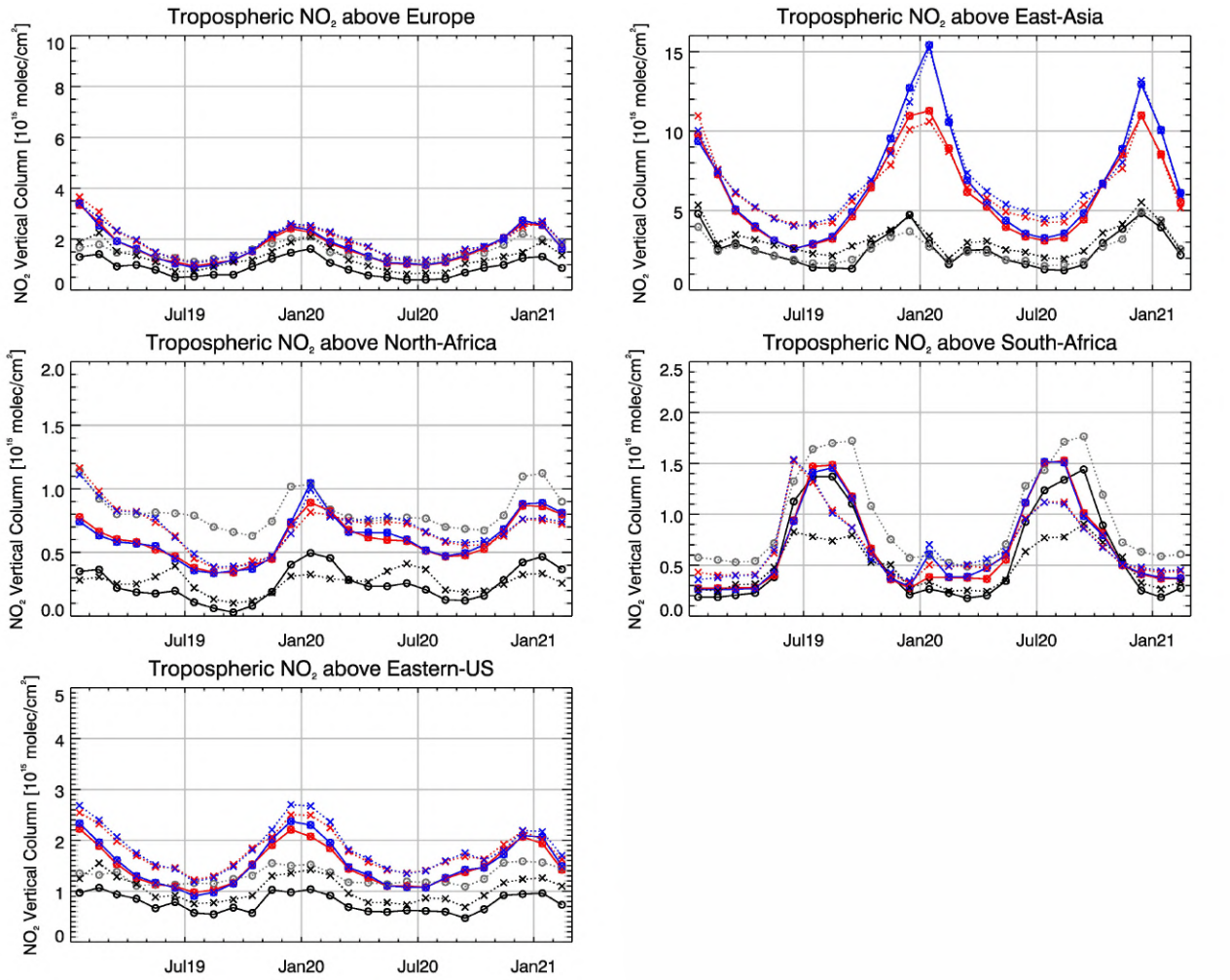


Figure 5.1.2: Time series of average tropospheric NO_2 columns [10^{15} molec cm^{-2}] from (black and grey) satellite retrievals, (blue) control and (red) o-suite model results between Jan 2019 and Feb 2021. The solid lines with circles show comparisons based on TROPOMI (in black the IUP-Bremen product and in grey the operational offline product), the dotted lines with crosses show comparisons for GOME-2C. The upper panels and the lowest panel represent regions dominated by anthropogenic emissions, and the panels in the middle represent those dominated by biomass burning.

5.2 Evaluation against ground-based DOAS observations

In this section, we compare the NO_2 columns of the CAMS products with UVVIS DOAS profile measurements at Xianghe 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%. 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 for all DOAS OFFAXIS measurements, see 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

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

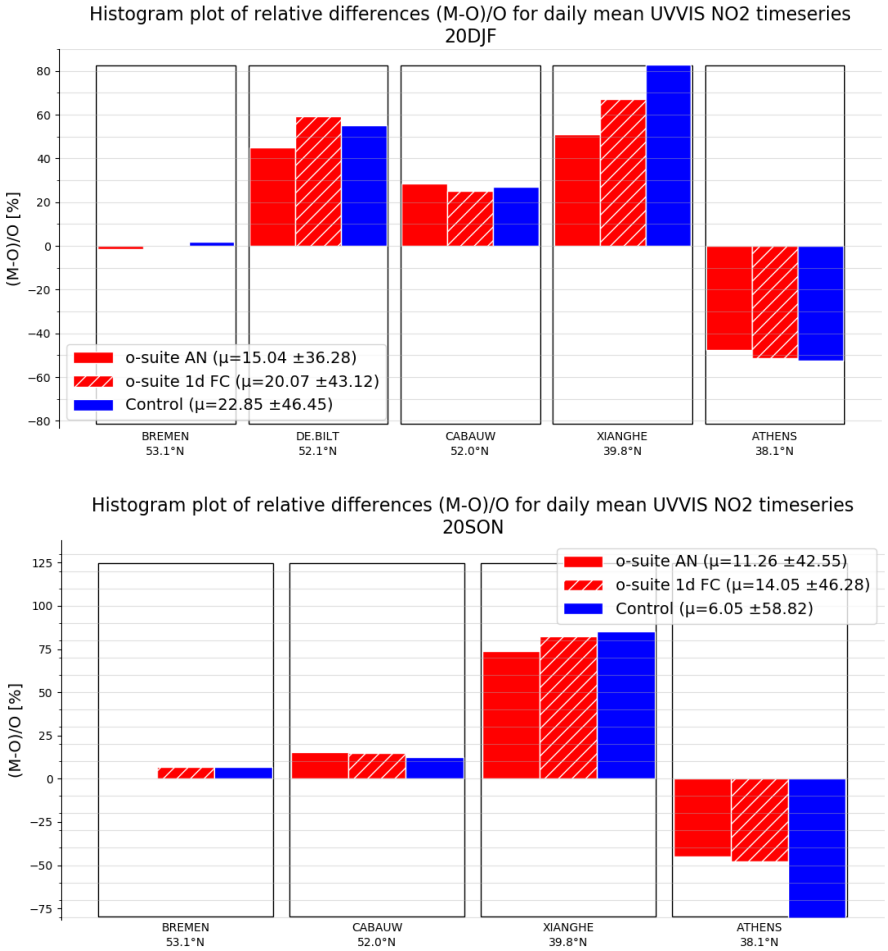


Figure 5.2.1: Table diagram showing the seasonal bias September-November 2020 (top) and Dec-February 2021 (bottom) for five stations, sorted by latitude.

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.

Figure 5.2.1 shows the biases for the latest validation periods Sept-Nov 2020 and Dec-Feb 2021 at the different sites. The corresponding time series are shown in Fig. 5.2.2. The o-suite is able to capture only few of the high pollution events for Bremen, De Bilt and Cabauw. In Athens no high pollution events are captured by the o-suite.

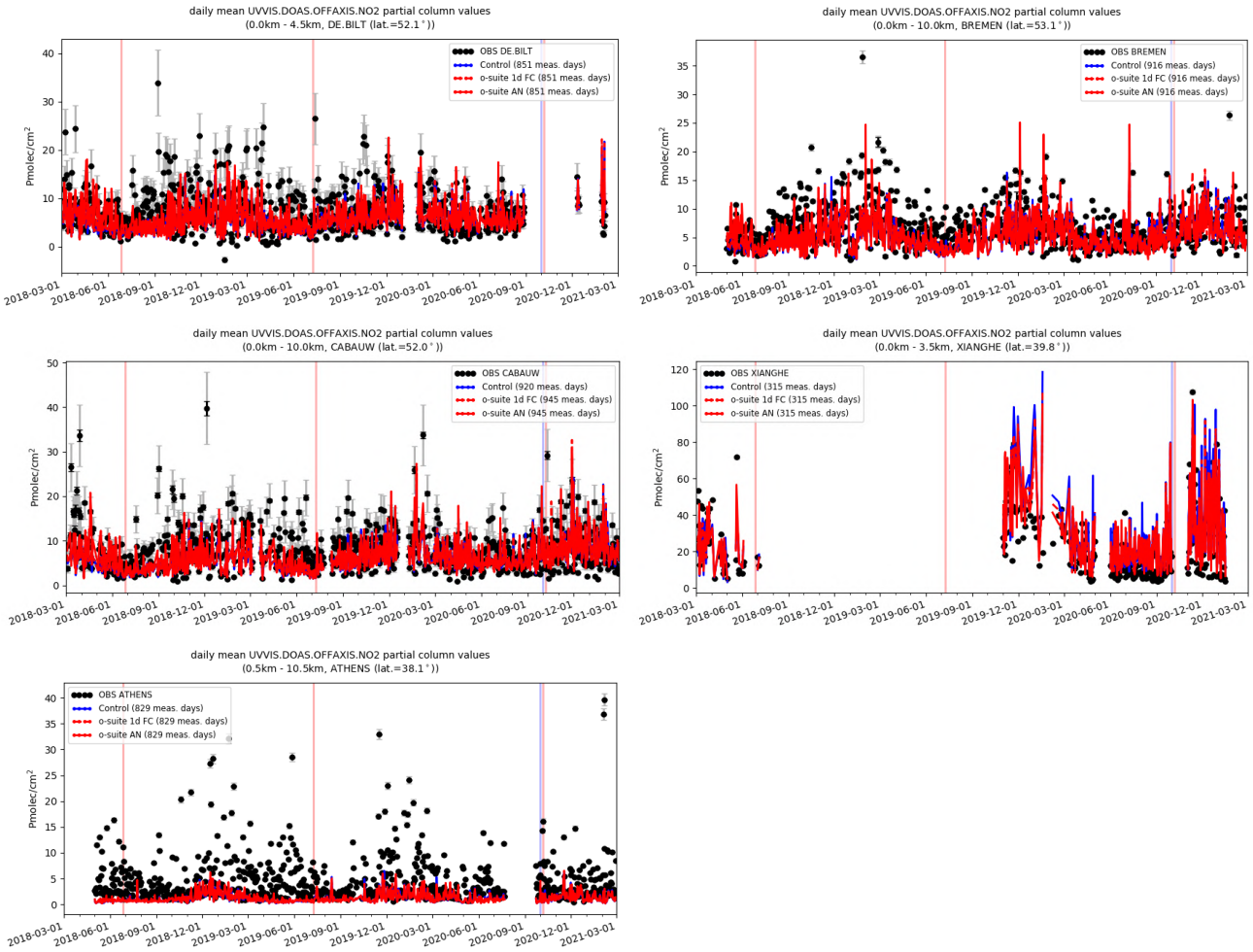


Figure 5.2.2: Time series of NO_2 partial columns at the five different sites. For all sites except Athens, background concentrations are well captured by the CAMS products. The o-suite and control runs show little difference.

5.3 Evaluation against surface nitrogen dioxide observations over China

The surface NO_2 validation over China is based on station observations from more than 1,500 in situ stations covering all major cities in China, operated by the China National Environmental Monitoring Center, reporting the pollutants PM10, PM2.5, O_3, NO_2, SO_2, and CO (e.g., Bai et al., 2020). The measurements were collected within the EU MarcoPolo and Panda projects. Individual station data was clustered for 37 megacities (e.g., 10-20 stations per city) and the observed surface NO_2 values are compared with the simulated NO_2 values calculated for the corresponding o-suite grid point.

Table 5.3.1 shows the names, coordinates, observed and simulated ozone values as well as validation metrics namely the MNMBs and correlations obtained for the o-suite run. The spatial distribution of MNMBs and the correlation coefficients of the o-suite over China are shown in Fig. 5.3.1, where it is evident that correlations over most megacities in the entire China (with few exceptions namely Kunming, Xiamen and Xian) are highly significant (0.32<r<0.73).

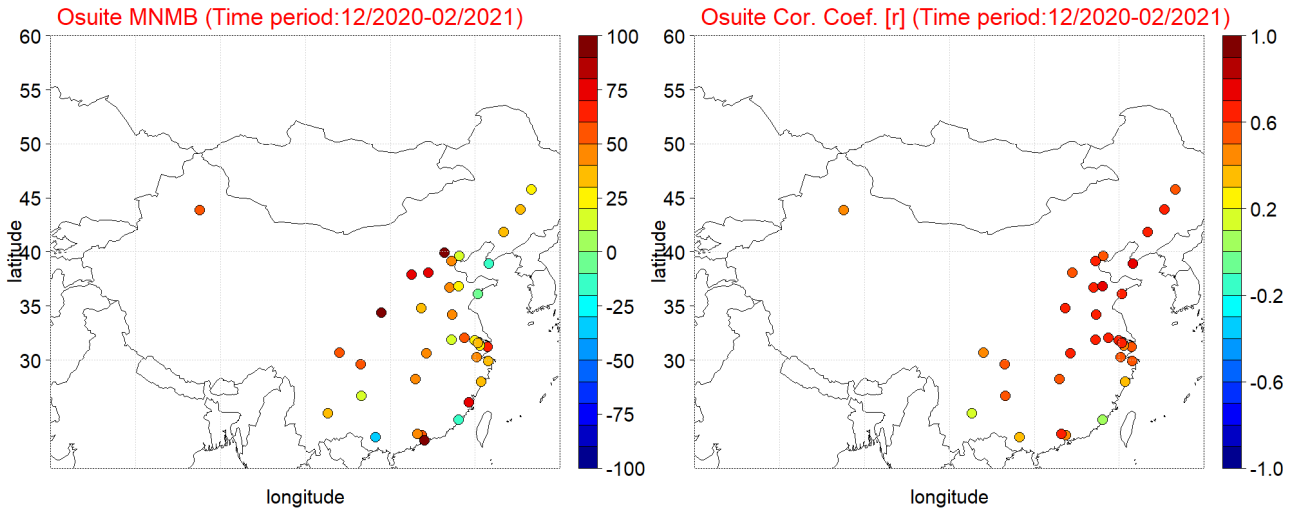


Figure 5.3.1: Spatial distribution of MNMB in % (left) and correlation coefficient (right) of the o-suite run compared to observational data during the period from 1 December 2020 to 28 February 2021.

The o-suite mostly overestimate surface NO_2 values. More specific o-suite MNMBs for NO_2 vary between 20% and 30% depending on the Megacity in North-western China while at Urumqi Megacity in North-eastern o-suite MNMB is 55%. For stations in the latitudinal belt 30°N-40°N the o-suite strongly overestimate surface NO_2 values with MNMBs varying between 20% and 75% (exceptions are Dalian, Hefei and Qingdao cities where o-suite MNMBs are close to zero as well as Beijing where o-suite MNMB exceed 110%). For the Megacities southern of the 30° N parallel, the o-suite MNMBs vary between 17% and 77% (exceptions are Nanning and Xiamen cities where o-suite underestimate surface NO_2 values by -36% and -18% respectively, and Shenzhen city where o-suite MNMB is almost 100%). The control run surface NO_2 values are within ±1.5 ppb compared to o-suite and correlations between control run NO_2 and observations are almost identical with the corresponding o-suite correlations.

The above-mentioned findings concerning CAMS o-suite biases and correlations are also observed in individual time series at selected cities plotted in Figure 5.3.2.

It should that the NO_2 overestimation from CAMS NRT runs explains in a significant degree the O_3 underestimation over China (see section 3.6). This is illustrated in figure 5.3.3 where it is evident that $MNMB_{NO2}$ and $MNMB_{O3}$ over the 37 under study Megacities are statistically significantly anticorrelated (r=-0.7).

Table 5.3.1: Names, coordinates, observed and simulated NO_2 values as well as validation metrics (MNMBs and correlations for the period DJF 2020-2021) obtained with the 2 forecast runs (o-suite and control), for each one from 37 China Megacities under study. MNMBs and correlations with blue denote stations where control run performs better while with red are denoted stations where o-suite performs better.

			NO_2 (ppb)			MNMB (%)		Cor. Coef.	
MegaCity	**Lat**	**Lon**	**Observed**	**o-suite**	**control**	**o-suite**	**control**	**o-suite**	**control**
Beijing	39.92	116.38	17.4	65.2	66.6	112.7	113.7	0.59	0.61
Changchun	43.89	125.33	16.6	23.5	24.1	31.3	33.4	0.62	0.61
Changsha	28.20	112.97	20.1	30.9	31.0	42.6	43.1	0.54	0.54
Changzhou	31.81	119.97	25.4	33.4	32.9	29.6	28.7	0.68	0.60
Chengdu	30.66	104.07	22.3	39.2	41.1	53.0	56.8	0.44	0.45
Chongqing	29.56	106.55	23.9	45.7	47.4	57.7	61.1	0.58	0.60
Dalian	38.91	121.60	16.5	15.1	15.4	-19.8	-17.6	0.72	0.76
Dongguan	23.02	113.75	20.0	38.1	37.9	55.4	55.5	0.42	0.44
Fuzhou	26.08	119.31	11.1	26.7	26.8	77.9	78.7	0.45	0.47
Guangzhou	23.13	113.25	24.6	41.1	41.0	50.0	50.0	0.68	0.69
Guiyang	26.65	106.63	11.8	14.4	14.6	17.8	19.3	0.56	0.53
Hangzhou	30.25	120.17	22.9	35.5	35.3	45.7	44.8	0.56	0.51
Harbin	45.75	126.63	18.9	23.3	23.9	20.1	22.5	0.54	0.55
Hefei	31.85	117.27	24.8	28.7	27.7	15.0	12.9	0.69	0.61
Jinan	36.67	116.98	22.7	35.9	35.0	45.0	43.0	0.64	0.63
Kunming	25.04	102.71	14.3	22.4	22.7	35.8	36.5	0.12	0.10
Nanjing	32.05	118.77	22.6	36.7	35.9	52.4	50.5	0.64	0.54
Nanning	22.82	108.32	17.1	12.4	12.4	-35.9	-34.9	0.32	0.32
Ningbo	29.87	121.54	21.8	30.9	30.8	32.9	32.6	0.60	0.57
Qingdao	36.07	120.38	21.4	20.5	19.4	-6.7	-9.7	0.68	0.70
Shanghai	31.22	121.47	24.7	47.3	47.4	63.0	63.4	0.51	0.46
Shenyang	41.80	123.40	20.2	28.6	29.1	31.7	33.8	0.69	0.72
Shenzhen	22.54	114.06	15.4	49.2	48.7	97.8	97.4	0.40	0.38
Shijiazhuang	38.04	114.51	21.1	45.7	46.2	73.4	75.1	0.51	0.56
Suzhou	31.30	120.60	23.6	31.4	31.1	28.9	28.0	0.47	0.38
Taiyuan	37.87	112.55	21.8	53.0	54.5	75.6	78.1	0.64	0.64
Tangshan	39.63	118.18	27.1	32.3	32.6	19.4	19.8	0.57	0.58
Tianjin	39.13	117.25	25.2	42.4	42.3	47.8	48.0	0.67	0.68
Urumqi	43.83	87.62	32.8	60.9	62.7	55.1	57.3	0.45	0.43
Wenzhou	27.99	120.70	18.9	27.5	27.1	35.7	34.6	0.37	0.35
Wuhan	30.58	114.30	26.9	41.6	41.3	43.9	43.1	0.65	0.58
Wuxi	31.57	120.33	24.0	35.0	34.5	39.7	39.0	0.62	0.54
Xiamen	24.48	118.09	28.3	26.7	26.8	-18.9	-18.5	0.02	-0.01
Xi'an	34.34	108.94	10.8	48.7	50.6	122.3	124.7	0.10	0.11
Xuzhou	34.21	117.28	22.2	35.8	34.3	45.6	42.3	0.64	0.59
Zhengzhou	34.76	113.65	22.9	34.0	34.6	39.8	42.1	0.64	0.66
Zibo	36.78	118.05	22.1	30.7	29.8	28.4	25.9	0.73	0.70

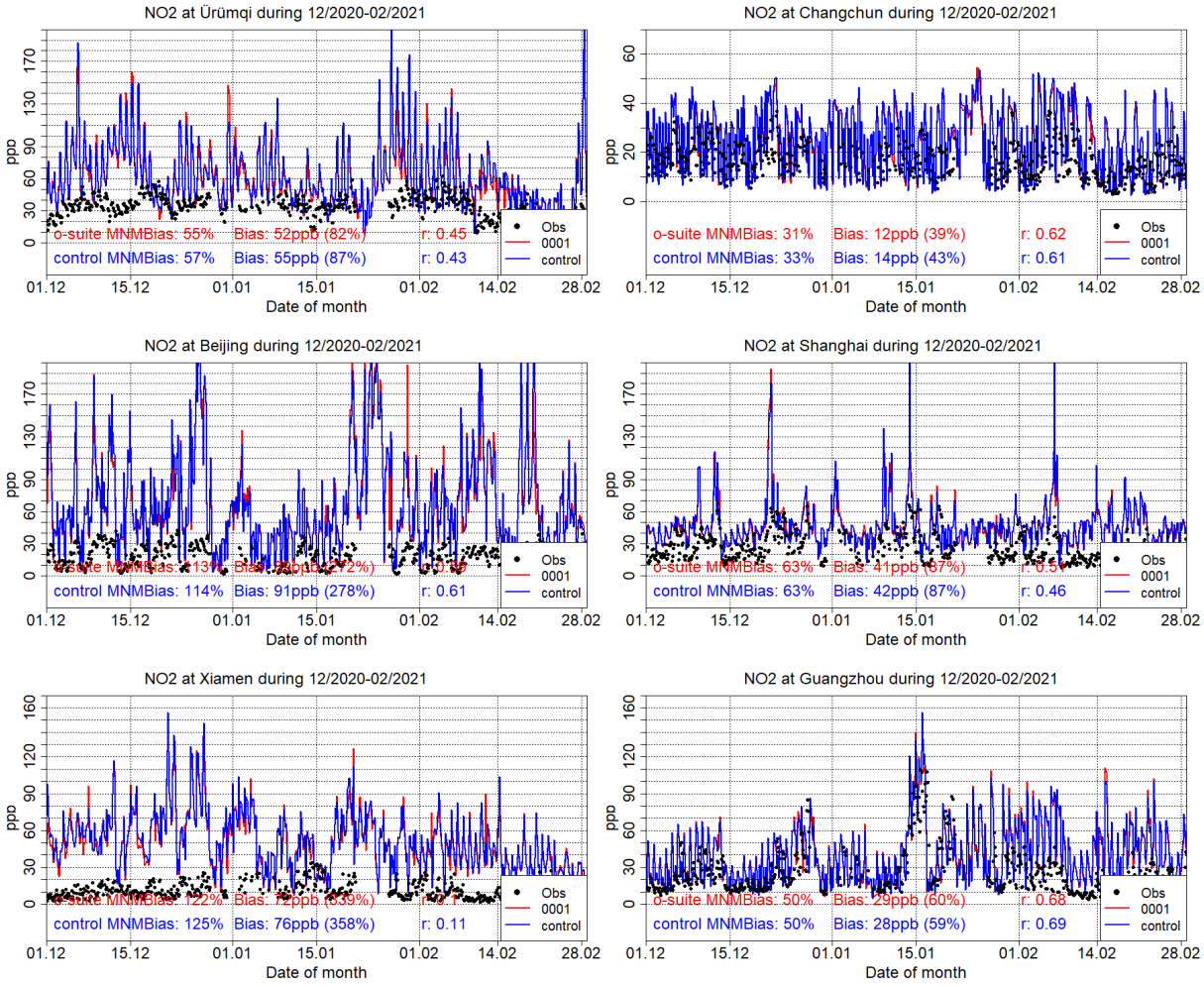


Figure 5.3.2: Surface NO_2 time series for the o-suite (red) compared to MarcoPolo-Panda project observations at Urumqi (43.83°N, 87.62°E, 1st row left), at Changchun (43.89°N, 125.33°E, 1st row right), at Beijing (39.92°N, 116.38°E, 2nd row left), at Shanghai (31.22°N, 121.47°E, 2nd row right), at Xiamen (24.48°N, 118.09°E, 3rd row left), and at Guangzhou (23.13°N, 113.25°E, 3rd row right). Period December 2020 – February 2021.

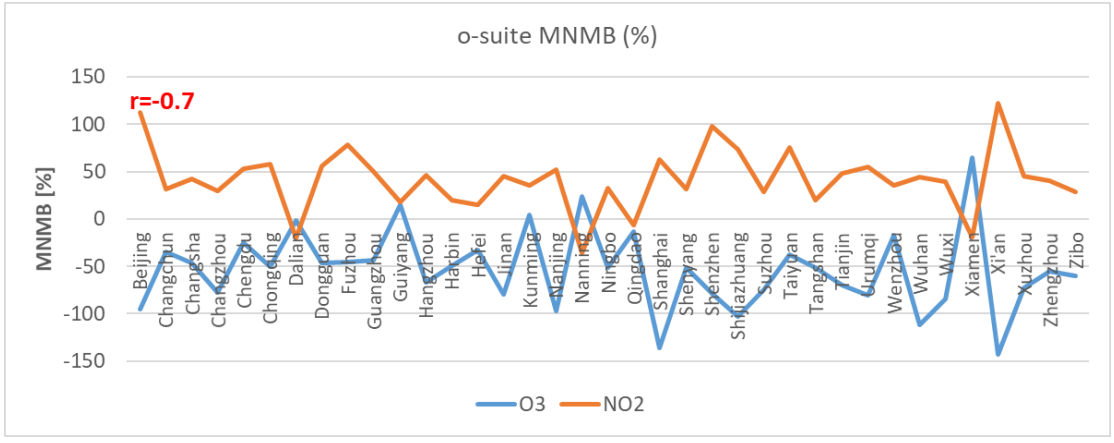


Figure 5.3.3: O-suite ozone and nitrogen dioxide MNMB in % for the 37 under study China Megacities during the period from 1 December 2020 to 28 February 2021.

6. Formaldehyde

6.1 Validation against satellite data

In this section, simulations of tropospheric formaldehyde are compared to TROPOMI/Sentinel-5P data (IUP-UB v1.0) and to GOME-2/MetOp-B (IUP v0.9, preliminary), using the CAMS o-suite as a-priori in the retrievals. The HCHO retrievals are described in Alvarado et al. (2019). The satellite data (tropospheric columns only) are always taken at approximately the same local time, roughly 09:30 LT for the GOME-2 instruments and 13:30 LT for TROPOMI and at clear sky only. The satellite data were gridded to model resolution (currently 0.4° x 0.4° degree). 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 from TROPOMI and GOME-2B are compared to model results for February 2021 as an example for the last winter. The TROPOMI based map comparisons show a very good agreement, while the comparison to GOME-2B shows a positive bias over main emission regions of HCHO and over the ocean at higher southern latitudes. Differences in comparison results between the sensors are in principle due to differences in observation time and the retrieval products.

Time series comparisons between the o-suite and TROPOMI as well as GOME-2B are shown in Figure 6.1.2 for data since Jan 2019. The agreement to the satellite observations is in general good for regions dominated by biogenic emissions (East-Asia and Eastern-US), but the seasonality is underestimated over East-Asia compared to TROPOMI. Both model runs are positively biased over regions dominated by biogenic and biomass burning emissions (North-Africa and Indonesia), with the control showing better results than the o-suite here. The peak over Indonesia in September 3029 is largely overestimated by the o-suite and control. Differences in comparison results are in principle due to differences in observation time or differences in the retrieval products.

More HCHO evaluation plots can be found on the CAMS website, see table 1.2 and

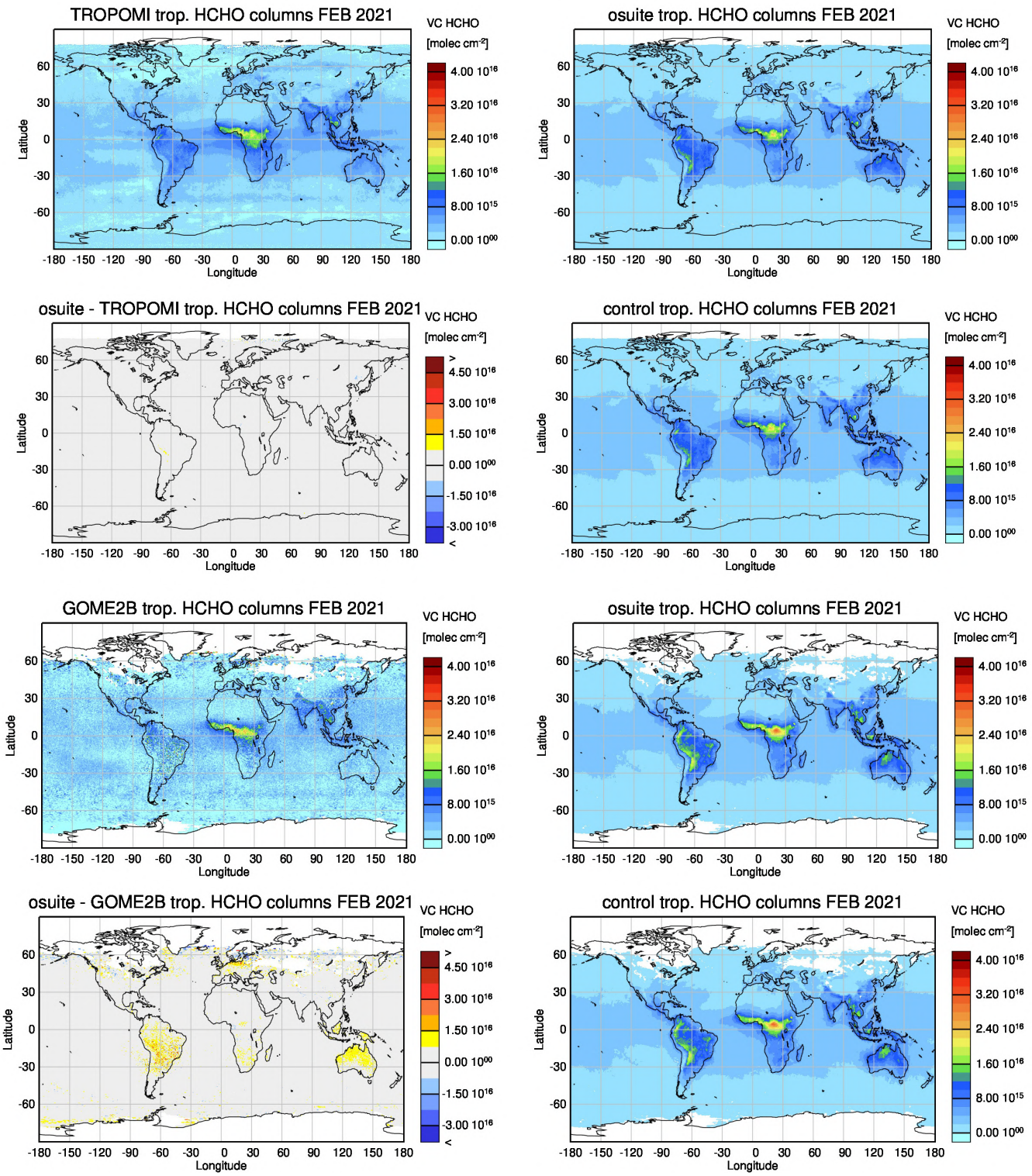


Figure 6.1.1: Global map comparisons of satellite retrieved, and model simulated tropospheric HCHO columns [molec cm^{-2}] for February 2021 based on (top two rows) TROPOMI and (bottom rows) GOME-2B. The columns show (first row, left to right) satellite observations, o-suite, (second row, left to right) the difference between o-suite and satellite observations, the control run. The satellite data were gridded to model resolution (0.4° x 0.4° degree) and the CAMS o-suite was used as a-priori in the retrievals.

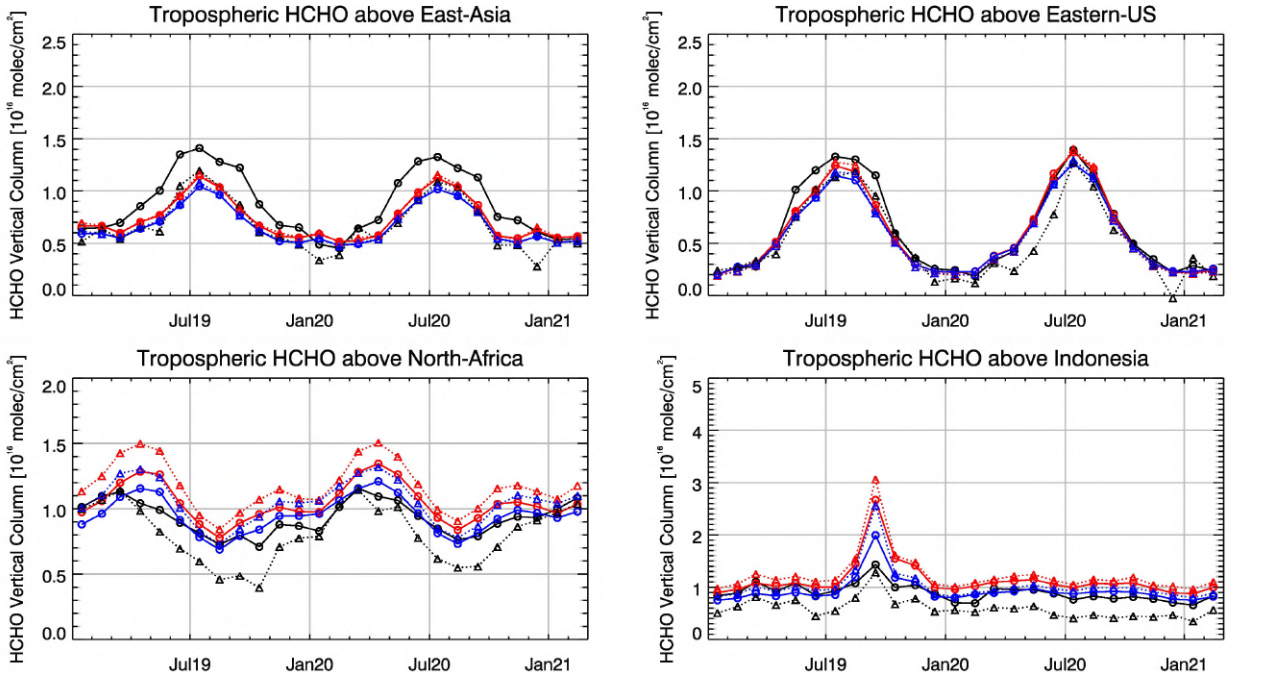


Figure 6.1.2: Time series of average tropospheric HCHO columns [10^{16} molec cm^{-2}] from (black and grey) satellite retrievals, (blue) control and (red) o-suite model results since Jan 2019. The solid lines with circles show comparisons based on TROPOMI, the dotted lines with triangles show comparisons for GOME-2B. 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).

6.2 Evaluation against ground-based DOAS observations

In this section, we compare the HCHO columns of the CAMS products 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)). The validation methodology 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 CAMS 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.

Figure 6.2.1 shows the absolute biases Dec – Feb 2021 at the different sites and indicates strongly reduced biases for the different sites. At all three sites high pollution events are not captured by the CAMS runs and leads to a higher overall underestimation (Fig 6.2.2). 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 products, the high emission events are not. A longer time series is required to analyse a seasonal dependence in the bias, at Cabauw and De Bilt, however a seasonal dependence may be observed with an underestimation during summer and overestimation during winter months.

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

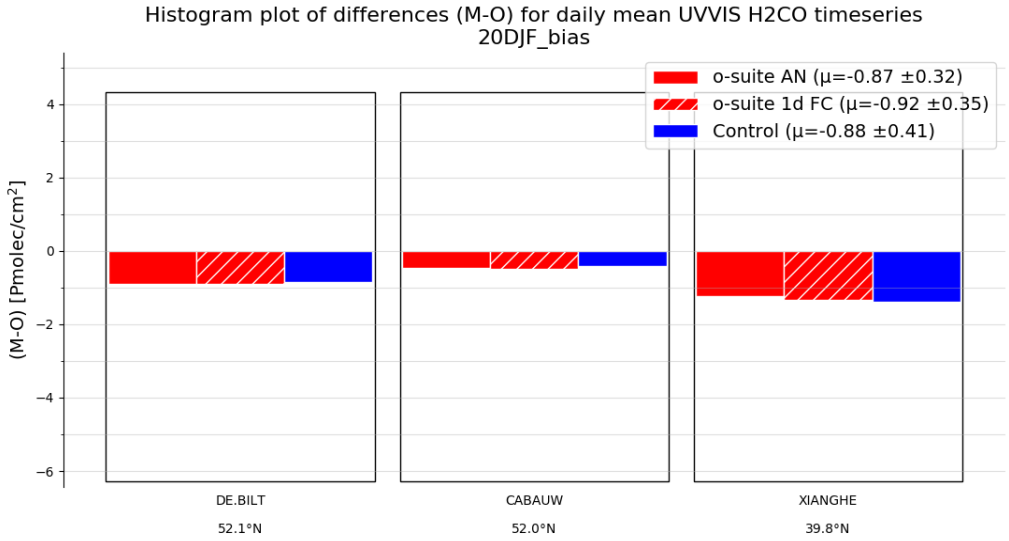


Figure 6.2.1: Table diagram showing the seasonal absolute bias in DJF 2021 for three stations, sorted by latitude

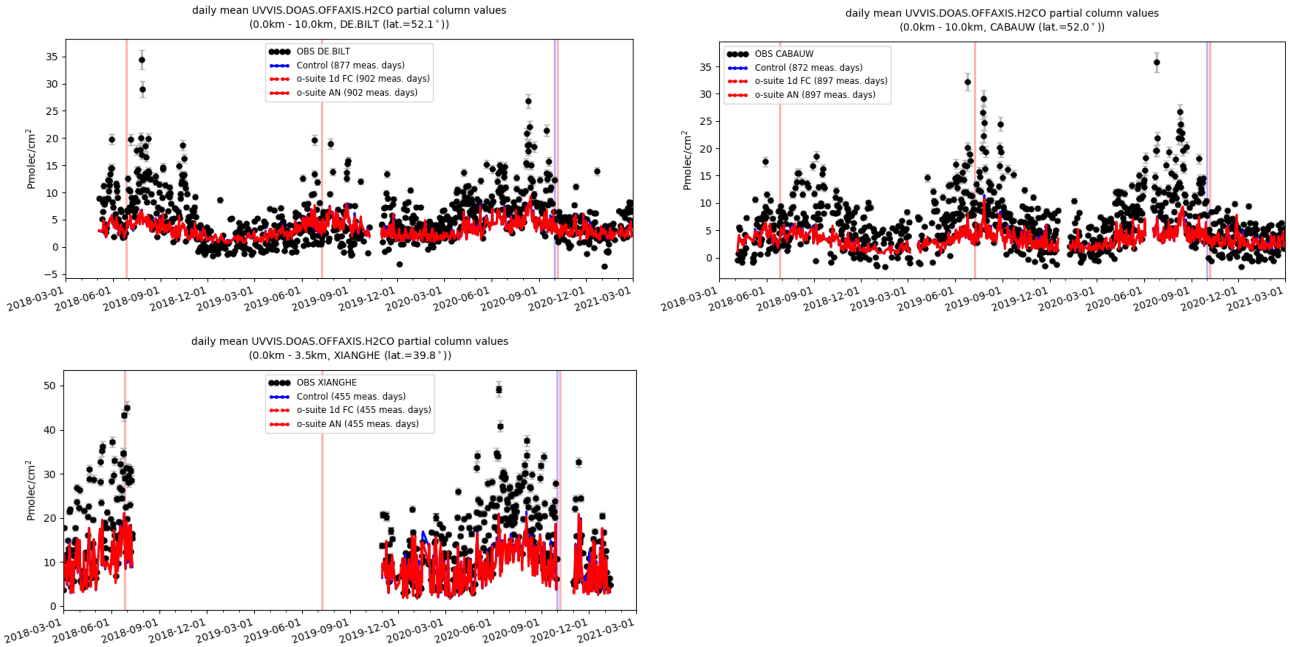


Figure 6.2.2: Time series of HCHO partial columns at the three different sites. All CAMS products underestimate the high peak HCHO concentrations. At Cabauw and De Bilt a seasonal dependence in the bias can be seen. Period March 2018 – February 2021.

7. Water vapour

Like for ozone and CO, water vapour has been sampled mostly sampled by the two Lufthansa planes during DJF 2021 but also by Hawaiian Airlines. The availability of water vapour data at Frankfurt is similar to that of CO and ozone (see IAGOS ozone and CO sections).

The results from the o-suite and control run are mostly similar (Fig.7.1, Fig. 7.2) and the variability of water vapour during DJF 2021 is well represented by the models in all layers. The two runs agree well with the observations in the lowest layers with small positive bias as shown on Fig. 7.2. On average over the full period the MNMB is smaller than 10% in absolute value (Fig. 7.3) and the correlation coefficient is greater than 90% (Fig. 7.3).

The agreement is worse in the upper layers, with larger biases and smaller correlation (Fig. 7.2 and Fig. 7.3). In the free troposphere, both large underestimations and overestimations are found (Fig. 7.2), and the correlation remains higher than 60% up to 8000 m (Fig. 7.3). In the UTLS, the bias of the models is mostly negative with on average an MNMB absolute value of more than 40% (Fig. 7.3).

Several examples of individual profiles at Frankfurt are shown on Fig. 7.4., which illustrate the results. Most of these profiles have complex shapes with extrema of water vapour in the mid-troposphere which are not always well reproduced by the CAMS configurations. Individual profiles from other regions of the world are also presented in Fig. 7.5-8 for respectively: North America, South Africa, India, Eastern and South-eastern Asia, Central America and the Caribbean. For these other regions the models show overall results similar to those observed at Frankfurt.

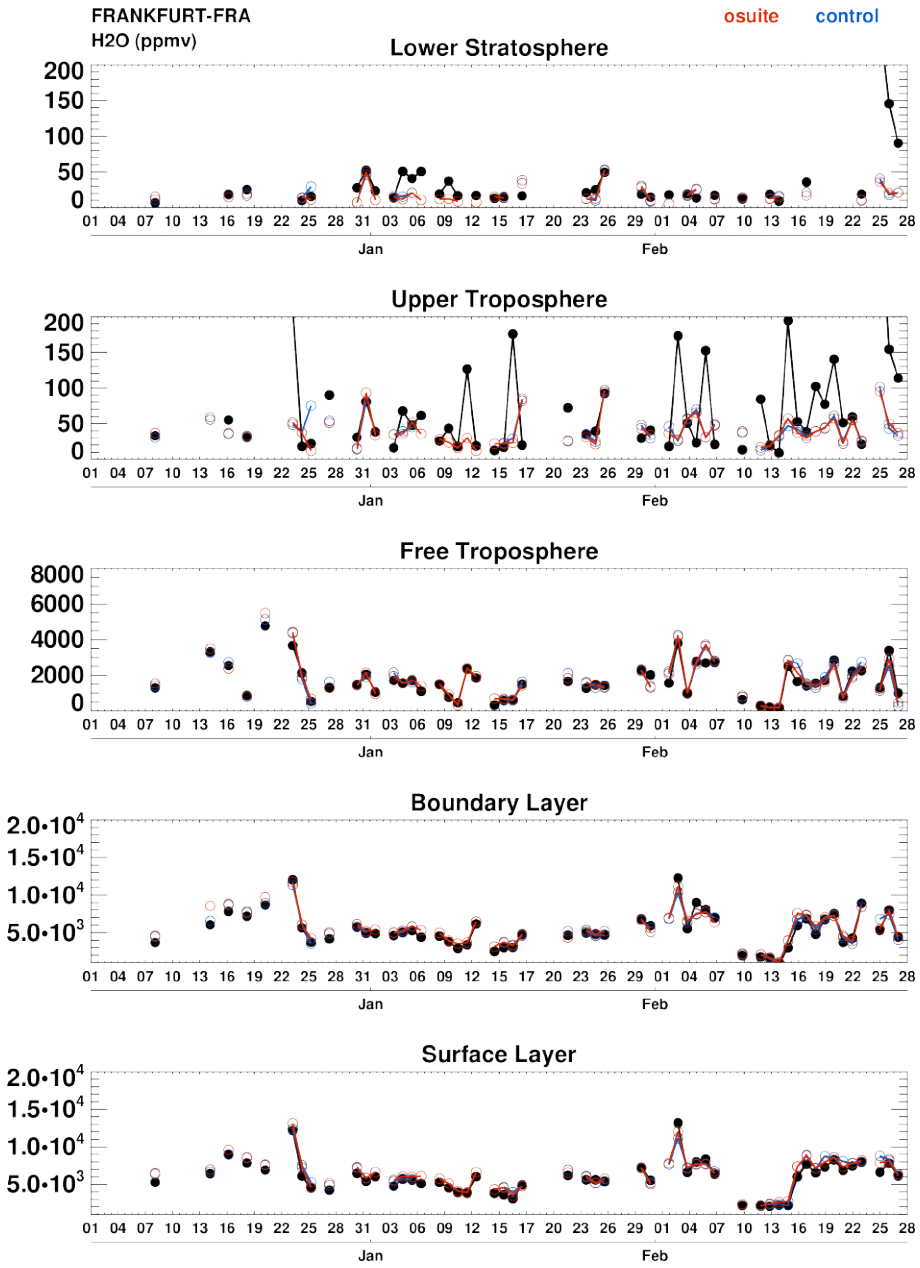


Figure 7.1: Time series of daily mean water vapour over Frankfurt during DJF 2021 for 5 layers: Surface Layer, Boundary Layer, Free Troposphere, Upper Troposphere and Lower Stratosphere. IAGOS is shown in black, the o-suite in red and associated control run in blue. Units: ppmv.

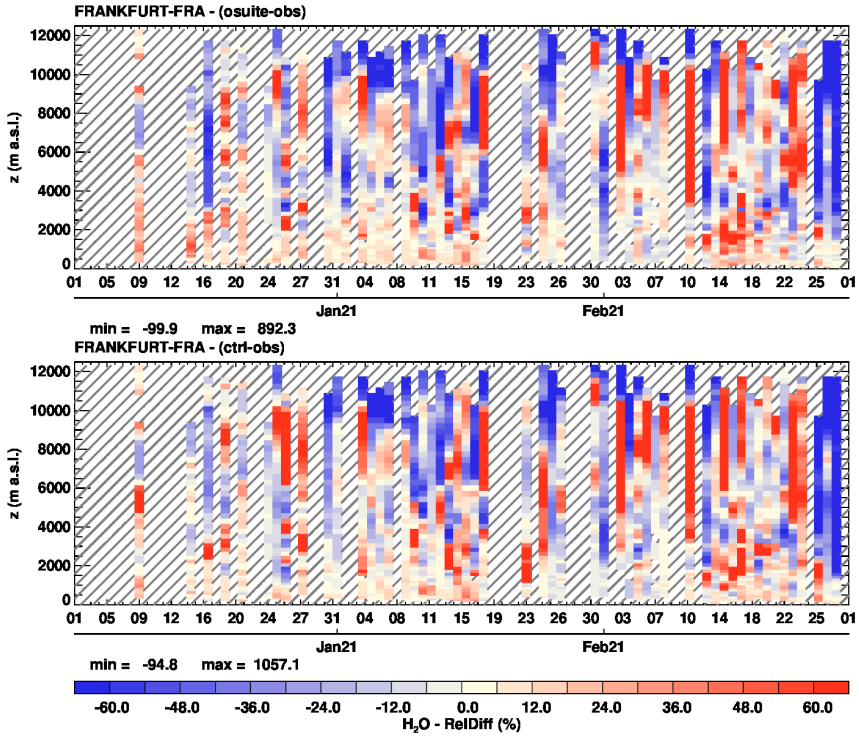


Figure 7.2: Time series of the relative differences ([model – observations]/observations) in daily profiles for water vapour over Frankfurt during DJF 2021. The top panel corresponds to o-suite the bottom panel to control run. Units: %.

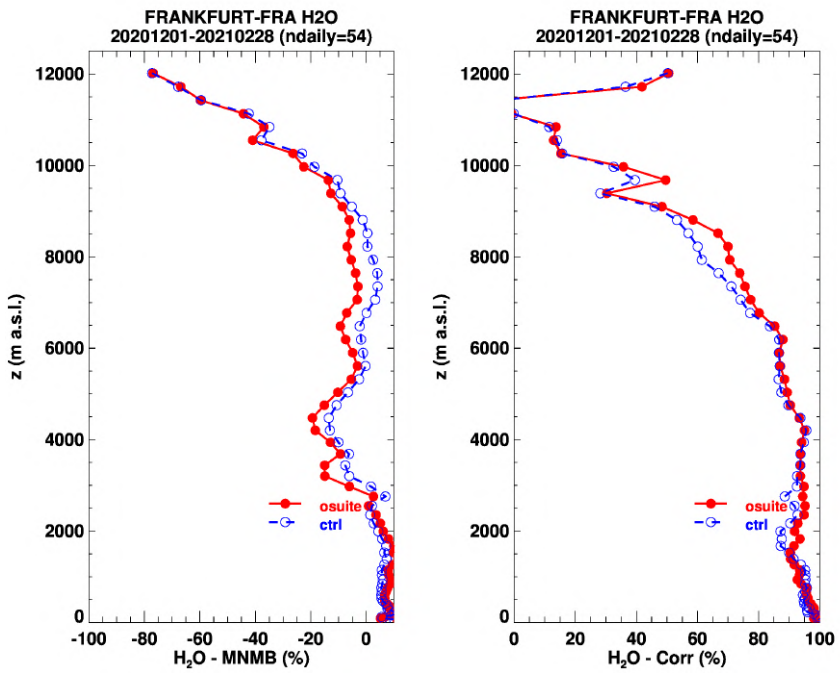


Figure 7.3: Model scores (MNMB and Correlation coefficient) for water vapour at Frankfurt calculated over the period DJF 2021. The left panel corresponds to MNMB and the right,panel to Correlation coefficient. The o-suite is shown in red and associated control run in blue. Units: %.

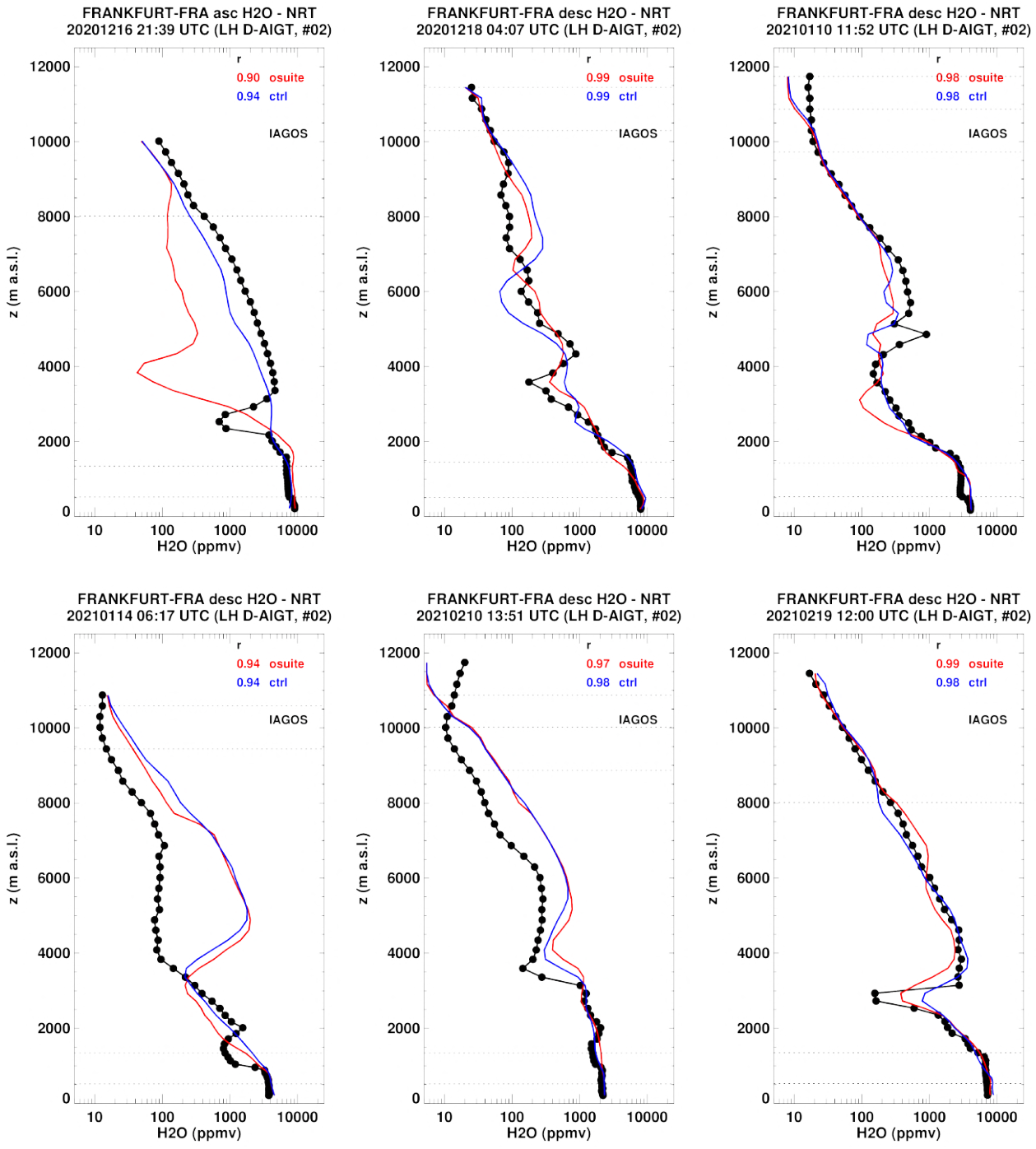


Figure 7.4: Selection of individual profiles for water vapour from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over Frankfurt during DJF 2021. Units: ppmv.

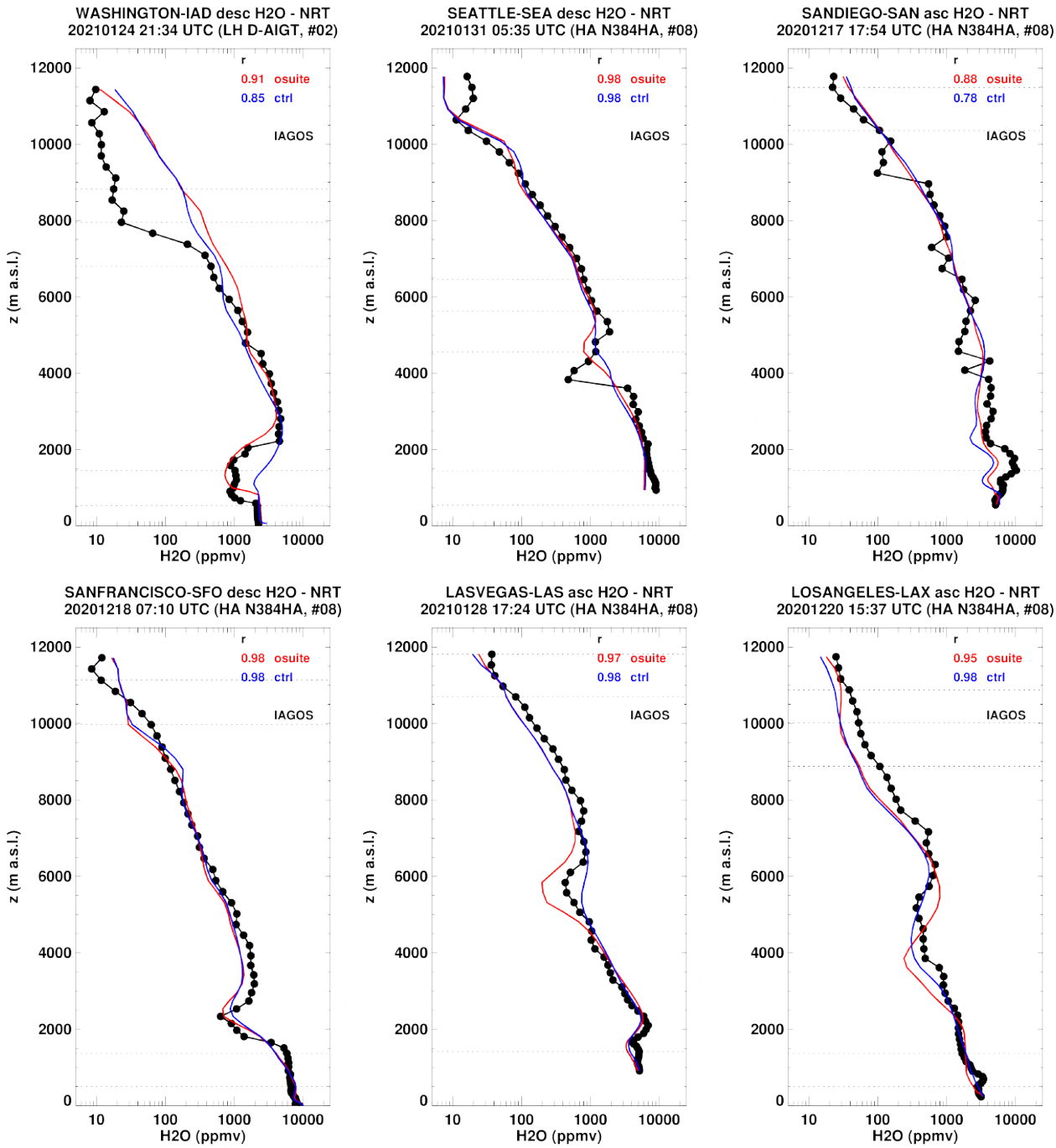


Figure 7.5: Selection of individual profiles for water vapour from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over the North America during DJF 2021. Units: ppmv.

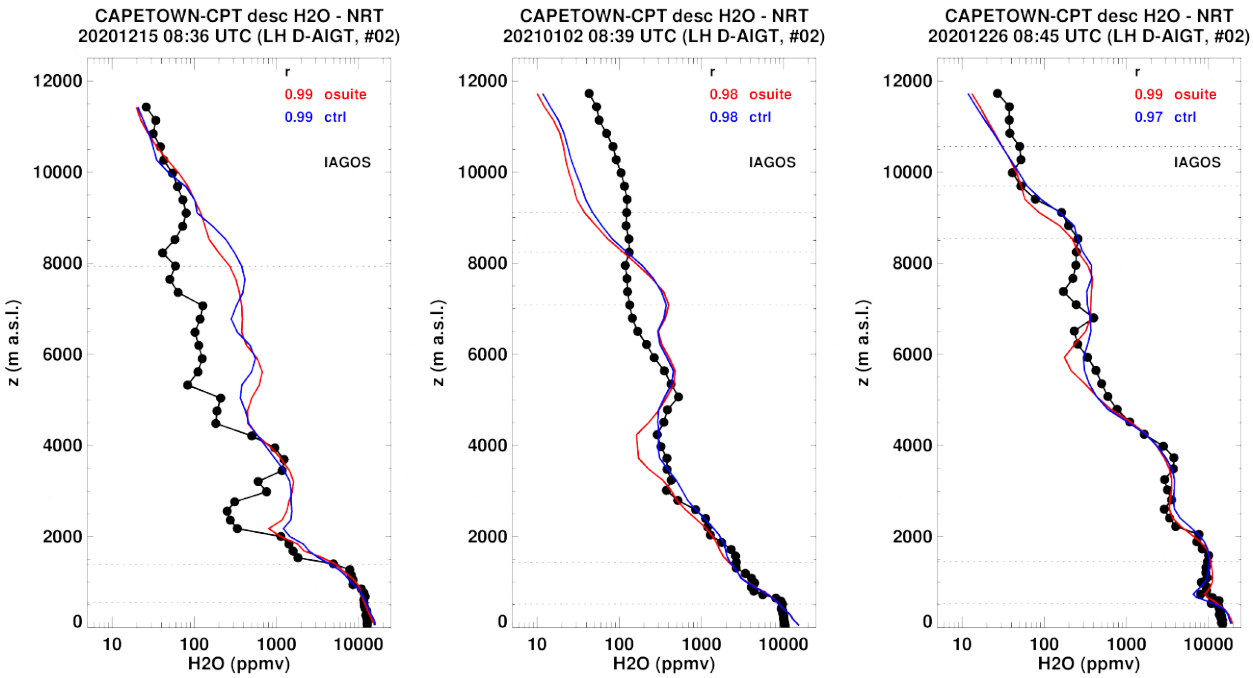


Figure 7.6: Selection of individual profiles for water vapour from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over the South Africa during DJF 2021. Units: ppmv.

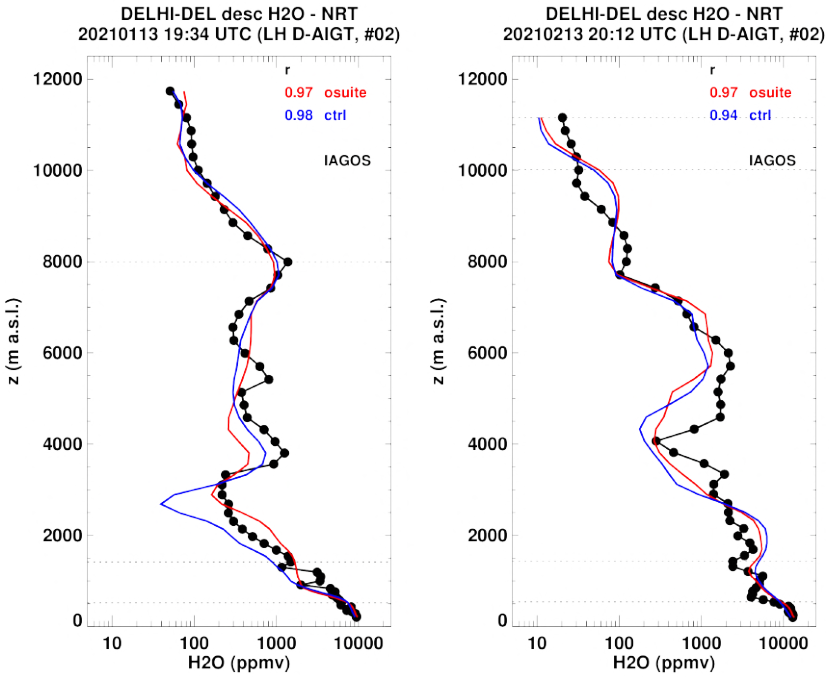


Figure 7.7: Selection of individual profiles for water vapour from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over the India during DJF 2021. Units: ppmv.

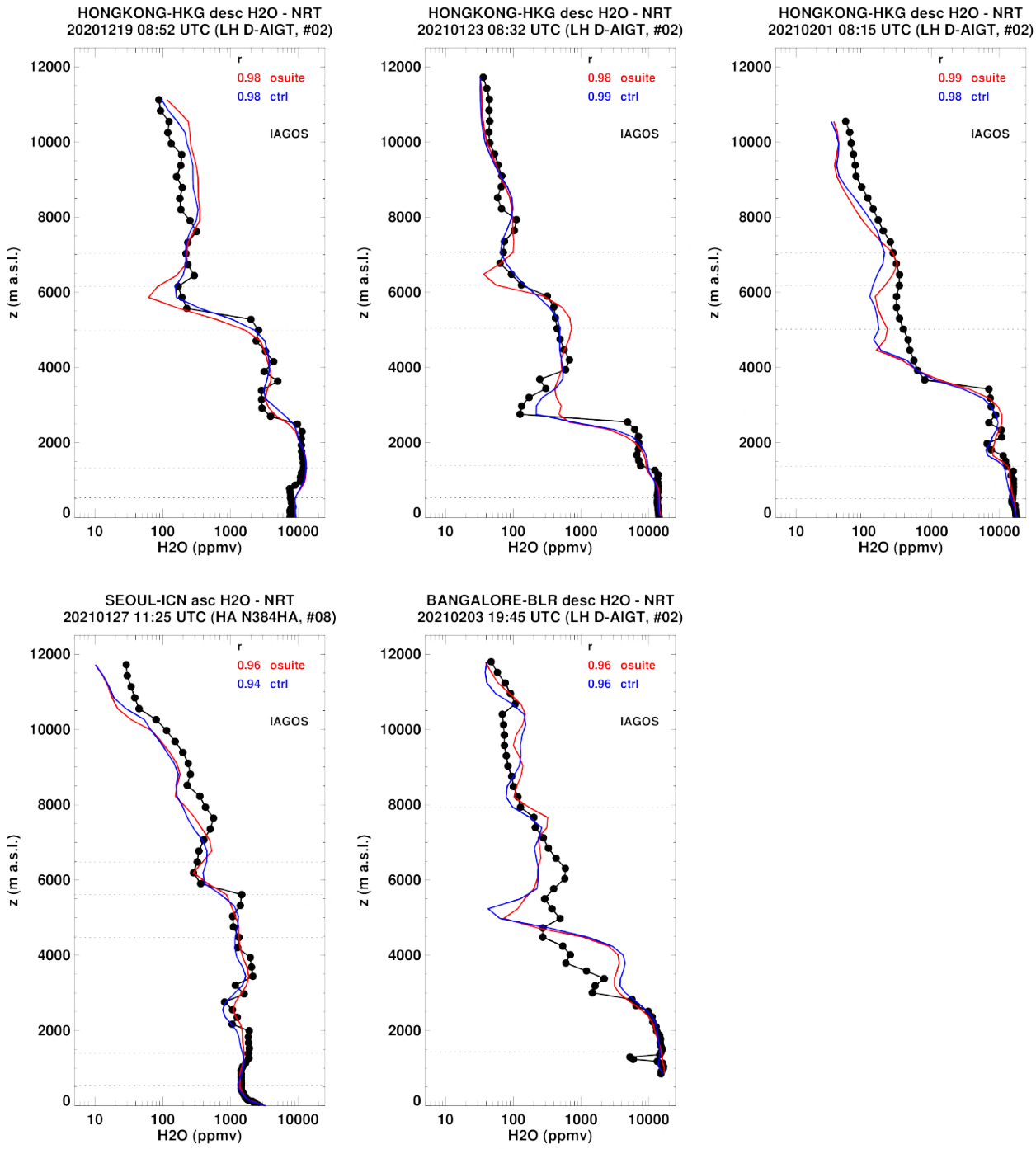


Figure 7.8.a: Selection of individual profiles for water vapour from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over the Eastern and Southeastern Asia during DJF 2021. Units: ppmv.

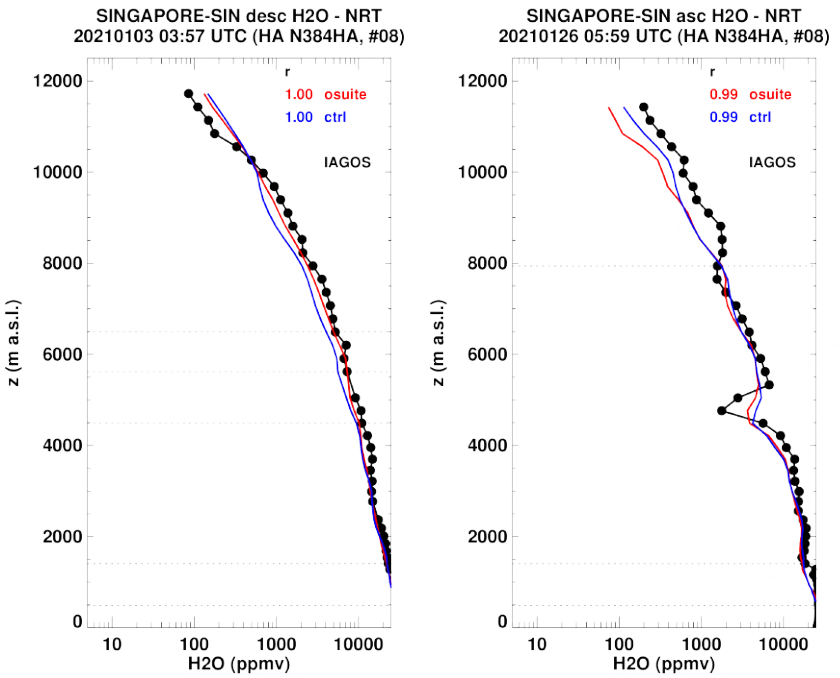


Figure 7.8.b: Selection of individual profiles for water vapour from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over the Eastern and Southeastern Asia during DJF 2021. Units: ppmv.

8. Aerosol

8.1 Global comparisons with Aeronet and EMEP

A 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 (https://aerocom-classic.met.no/cgi-bin/aerocom/surfobs_annualrs.pl?PROJECT=CAMS).

A second web interface (still under development) integrates NRT global surface observations of PM10, PM25, NO2, ozone and Aeronet AOD and Ångström Exponent (AE) covering comparisons since year 2019 and for the 2021/2020 DFJ period: https://aerocom-evaluation.met.no/main.php?project=cams84&exp=eval-2021-DJF.

Correlation, based on daily aerosol optical depth from NRT Aeronet observations, has been rather stable recently, possibly increasing in the DJF period. The o-suite forecast at +3 days shows only slightly lower correlation. See figure S.6. 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 comparisons to Aeronet.

The performance of the o-suite model exhibits some seasonal variation in AOD depending on region (Fig. 8.1.1). Noteworthy is the persistent AOD overestimation over North America (Fig. 8.1.1-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. 8.1.2) shows a positive bias against Aeronet in the Southern Hemisphere, but also at Northern mid-latitudes.

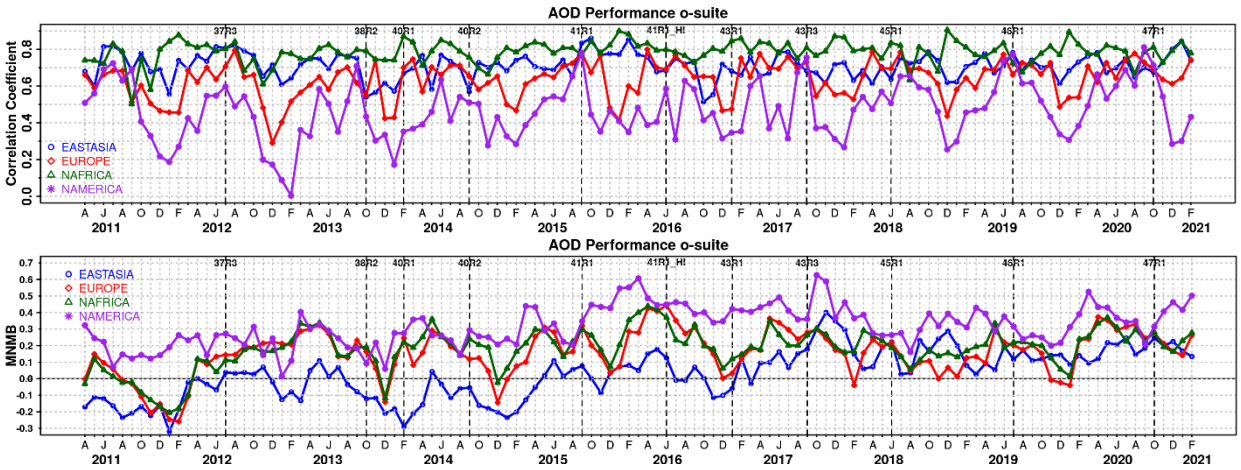


Figure 8.1.1. Correlation coefficient (top) and modified normalized mean bias (MNMB) (bottom) in AOD, since 2011, based on daily AOD comparison (Aeronet V3 level 1.5 data) in four world regions [East-Asia (blue); Europe (red); North Africa (green); North America (purple)] for the o-suite.

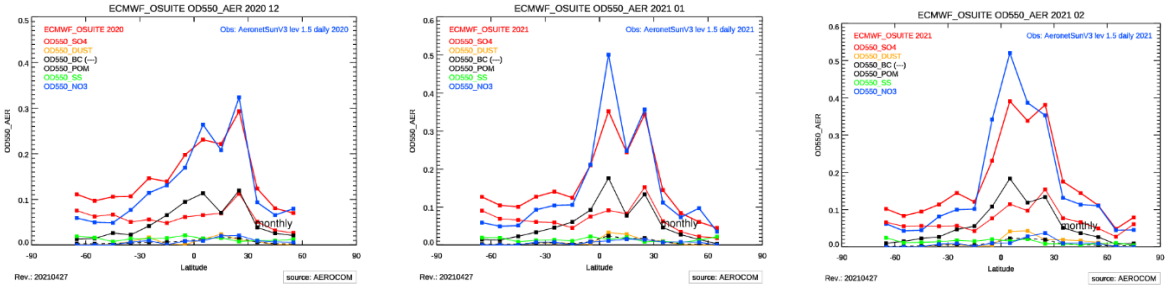


Figure 8.1.2. Aerosol optical depth of o-suite (red) compared to latitudinally aggregated Aeronet V3 level 1.5 data (blue) for the three months covered by this report.

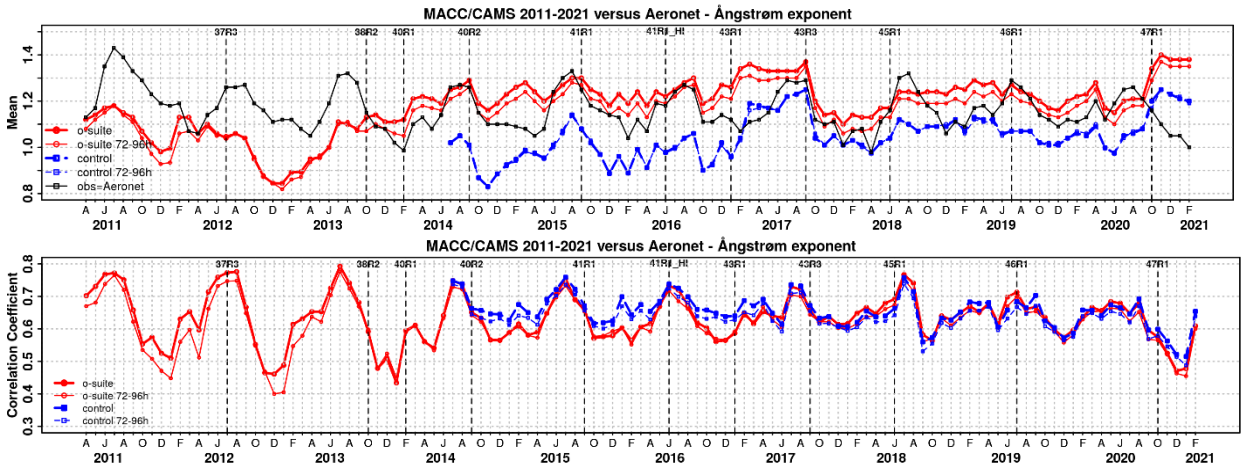


Figure 8.1.3. a) (top) Evolution of mean Ångström exponent in o-suite and control at Aeronet sites (Aeronet V3 level 1.5 data), based on matching monthly mean values. O-suite (thick red curve); o-suite at last forecast day (light red curve); control (blue dashed curve); control at last forecast day (light blue dashed curve). B) (bottom) Correlation using daily matching Ångström exponent.

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 DJF 2021 a larger positive bias (Figure 8.1.3-a) and lower correlation (Figure 8.1.3-b) along with the model update in October 2020. Correlation from all AE data is lower than for AOD.

Figure 8.1.3 together with 8.1.4 shows that model version changes are responsible for a shift in Ångström exponent. More sulphate and organic matter shift the size distribution to smaller sizes. Figure also shows considerable differences in organic and sulphate AOD in the assimilated IFS experiment and in the control simulation. The o-suite uses data assimilation to obtain an analysis of the 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. Table 8.1.1 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 less AOD (-45% compared to o-suite). 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.

Table 8.1.1. Mean global total and speciated AOD in the o-suite for the last two periods covered by the report and their change after 3 forecast days.

o-suite	Mean SON 2020 0-24h	Change wrt to first day on day 4	Mean DJF2020/21 0-24h	Change wrt to first day on day 4
AOD@550	0.164	-15%	0.148	-18%
BC-OD@550	0.0072	-20%	0.0044	-20%
Dust-OD@550	0.014	13%	0.0075	8%
OA-OD@550	0.045	-24%	0.034	-24%
SO4-OD@550	0.049	-25%	0.060	-26%
SS-OD@550	0.043	-6%	0.040	-7%
NO3-OD@550	0.0038	7%	0.0020	2%

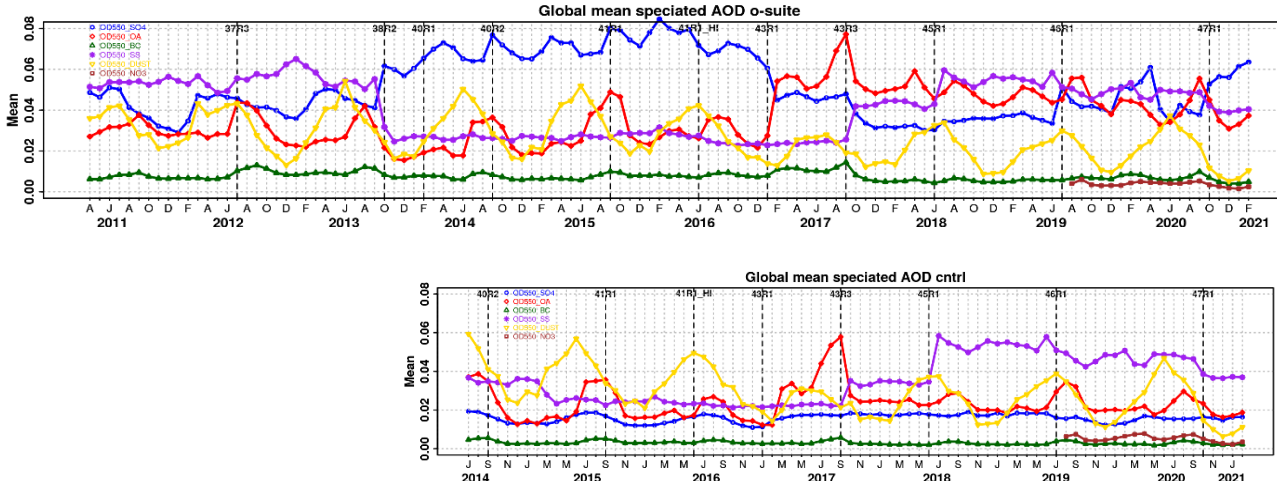


Figure 8.1.4. Evolution of the aerosol components of total AOD@550nm [OD550_SO4 = sulphate(blue); OD550_OA = organics(red); OD550_BC = black carbon(green); OD550_SS = sea salt(purple); OD550_DUST = dust(yellow); OD550_NO3 = nitrate(brown)] in o-suite and control simulation.

Global PM10 and PM25 daily NRT data from surface observations (Airnow, EEA and Marco Polo) can be used to evaluate the surface PM concentrations in the IFS aerosol model. NRT data for the DFJ 2020/2021 period suggest on average for North America, Europe, and China a PM10 MNMB bias of -31%, +2%, -4% respectively and for PM2.5 a MNMB bias of -4%, -10% and +5% respectively. Inter-regional variations of the bias are large as can be seen in figure 8.1.5. A detailed web interface integrates NRT global surface observations of PM10, PM25, NO2, ozone and Aeronet AOD and AE, available since January 2020 (https://policy.atmosphere.copernicus.eu/aeroval.php#).

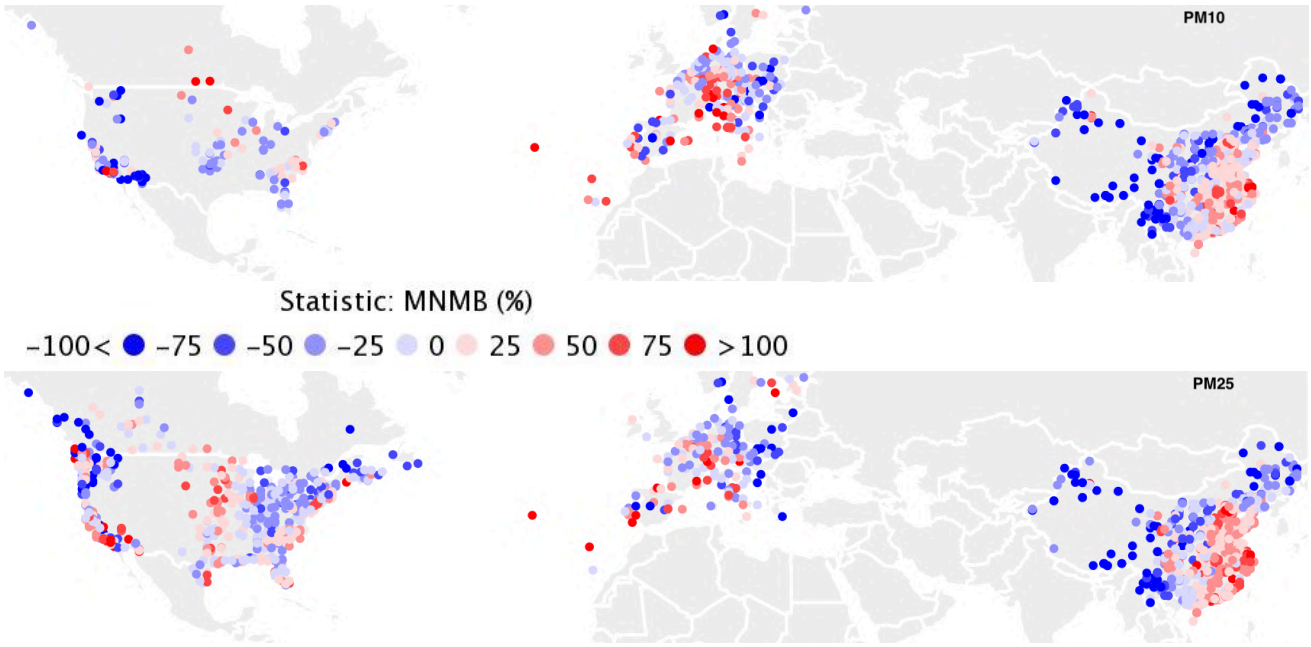


Figure 8.1.5. MNMB Bias [%] based on daily DJF 2020/2021 values of PM10 and PM2.5 for the IFS o-suite against the combination of AirNow, EEA and Chinese monitoring data obtained through the EU Marco Polo project.

For a longer inspection of the PM performance of the IFS model we utilize a climatological average constructed from observational data in the period 2000-2009 as available in the EBAS database hold at NILU. Climatological, monthly surface concentration of particulate matter below 10 µm (PM10) and below 2.5 µm (PM25) stem from 160 background IMPROVE and EMEP stations, thus representing North America and Europe. Figure 8.1.6 shows the evolution of mean observed and simulated PM10 and PM2.5. It seems that also against this PM climatology there is a slightly negative bias in the latest period. Shown is also the statistics of model data within factor 2 of observed data, 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. The o-suite is also better than the control simulation most of the times, at least for PM10. With the July 2019 model upgrade the PM2.5 performance of the o-suite is very similar to the control.

A similar method is used to evaluate surface concentrations. The MNMB bias of the o-suite against a climatology is shown in Fig. 8.1.7. Organic aerosols seem to be overestimated, while ammonium, sulphate, black carbon, and sea salt seem to be underestimated.

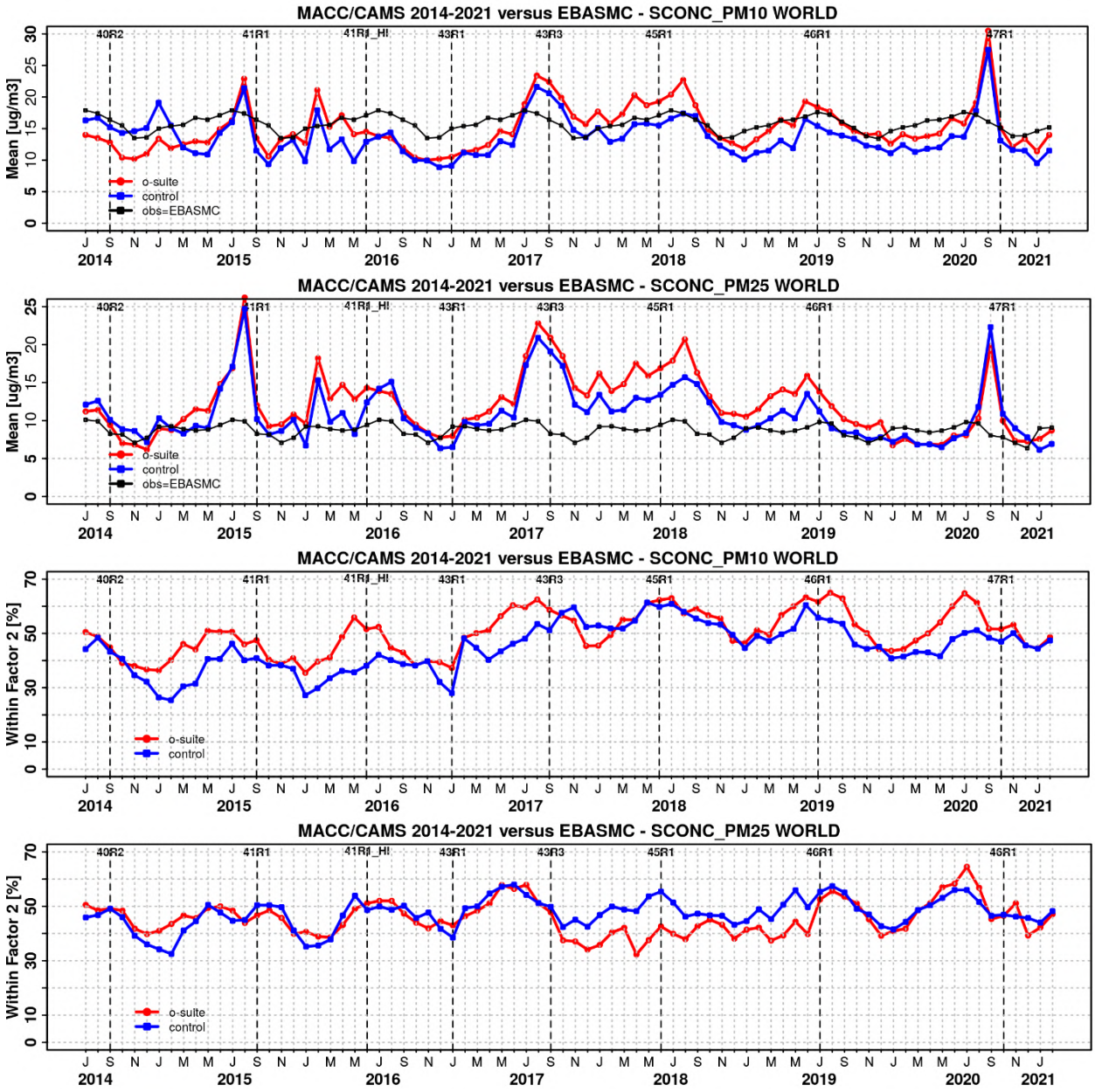


Figure 8.1.6. Temporal evolution of monthly mean average PM10 and PM2.5 concentrations at EMEP (Europe) and IMPROVE sites (North America) and model data fraction within a factor 2 of observed; ca 160 sites, observed data are averaged from data available in EBAS from 2000-2009.

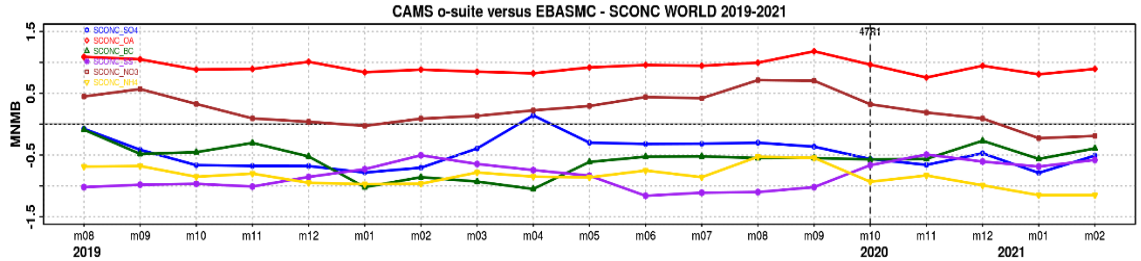


Figure 8.1.7. Evolution of MNMB Bias of simulated surface concentrations in o-suite against a climatology of speciated aerosol concentrations, constructed as for PM, mainly over Europe and North America.

8.2 Validation of dust optical depth against AERONET, and comparisons with the Multi-model Median from SDS-WAS

The 72-hour forecasts (on a 3-hourly basis) of dust aerosol optical depth (DOD) from CAMS o-suite and control have been validated for the period 1 December 2020 – 28 February 2021 against the AERONET Spectral Deconvolution Algorithm (SDA) 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/). At those sites where the SDA products are available, the dust AOD evaluation will be complemented with AOD-coarse, which is fundamentally associated with maritime/oceanic aerosols and desert dust. Since sea-salt is related to low AOD (< 0.03; Dubovik et al., 2002) and mainly affects coastal stations, high AOD-coarse values are mostly related to mineral dust.

During this season, satellites (see MODIS in Figure 8.2.1) show lower dust activity (AOD < 0.3). The highest AOD values are observed in the Bodélé. Higher AOD values (> 0.3) in the Gulf of Guinea are associated with biomass burning from the Savannah fires. In the Middle East, the AOD seasonal average is under < 0.3 with maximum AOD in the south-eastern Arabia Peninsula. Both CAMS runs show maximum DOD values in the Bodélé and surrounding regions in Chad and Mali, as well as Mauritania, Dakar, and Algeria. They appear overestimated compared to the SDS-WAS multi-model ensemble (see Figure 8.2.1) and the MODIS observations. The Atlantic transport and Tunisia, Lybia and Egypt seem to be underestimated. In the Middle East, both CAMS runs show a maximum in the central north of Saudi Arabia and Iraq; the other regions have DOD < 0.06.

From December to February, s o-suite (control) reproduce the daily variability of AERONET dust-filtered observations (see Figure 8.2.2), with a correlation coefficient of 0.61 (0.56) averaged over all AERONET sites, which is lower than the SDS-WAS multi-model product which has a correlation coefficient of 0.84. Regarding the mean bias (MB), the o-suite tends to underestimate the AERONET observations with an MB of -0.06 for o-suite and -0.05 for control in comparison with the SDS-WAS multi-model that shows slightly lower underestimations (MB of -0.03).

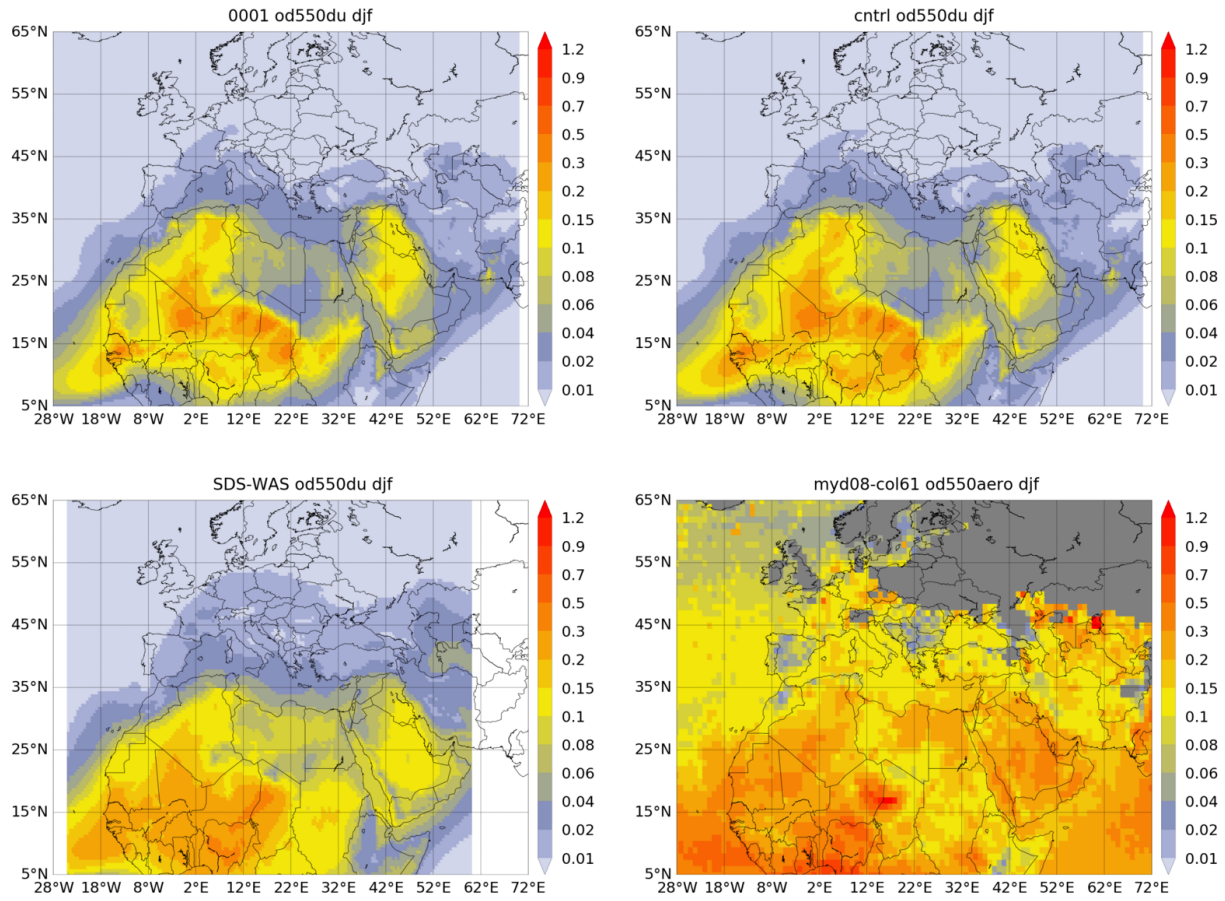


Figure 8.2.1: Averaged DOD 24h forecast from o-suite (top left) and control (top right), DOD of the multi-model SDS-WAS Median product (bottom left) as well as AOD from MODIS/Aqua Collection 6.1 Level 3 combined Dark Target and Deep Blue product (bottom right) for the study period.

Over desert dust sources in the Sahara (see Table 8.2.1 as well as Tamanrasset INM AERONET site in Figure 8.2.3a), the CAMS runs do reproduce the daily variability with a correlation coefficient 0.86 for o-suite and 0.84 for control. However, DOD is slightly overestimated (MB of 0.01 for o-suite and 0.03 for control). As shown in Tamanrasset INM (Figure 8.2.3a), the overestimations observed in the control run are reduced in the o-suite. The SDS-WAS Multi-model result for the Sahara shows better skills for this season (with a seasonal correlation of 0.94 and MB of 0.01). In the Middle East (see Kuwait University in Figure 8.2.3a), the comparison with AERONET observations shows similar correlations coefficients (0.49 for o-suite and 0.45 for control) and DOD is underestimated with MB (MB of -0.05 for o-suite and for control, see Table 8.2.1). Also in the Middle East, the SDS-WAS Multi-model presents better results with a seasonal correlation of 0.84 and MB of -0.03.

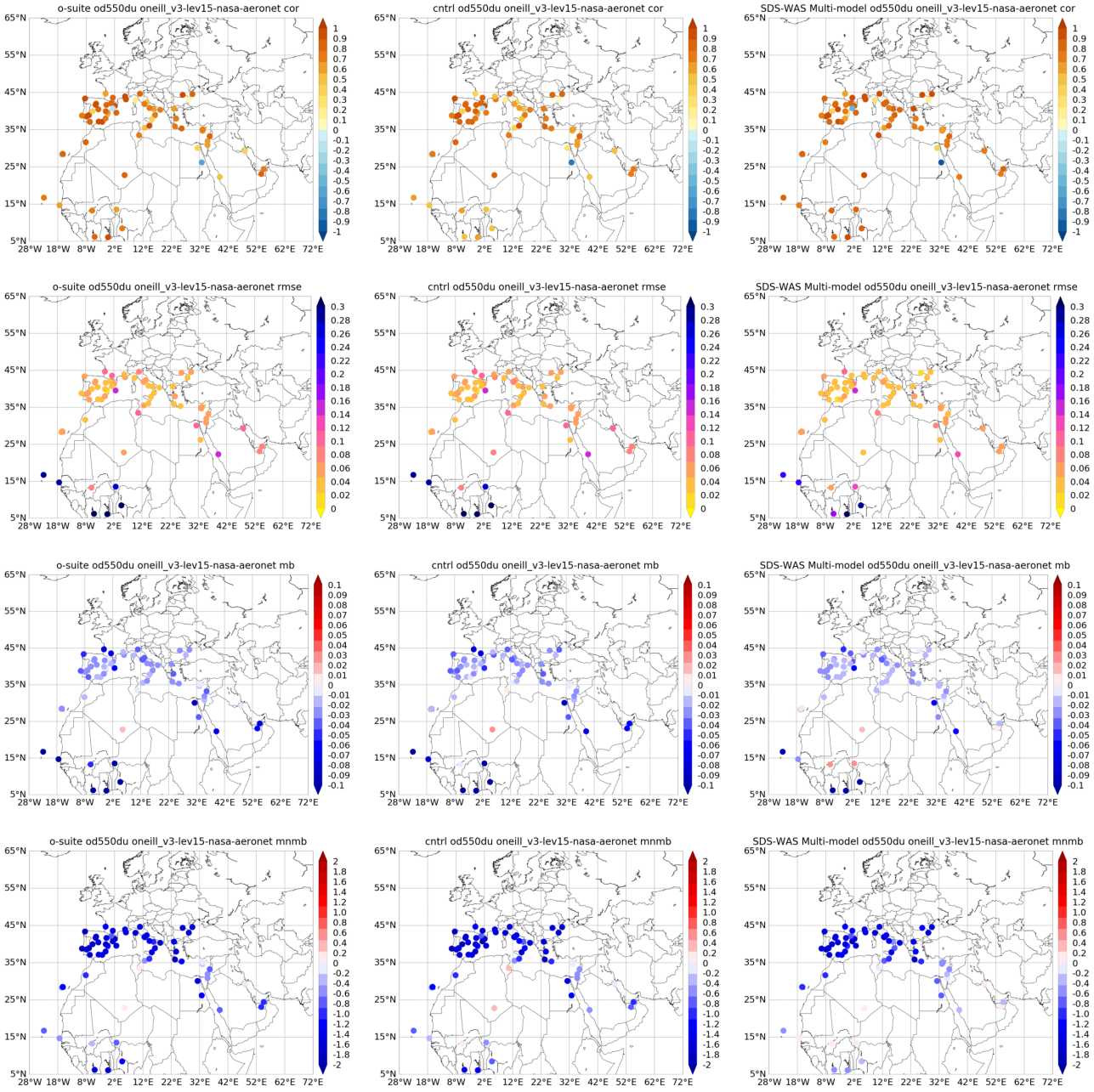


Figure 8.2.2: Skill scores (correlation coefficient, RMSE, MB and MNMB) for 24-hour forecasts of CAMS o-suite (left column), control (central column) and DOD Multi-model SDS-WAS Median (right column) for the study period. AOD-coarse from AERONET SDA is the reference.

Table 8.2.1: Skill scores (MB, MNMB, RMSE and r) of 24h forecasts (on 3hourly basis) for CAMS o-suite, CAMS control and SDS-WAS Multi-model Median for the study period, and the number of data (NDATA) used. DOD (SDA AOD coarse product) from AERONET is the reference.

		Control				o-suite DOD				SDS-WAS Median DOD			
	NDATA	MB	MNMB	RMSE	r	MB	MNMB	RMSE	r	MB	MNMB	RMSE	r
Sahara	228	0.03	0.26	0.07	0.84	0.01	0.10	0.05	0.86	0.01	0.15	0.03	0.94
Sahel	897	-0.21	-0.68	0.35	0.34	-0.24	-0.96	0.35	0.53	-0.07	-0.14	0.19	0.82
Tropical North Atlantic	126	-0.13	-0.73	0.29	0.57	-0.13	-0.78	0.28	0.74	-0.09	-0.49	0.23	0.80
Subtropical North Atlantic	437	-0.01	-0.70	0.07	0.50	-0.01	-0.71	0.07	0.53	0.00	-0.46	0.08	0.55
North Western Maghreb	172	-0.02	-1.01	0.03	0.82	-0.02	-1.06	0.03	0.83	-0.02	-1.04	0.03	0.77
Western Iberian Peninsula	611	-0.03	-1.64	0.05	0.78	-0.03	-1.66	0.05	0.83	-0.03	-1.53	0.05	0.86
Iberian Peninsula	731	-0.02	-1.68	0.04	0.77	-0.03	-1.70	0.04	0.81	-0.02	-1.60	0.04	0.84
Western Mediterranean	973	-0.03	-1.46	0.07	0.75	-0.02	-1.45	0.06	0.85	-0.02	-1.36	0.06	0.79
Central Mediterranean	1062	-0.02	-1.01	0.06	0.69	-0.02	-1.10	0.06	0.69	-0.02	-1.04	0.04	0.85
Eastern Mediterranean	1543	-0.03	-0.80	0.06	0.60	-0.03	-0.84	0.06	0.63	-0.02	-0.71	0.05	0.73
Eastern Sahara	-	-	-	-	-	-	-	-	-	-	-	-	-
Middle East	1181	-0.05	-0.63	0.11	0.45	-0.05	-0.74	0.11	0.49	-0.02	-0.11	0.08	0.70
All sites	9024	-0.05	-1.00	0.15	0.56	-0.06	-1.08	0.15	0.61	-0.03	-0.80	0.10	0.84

In the Sahel (Table 8.2.1 and Banizoumbou in Figure 8.2.3a), the o-suite enhances the underestimation observed in the control run (MB of -0.21 for control and -0.24 for o-suite) even though the o-suite better reproduces the observed daily variability (with a correlation of 0.53 for o-suite in comparison to 0.34 for control). In the Sahel a larger discrepancy with the SDS-WAS Multi-model result is observed. The underestimations observed in the o-suite in the Sahel are also spread to the Tropical North Atlantic (MB of -0.13 for o-suite and control, see Table 8.2.1). The daily variability is better captured by CAMS in this region with correlation coefficients of 0.57 for control and 0.74 for o-suite.

In the case of the North-Western Maghreb (see Table 8.2.1 and Saada in Figure 8.2.3b), o-suite and control show higher correlation coefficients (0.82 for control and 0.83 for o-suite) and o-suite also reduces overestimations observed in control (MB of -0.02 for control and o-suite). Over the Iberian Peninsula and the Mediterranean, both CAMS runs show correlations between 0.60 and 0.85 and slightly overestimations in o-suite (MB between 0.03 and 0.05) than control (MB between -0.03 and -0.02). During this season, dust transport is limited, and is concentrated in western-central Mediterranean (see Saada and Lampedusa in Figure 8.2.3b).

The comparison of the 1- to 3-day forecasts shows that the prediction is stable during the forecasts in comparison with AERONET dust-filtered observations with correlation coefficients of 0.61 (0.56), 0.54 (0.54), and 0.52 (0.54) respectively for 24, 48 and 72h forecasts for all the sites, for o-suite (control).

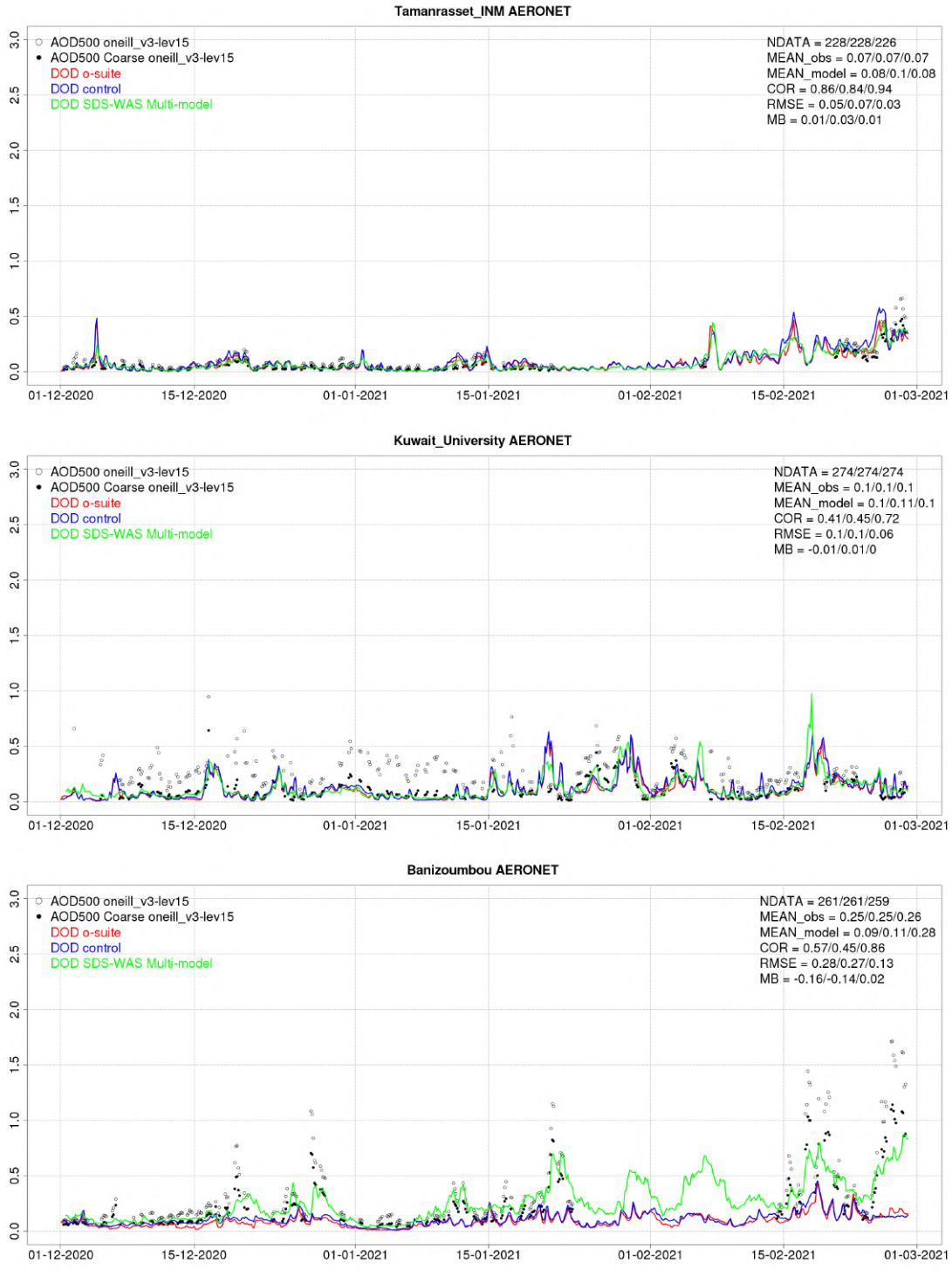


Figure 8.2.3a: AOD and Angstrom Exponent from AERONET Direct-sun (black dots), DOD o-suite (red line), DOD control (blue line) and DOD Multi-model SDS-WAS Median (green line) for the study period over Tamanrasset INM (Sahara), Kuwait University (Middle East) and Banizoumbou (Sahel). Skill scores for the three models (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).

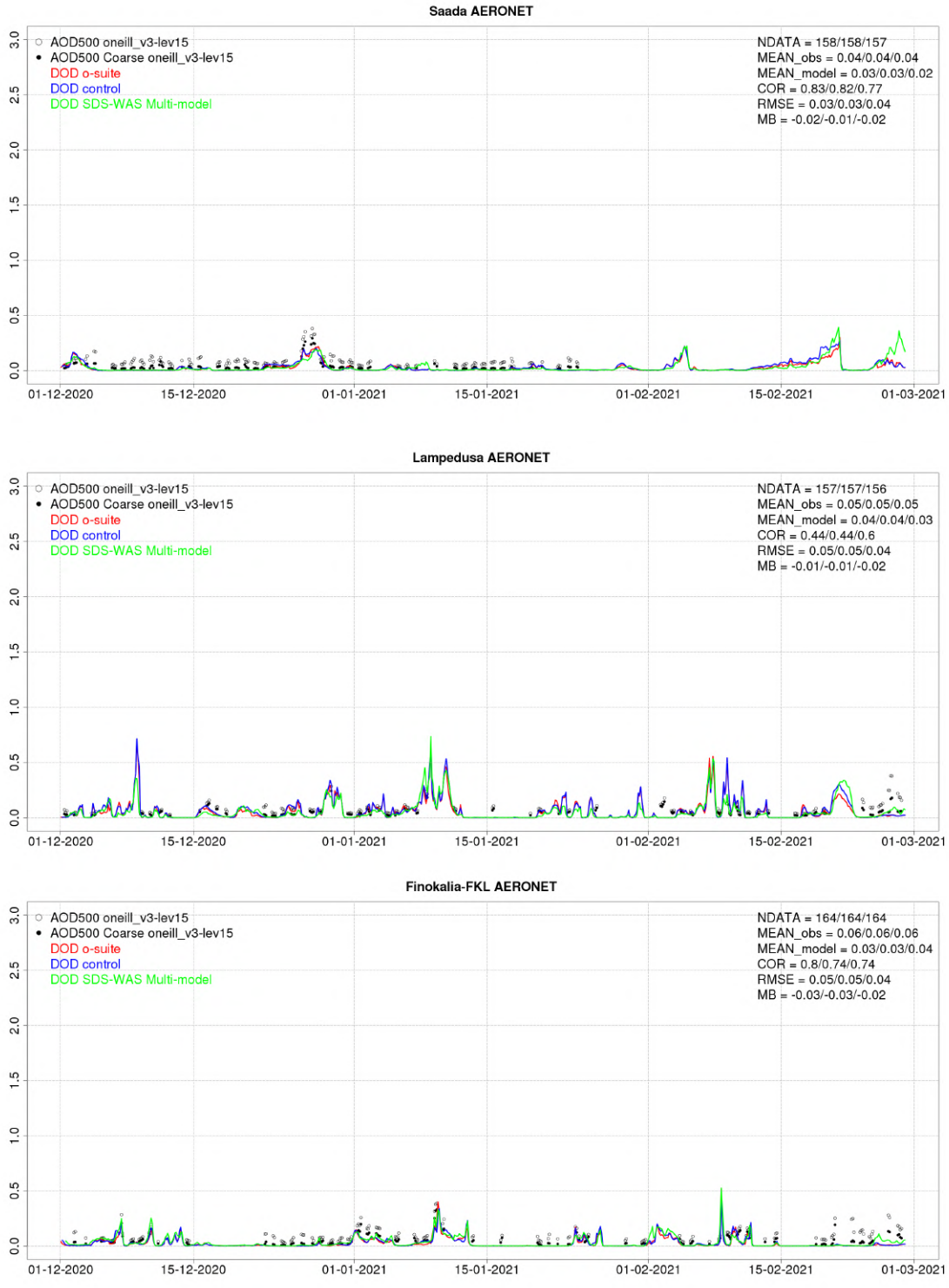


Figure 8.2.3b: AOD and AOD-coarse from AERONET SDA (black dots), DOD o-suite (red line), DOD control (blue line) and DOD Multi-model SDS-WAS Median (green line) for the study period over, Saada (NW Maghreb), Lampedusa (Central Mediterranean) and Finokalia-FKL (Eastern Mediterranean) Skill scores for each 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).

8.3 Aerosol validation over Europe and the Mediterranean

Three-hourly aerosol optical depth (AOD) and surface concentration (PM10 and PM2.5) from the o-suite and control run have been validated against AERONET AOD direct-sun cloud-screened and EEA PM10 and PM2.5 observations.

Aerosol optical depth over the Mediterranean

During winter, both CAMS runs do reproduce the daily variability of AERONET AOD observations, although they present a general overestimation in the whole Mediterranean Basin (see Figure 8.3.1). The correlation coefficient decreases from (0.78, 0.68 and 0.39) for control to (0.84, 0.71 and 0.41) for o-suite, and MB slightly increases from (-0.01, 0.03 and 0.04) for control to (0.04, 0.07 and 0.11) for o-suite respectively for the Western, Central and Eastern Mediterranean. Overestimations are linked to an enhanced background of aerosols that is not directly linked to natural contributions. This is shown in the Barcelona (Spain, Western Mediterranean), Lecce (Italy, Central Mediterranean) and Sede Boker (Israel, Eastern Mediterranean) AERONET sites (see Figure 8.3.2). During February, two intense Africa dust outbreaks were observed in the Western Mediterranean (see Barcelona in Figure 8.3.2). A more detailed analysis of these events is considered in Section 12.

Surface aerosol concentrations in Europe

At surface levels, both CAMS runs show a higher correlation coefficient in north-western Europe (above 0.7) in comparison with the 3-hourly EEA PM10 and PM2.5 observations (see Figure 8.3.3 and 8.3.4). For PM10, both CAMS runs show underestimations (MB under -4 $\mu g/m^3$) except in Central Europe and central-western Mediterranean which appear overestimated (MB above 4 $\mu g/m^3$). The PM2.5 comparison shows smaller differences, but the overestimations observed in PM10 for the o-suite in Central Europe are also detected in the PM2.5 o-suite comparisons (MB above 4 $\mu g/m^3$, see Figure 8.3.4).

During winter, lower PM values are observed linked to higher precipitation. As also observed in the AOD comparison, February shows two high PM10 and PM2.5 peaks that are coincident with dust outbreaks, associated with concentrations above 50 $\mu g/m^3$ for PM10 in southern and central European sites (see Figure 8.3.5). In most of the cases the EEA surface observations are overestimated (see Figure 8.3.5). PM10 levels are systematically overestimated in Malta (Central Mediterranean, MT000007 in Figure 8.3.5a). These overestimations in the island of Malta are associated with coarse particles (as it is indicated the low observed PM2.5/PM10 ratio) with maritime origin. These extreme PM10 peaks are also observed in other sites (as FR35005 in Figure 8.3.5a).

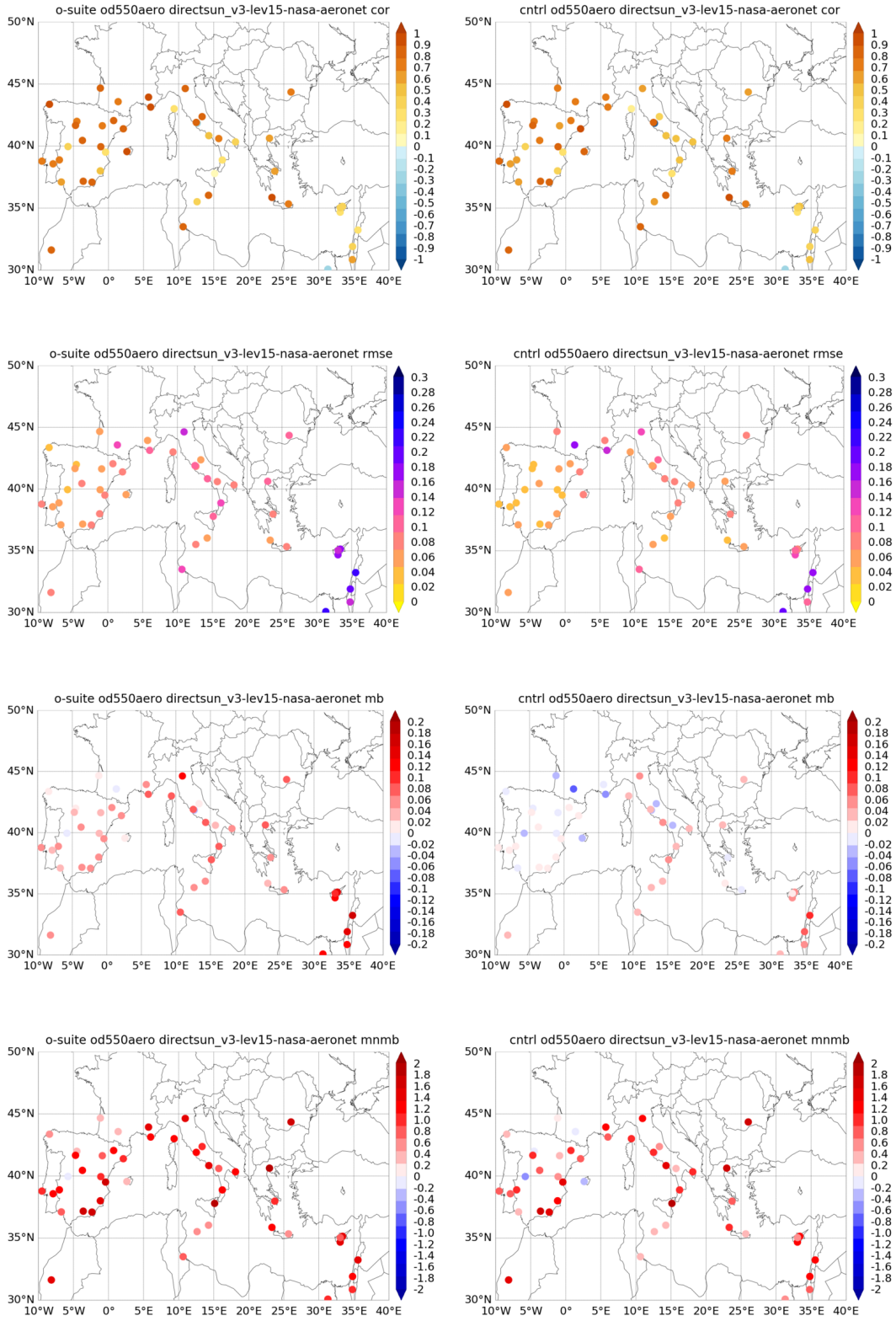


Figure 8.3.1: Skill scores (correlation coefficient, RMSE, MB and MNMB) for 24-hour forecasts of CAMS o-suite and control for the study period. AOD from AERONET direct sun is the reference.

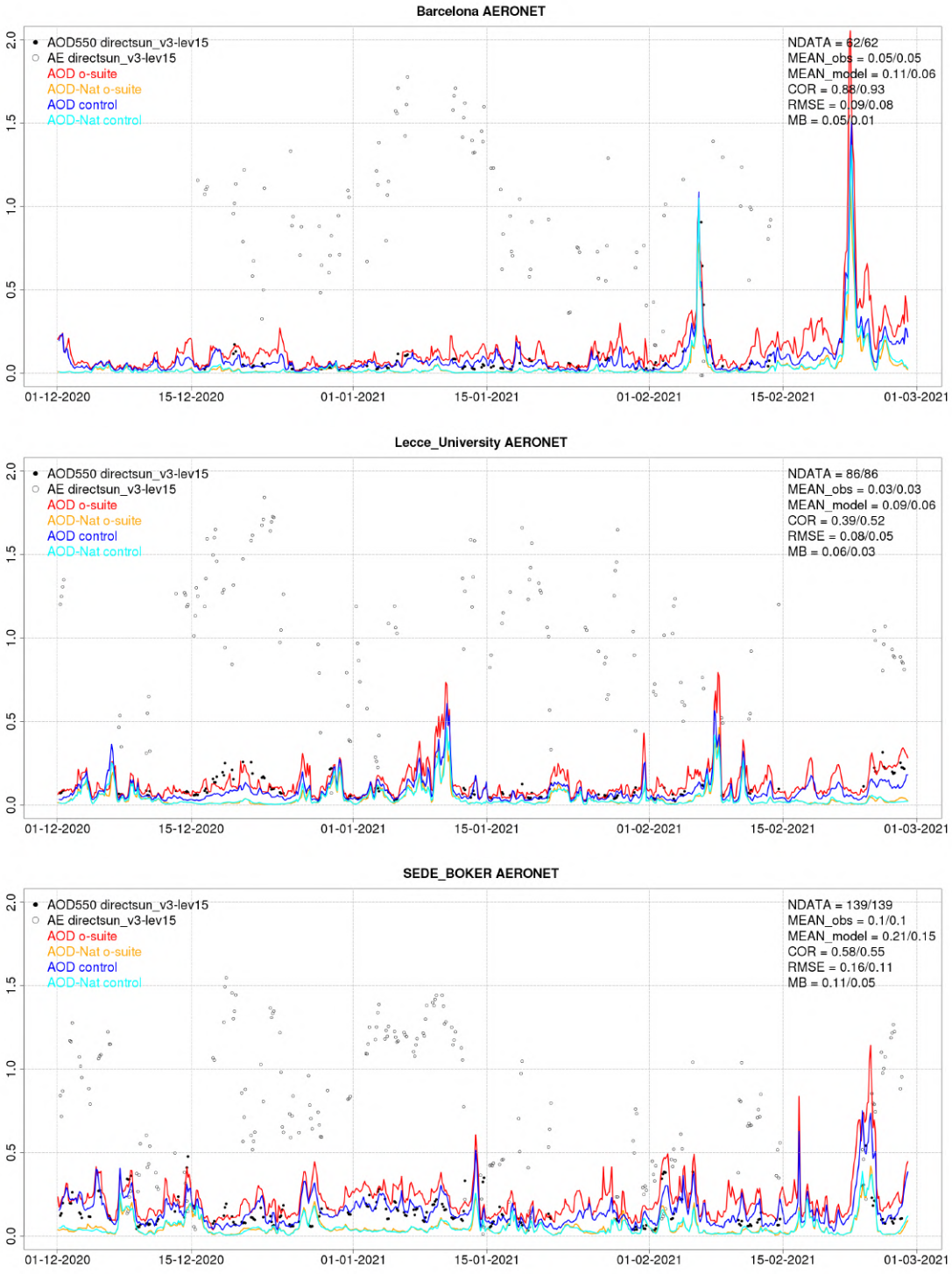


Figure 8.3.2: AOD from AERONET (black dot), AOD o-suite (red line), AOD control (blue line), AOD-Nat o-suite (orange line), AOD-Nat control (cyan line), for the study period over Barcelona (Spain), Lecce University (Italy) and SEDE BOKER (Israel). AOD-Nat corresponds to the natural aerosol optical depth that includes dust and sea-salt. Skill scores per 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).

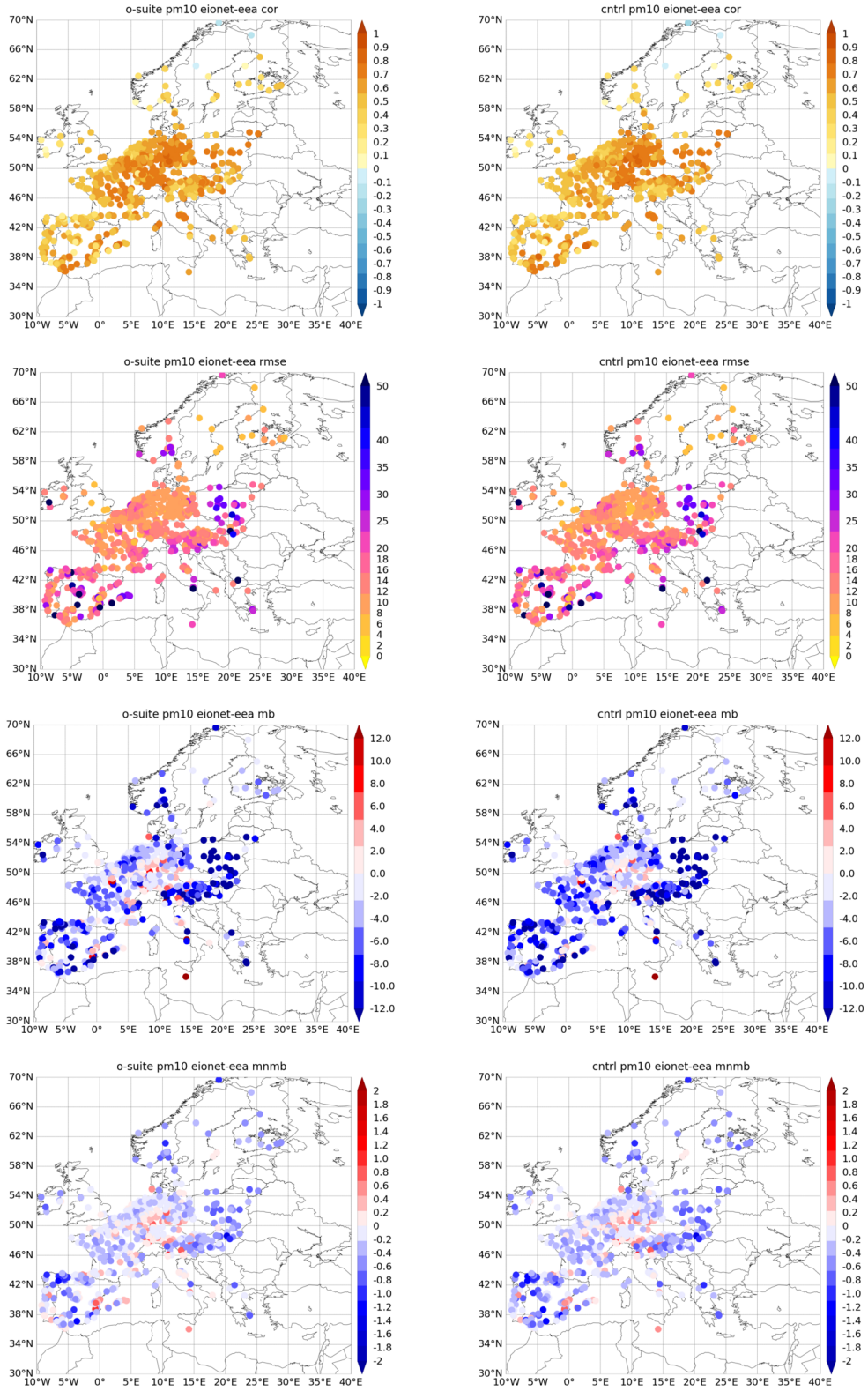


Figure 8.3.3: Skill scores (correlation coefficient, RMSE, MB and MNMB) for 24-hour forecasts (at 3hourly basis) of CAMS o-suite and control for the study period. 3hourly PM10 from EIONET is the reference. Only global scale representative stations are displayed.

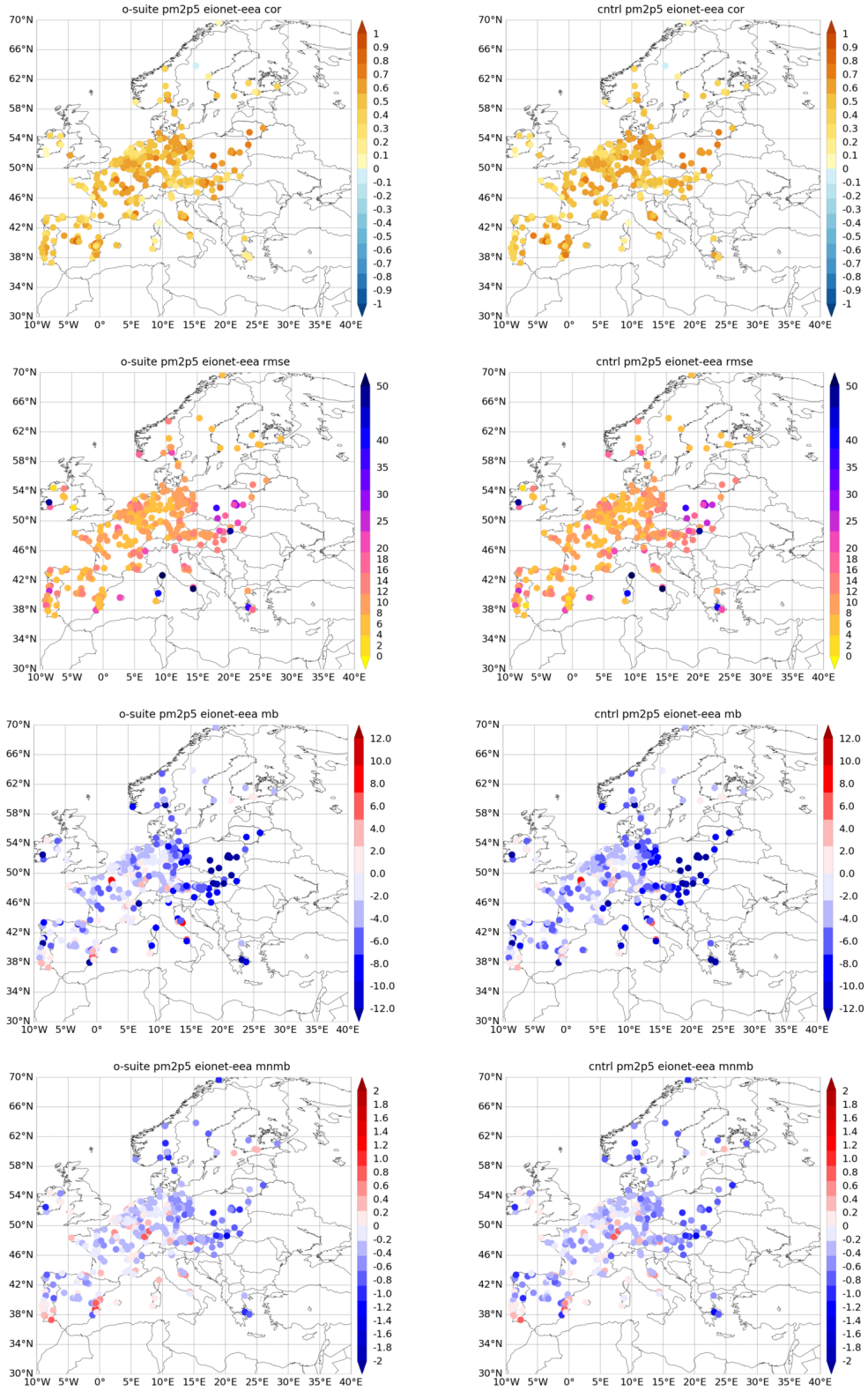


Figure 8.3.4: Skill scores (correlation coefficient, RMSE, MB and MNMB) for 24-hour forecasts (at 3hourly basis) of CAMS o-suite and control for the study period. 3hourly PM2.5 from EIONET is the reference. Only global scale representative stations are displayed.

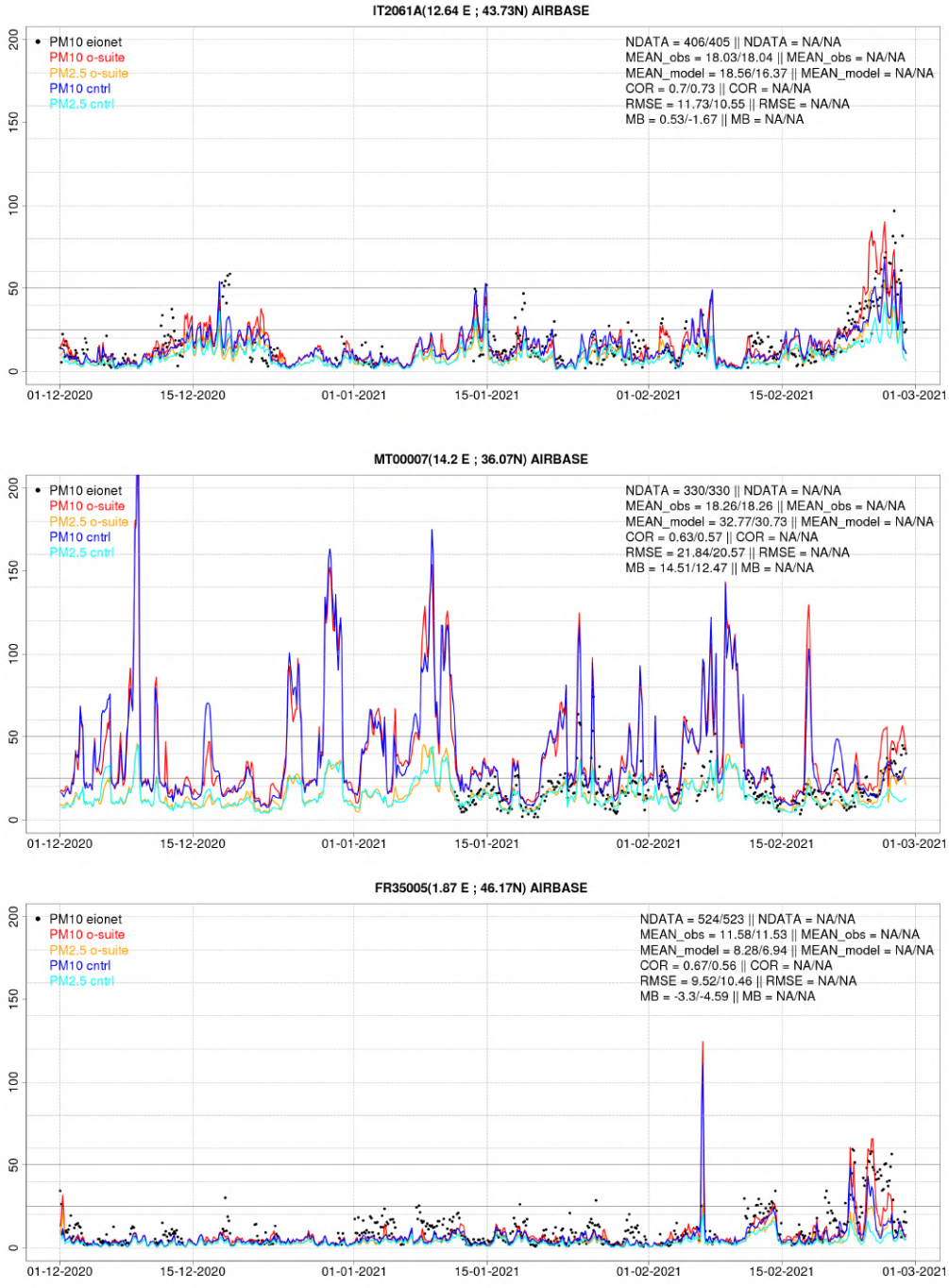


Figure 8.3.5a: PM10 and PM2.5 Airbase observations (black and grey dots, respectively), PM10 and PM2.5 o-suite (red and orange lines, respectively) and PM10 and PM2.5 control (blue and cyan lines, respectively) for the study period over IT2061A (Italy), MT00007 (Malta) and FR35005 (Central France). Skill scores per 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).

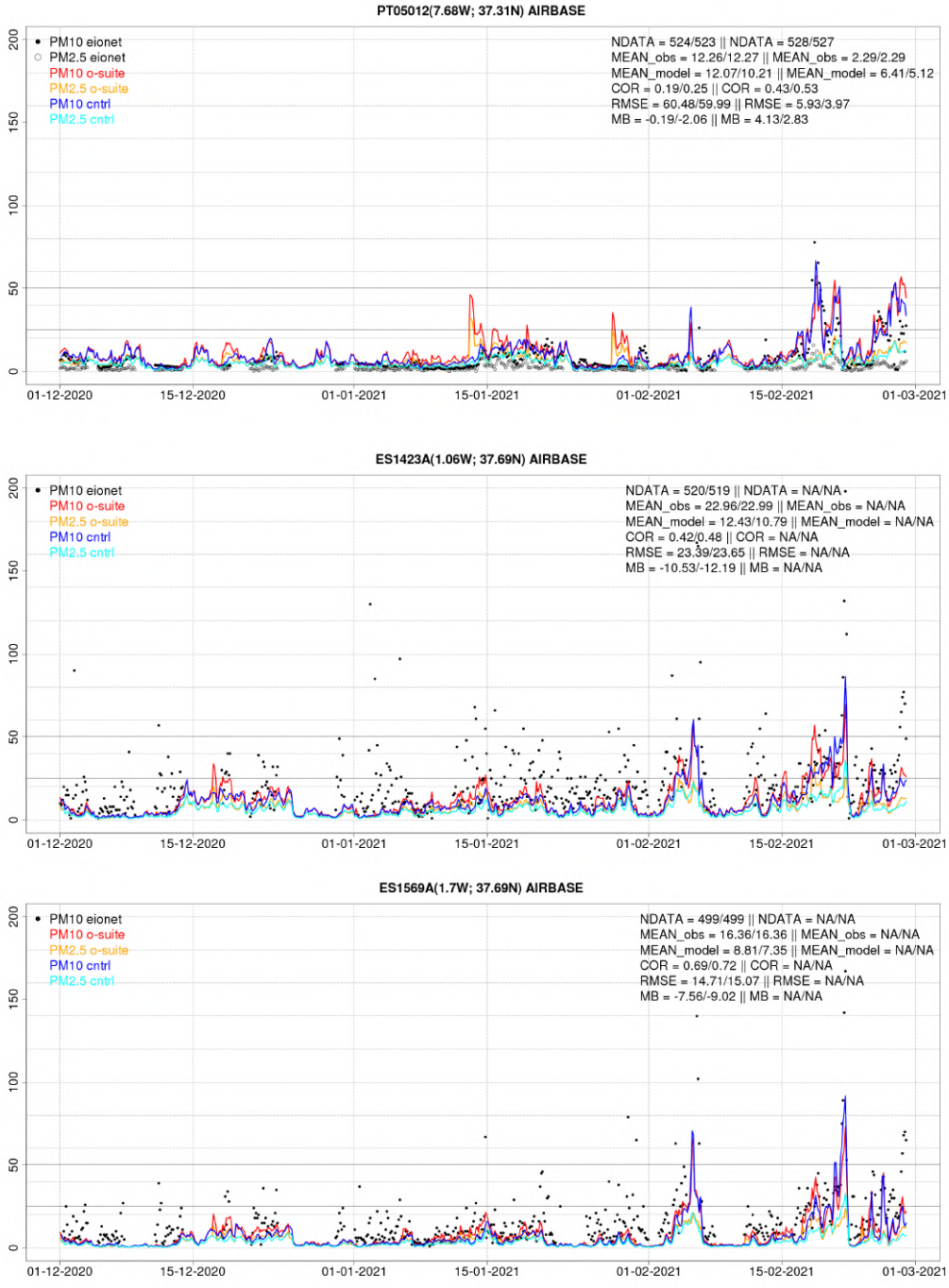


Figure 8.3.5b: PM10 and PM2.5 Airbase observations (black and grey dots, respectively), PM10 and PM2.5 o-suite (red and orange lines, respectively) and PM10 and PM2.5 control (blue and cyan lines, respectively) for the study period over PT05012 (Portugal), ES1423A and ES1569A (Spain, SE Iberian Peninsula). Skill scores per 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).

8.4 Ceilometer backscatter profiles

Technical specifications of the German CHM15k ceilometer network, evaluated parameters and methods are described in report CAMS-84 D8.1 and more recently have been published in a GMD article (Flentje et al., 2021, https://gmd.copernicus.org/articles/14/1721/2021/). This paper provides a summary of model performance over Germany in the years 2016-2019. In this section, the temporal and vertical variation of the attenuated backscatter coefficient (absc) profiles are evaluated, statistically as bias (model minus observation), modified normalized mean bias (MNMB), correlation, and standard deviation of o-suite and control run 'hdir' vs ceilometers, and summarized in Taylor plots. All evaluations refer to the domain Germany and, if not noted otherwise, means or medians are presented for 21 ceilometer stations selected according to calibration performance and data coverage.

The model aerosol optical depth (AOD) and ceilometer overviews are used to select periods with significant aerosol plumes over Germany for specific attention and case studies. The CAMS attenuated backscatter (derived from aerosol mass mixing ratios by means of a forward operator built by DWD) and AOD are speciated for contributions of mineral dust (MD), sea salt (SS), carbonaceous matter (CM), black (BC) and organic carbon (OC), sulphate (SU), nitrate (NI), as well as ammonia (AM).

Period Overview

Significant Sahara dust (SD) events occurred around 5.-6.2.2021 and around 22-26.02.2021. As before, the time and location of these events were forecasted well but the dust load was strongly overestimated. This behaviour has also been discussed in previous reports. Around 24-26 Feb 2021 a sulphate plume, freshly emitted by the Etna volcano (Sicily) reached Germany, but its transport was not associated with high SO_4 AOD ≈ 0.65 over Germany. Further, elevated SO_4 AOD around 0.35 occurred on other days as well. No elevated sulphate plume was detected by the ceilometers. It has to be checked whether precursor SO_2 from volcanic emission was switched on in the model or could have been introduced by the assimilation. Surface in situ observations observed particle nucleation from SO_2 as very small particles (<40 nm) containing little mass load. Such small particles do not cause a signature evident in ceilometer profiles nor in AOD. AOD maps indicate a different origin than Etna.

Figures 8.4.1 and 8.4.2 show the temporal variation of bias respective modified normalized mean bias MNMB at different heights above ground (starting at 1km) during 2016-2020, for assimilation (reddish) and control (bluish) runs. The time series are smoothed by a 7-day moving average and indicate significant model upgrades by vertical dashed lines (cf. Flentje et al, 2021). MNMB are less dependent on the absolute backscatter values and thus also provide a more consistent scale of biases and height-covariations throughout the vertical profile. In MNMB the characteristic changes between the different IFS-AER cycles can be seen more clearly than in the absolute bias.

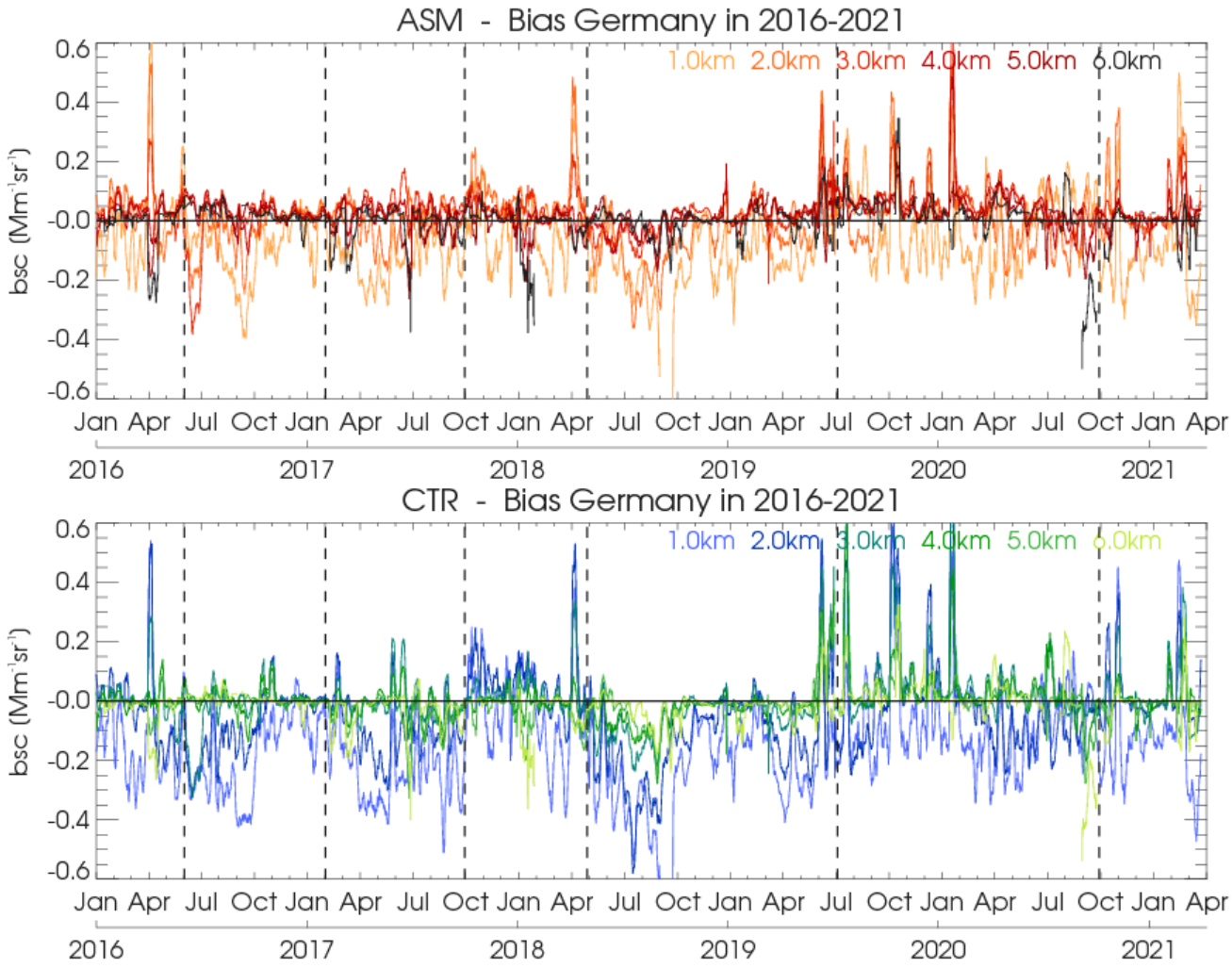


Figure 8.4.1: Bias of attenuated backscatter over Germany for different altitudes over the period 2016-2021 for o-suite (top) and control run (bottom), smoothed by a 7-day moving average. Altitudes 1, 2,…, 6 km above ground are shown, colour coded orange-to-black (o-suite) and blue-to-green, respectively.

Except for the 1 km level in the planetary boundary layer PBL and the 6 km level in the free troposphere FT the 7-day average bias during the 2016-2021 period was mostly within ±0.1 $Mm^{-1}sr^{-1}$. In the PBL, IFS-AER backscatter is biased low for both o-suite/ASM and (even lower) control/CTR with little change over the last 5 years. In the middle FT high bias peaks are usually associated with Saharan dust events, so for example in Feb 2021. Sporadic bias drops at 6 km (e.g., Sep 2020) are caused by remaining cloud artifacts that have not been captured by the cloud mask cleaning the ceilometer profiles. At the lowest level (≤ 1 km above ground) the model change to cycle 47r1 in Oct 2020 increases the low bias for the DJF 20/21 period. With cycles before 47r1, 'near-surface' backscatter (≈0.4 km above ground) tended to be overestimated during summers but was captured quite well in winter with cycle 46r1 (Flentje et al., 2021). The summer bias in cycle 47r1 will be discussed in the JJA-2021 report.

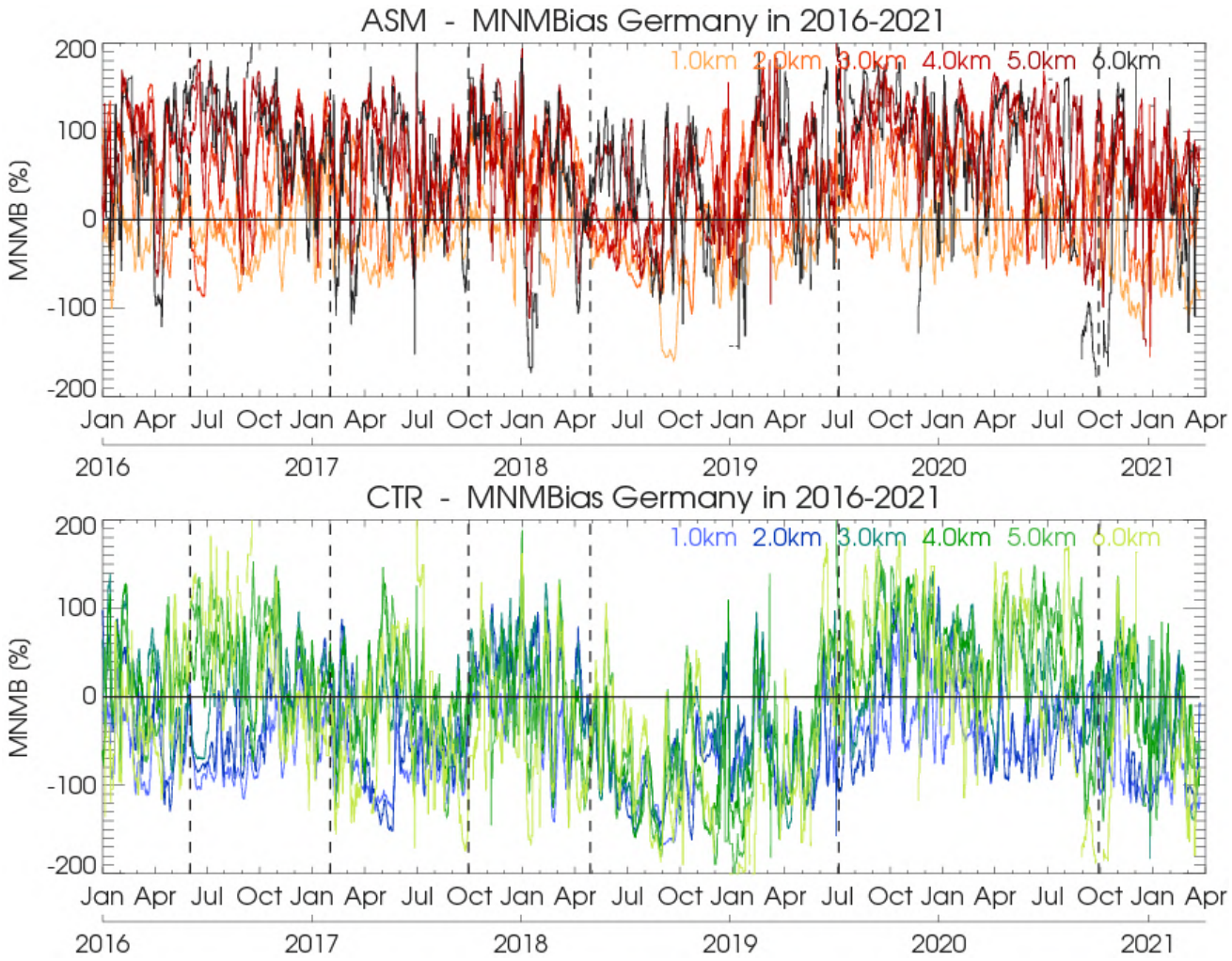


Figure 8.4.2: As Fig. 1 but for modified normalized mean bias MNMB.

According to the change log at ECMWF, the implementation of cycle 47r1 brought no changes to aerosol representation/sources, but possibly the changes to convective inhibition may have led to increased vertical transport, diluting surface emissions too fast to upper levels. A similar low bias, however, has also been observed with cycle 45r1 in parts of autumn 2018 already, independent from assimilation. A bias signature of C-19 induced emission reductions, possibly not fully realized in the model(?) is not seen. Owing to the 7-day moving averaging, short term anthropogenic pollution events, which previously have been underestimated by the IFS model, are unlikely to contribute to the low bias (there were not strong pollution events over Germany in autumn 2020 anyway). On the long term the o-suite exhibits a bit smaller bias then the control runs (bottom), indicating a positive impact of the assimilation (Fig. 8.4.1, 2). Fig 8.4.3 shows the PBL bias over the last year with a very small difference between o-suite and control in Jan and Feb 2021 and possibly a bit larger low bias with cycle 47r1 since Oct 2020.

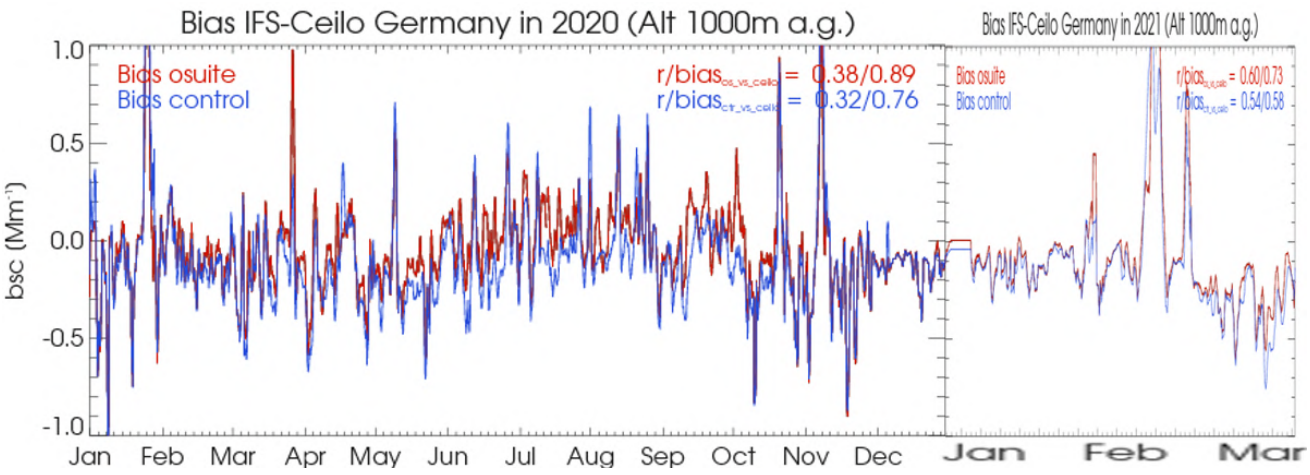


Figure 8.4.3: Daily average model-ceilometer bias in the PBL (1000 m a.g.) from 21 ceilometer stations distributed over Germany in 2020.

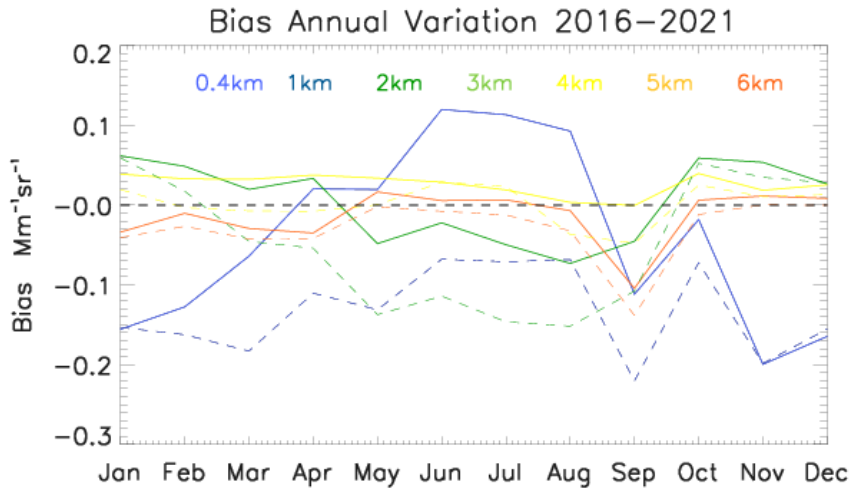


Figure 8.4.4: 5-year mean monthly attenuated backscatter bias seasonal cycles over Germany for different altitudes in the period 2016 to 2020.

Though seasonal regularities are disturbed by six irregular model updates in the 2016–2020 period, bias/MNMB in ASM show opposing seasonal cycles in the lower (0.4 km above ground) and the upper (2 km above ground) mixing layer with amplitudes of 0.2 $Mm^{-1}sr^{-1}$ = 40% (summer maximum) and 0.1 $Mm^{-1}sr^{-1}$ = 70% (summer minimum), respectively (Fig. 4). The seasonal amplitude is small at the intermediate 1 km level. The summer minimum is evident up to 3 km (MNMB even to 4 km), while it is variable due to Saharan dust events at 5 and 6 km altitude. Generally, the o-suite high bias near the surface (0.4 km above ground) in summer is larger than in the control run, while the low bias at 2 km is more expressed in the control run (Figure 8.4.4). Thus, assimilation seems to have a significant but opposing effect in the lower and the upper mixing layer.

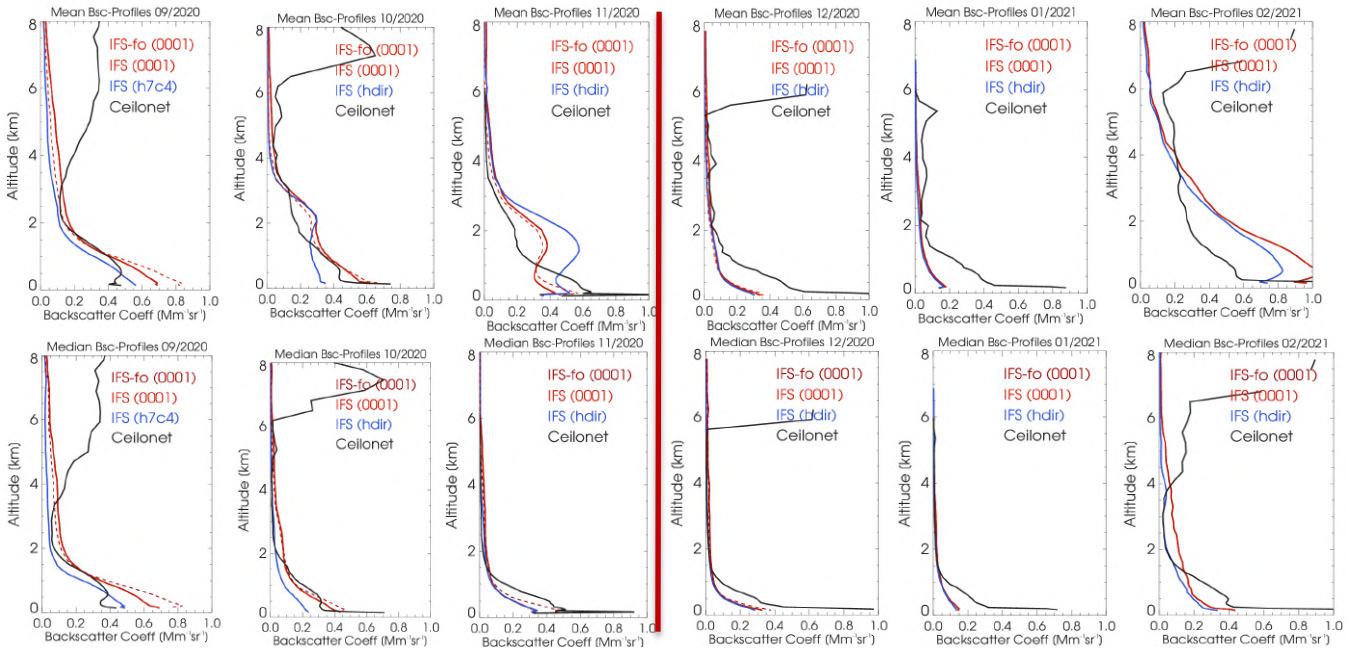


Figure 8.4.5: Left SON 2020, right DJF 2020/21 monthly mean profiles (upper panel) and median profiles (lower panel) of attenuated backscatter from o-suite (red), control run (blue), and ceilometers (black) combined from 21 German stations.

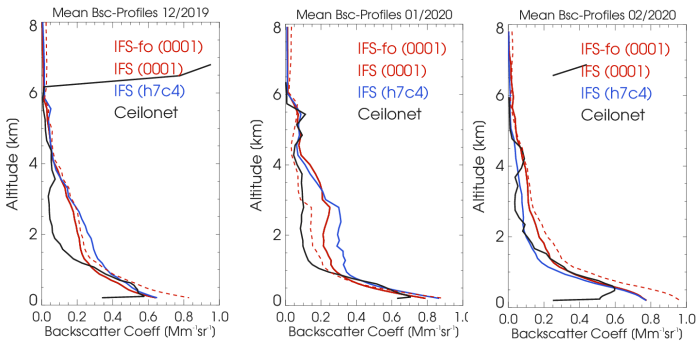


Figure 8.4.6: As Fig. 5 but monthly mean profiles for DJF 2019/20.

Mean and Median profiles (Fig. 8.4.5, 8.4.6):

With the new IFS cycle 47r1, implemented in Oct 2020, no changes to the aerosol representation came into IFS-AER. Thus, it was expected that nitrate NO_3, ammonia NH_4 and sulphate SO_4, would still be significantly overestimated near the surface as reported by Flentje et al., 2020. As they contribute roughly 10-30% to wintertime aerosol mass in the rural central European PBL, as neutralized forms NO_3NH_4 or $(NH_4)_2SO_4$, this should have a corresponding effect on backscatter in the lower mixing layer as well. Also, mineral dust and sea salt are likely still high biased over Germany in the present IFS-AER model.

Unlike before, modelled mean and median attenuated backscatter in the DJF 2020/21 period were much lower than observed with the ceilometers and there was little difference between o-suite and control run. Double checks of this result yet revealed no bug in data retrieval and processing but

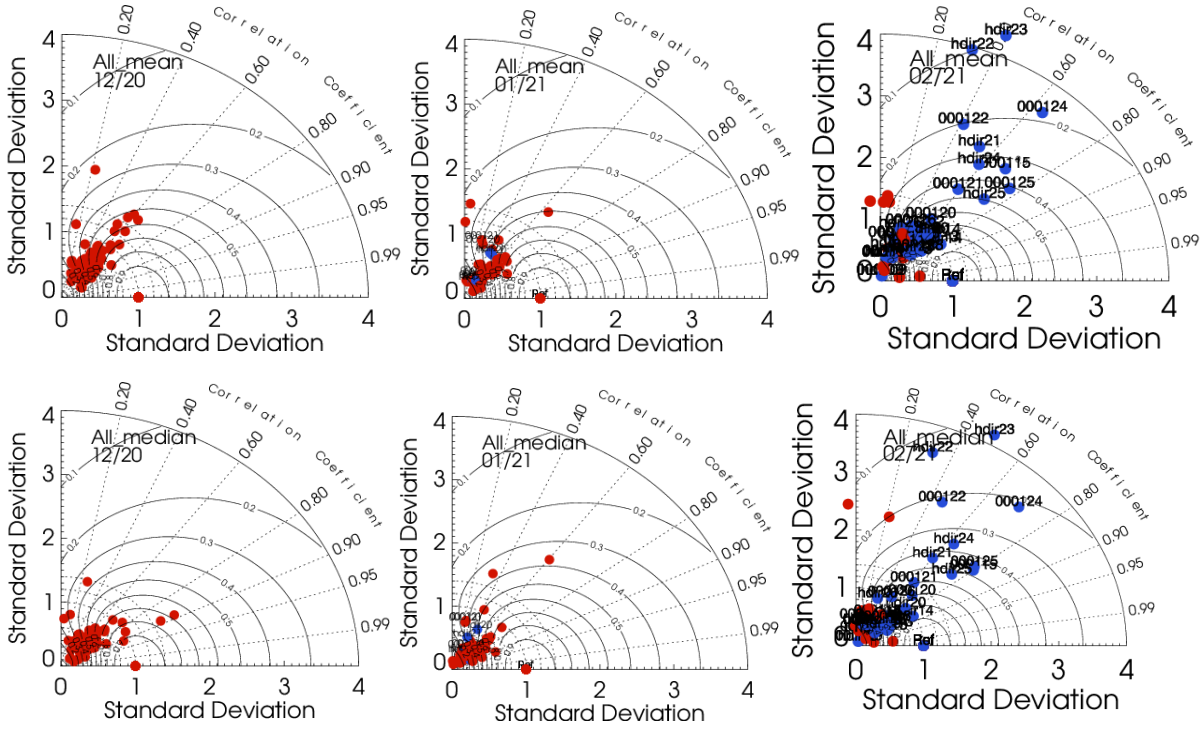


Fig. 8.4.7: Taylor polar plots with daily average standard deviation of vertical profiles vs correlation coefficient, averaged over 21 German ceilometer sites for Dec to February 2021 with mean values (top) and median values (bottom). O-suite red, control blue.

this will be further investigated. In Feb 2021 two strong Saharan dust events lead to the well-known high bias in the profiles and the event-related gap between mean and median profiles. Even higher bias of the o-suite profile indicates an adverse effect from the assimilation. In comparison to DJF 2019/20, there is a clear degradation of model performance with respect to the vertical backscatter profile for yet unknown reasons.

Taylor Plots (Fig. 8.4.7):

The average coefficient of correlation between modelled and observed vertical backscatter profiles ranges between r = 0.2 – 0.8 in DJF 2021 (Fig. 8.3.7). As shown in Figure 8.3.6 the Feb 21 profiles are strongly affected by 9 days Saharan dust events with AOD = 0.05 – 0.55 over Germany. December and January are low biased below 1.5 km height yielding normalized standard deviation NSD < 1 (≡ reference, i.e., the profile amplitude of the observations). A large difference between mean and median profiles indicates the strong events in the Taylor plots as well. December and January mostly exhibit correlations around r = 0.6 and NSD ≈ 0.2 - 1.5. In February the overestimation of dust load reflects in NSD > 1 on Saharan dust days and shows moderate correlation.

The o-suite and control run show very similar correlation and NSD during December and January, thus both are not distinguishable in the plots. This is quite suspicious and will further be investigated. In February, however, a normal spread with large NSD in presence of dust is observed, confirming the 'sanity' of the evaluation.

9. Stratosphere

9.1 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 February 2021 (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. (2019). Please note that recent scientific findings (https://tropo.gsfc.nasa.gov/shadoz/Archive.html, Thompson et al., 2017; Witte et al., 2017; 2018, Stauffer, et al. in preparation 2020) show a drop-off in Total Ozone at various global ozone stations in comparison with satellite instruments. This drop-off amounts between 5-10% for stratospheric ozone. Changes in the ECC ozone instrument are associated with the drop-off, but no single factor has been identified as cause yet.

The o-suite shows MNMBs within the range from -5 to 13% for Antarctic with maximum biases in October-November; from -2 to 8% for the Arctic (maximum in March-April); and within 5% for northern mid-latitudes. Figure 9.1.1 shows the results for the past year.

Fig. 9.1.2 compares the averaged profiles in each region during December 2020. The o-suite vertical distribution of stratospheric ozone is in good agreement with the observations, showing just slight overestimation. The control run captures well the vertical distribution over the Northern Hemisphere mid-latitudes and Arctic and shows overestimation over Antarctica.

For the period DJF-2021: The new control run (expid=hdir) shows strong improvement for the Arctic and the Northern mid-latitudes especially in the UTLS and stratosphere.

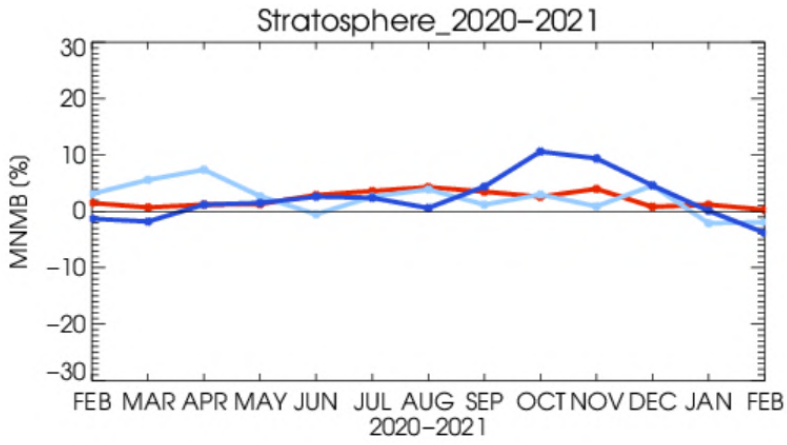


Figure 9.1.1: MNMBs (%) of ozone in the stratosphere from the o-suite against aggregated sonde data in the Arctic (light blue), Antarctic (dark blue) and northern mid-latitudes (red). Period December 2020 to February 2021. The stratosphere is defined as the altitude region between 90 and 10 hPa.

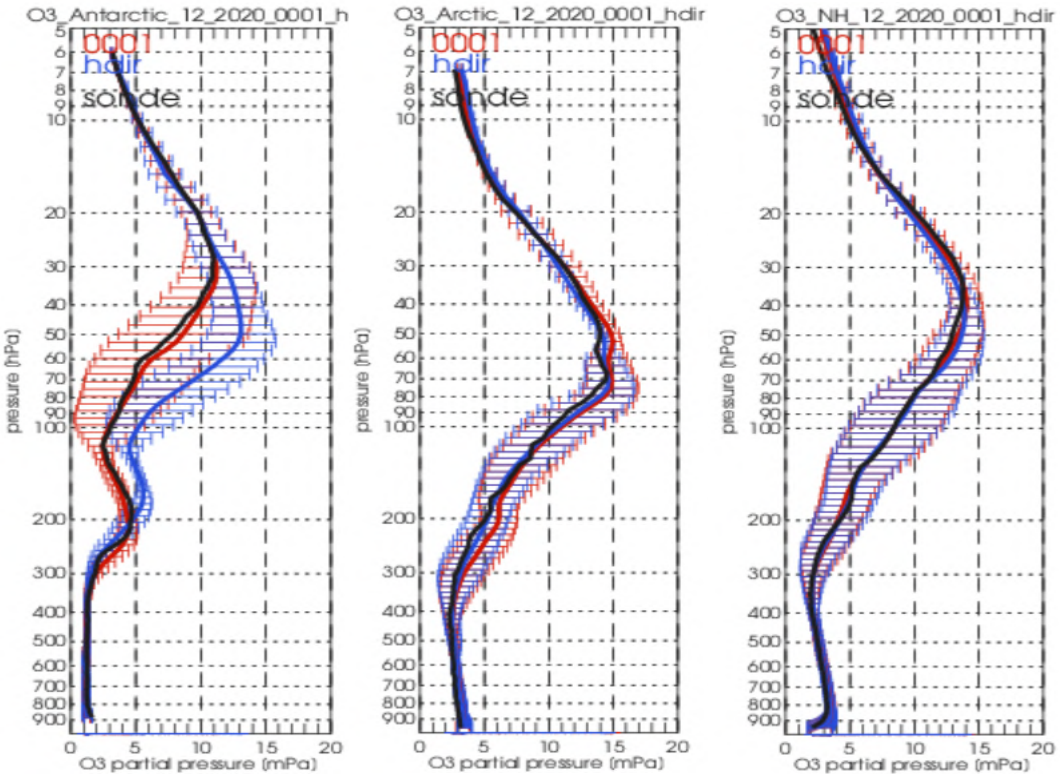


Figure 9.1.2: Comparison between mean O_3 profiles (units: mPa) of o-suite (red), and control (blue) in comparison with observed O_3 sonde profiles (black) for December 2020 for the various latitude bands: Arctic, NH-midlatitudes and Antarctic.

9.2 Validation against observations from the NDACC network

UVVIS column and FTIR stratospheric columns

Since the start of the CAMS27 project, the number of UVVIS Zenith ozone measurements have increased on NDACC. Currently 15 sites provided data in the recent quarter allowing for a representative picture on the latitude dependence of the CAMS data. Since 2019 also DOBSON measurements are supported by CAMS27 and delivered to NDACC more rapidly.

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 9.2.1. The averaged bias for the 15 UVVIS sites is 2% and within the reported measurement uncertainty of 5%, the averaged correlation is 0.86 during this quarter and did not change significantly compared to the previous quarter. A similar conclusion holds for the DOBSON comparisons: biases are comparable to the measurement uncertainty which typically lies between 2% and 3%.

The correlations between the individual sites and the CAMS runs are presented in the Taylor diagrams in Figure 9.2.2. Again, the o-suite AN and 1d FC perform very similarly in correlation coefficients. Figure 9.2.3 shows the evolution of the relative bias of the stratospheric columns against FTIR data since 2015.

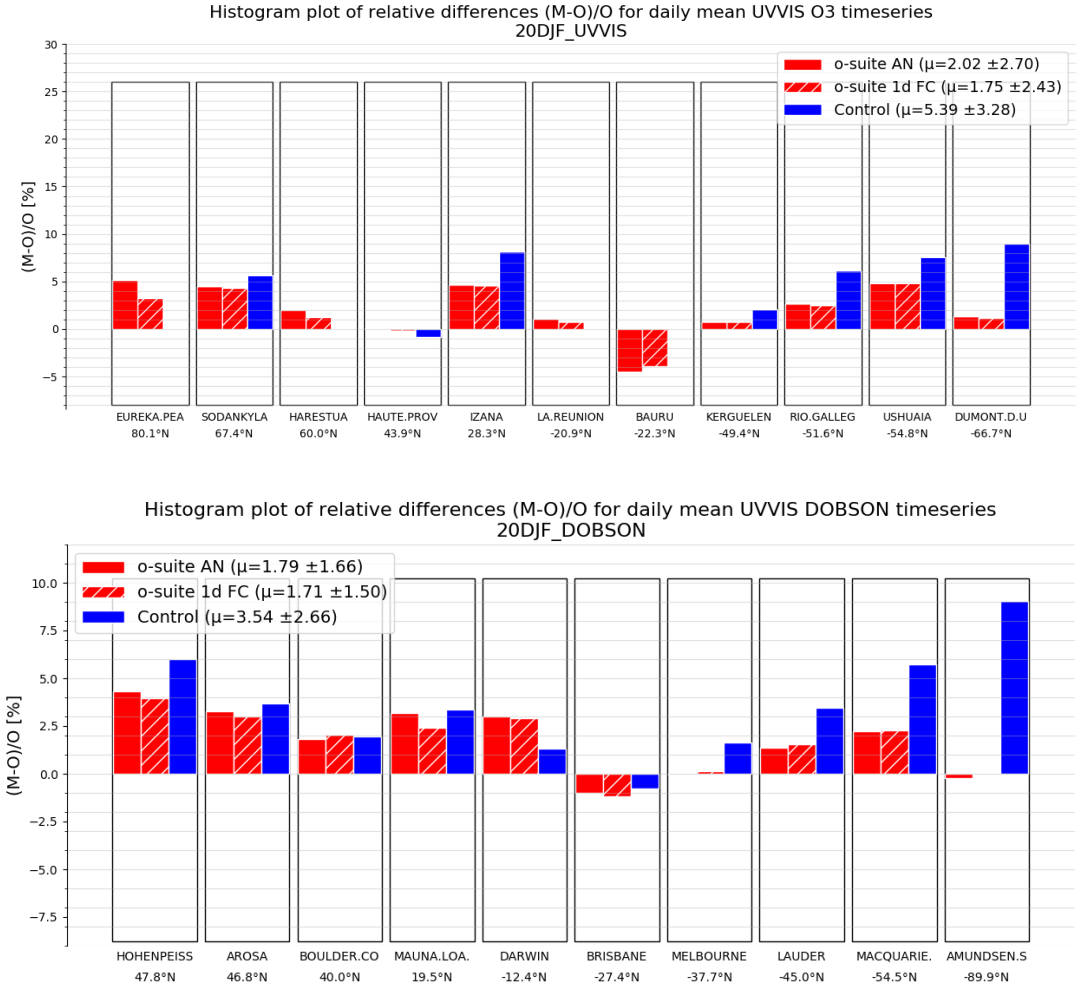


Figure 9.2.1 Relative biases during quarter DJF 2021 for 11 UVVIS DOAS stations measuring stratospheric ozone columns with ZENITH measurement geometry and 10 DOBSON instruments (stations sorted with decreasing latitude). The overall relative bias is positive for almost all latitudes and comparable to the typical measurement uncertainty of 5% for UVVIS ZENITH and 2.5% for DOBSON for most of the sites.

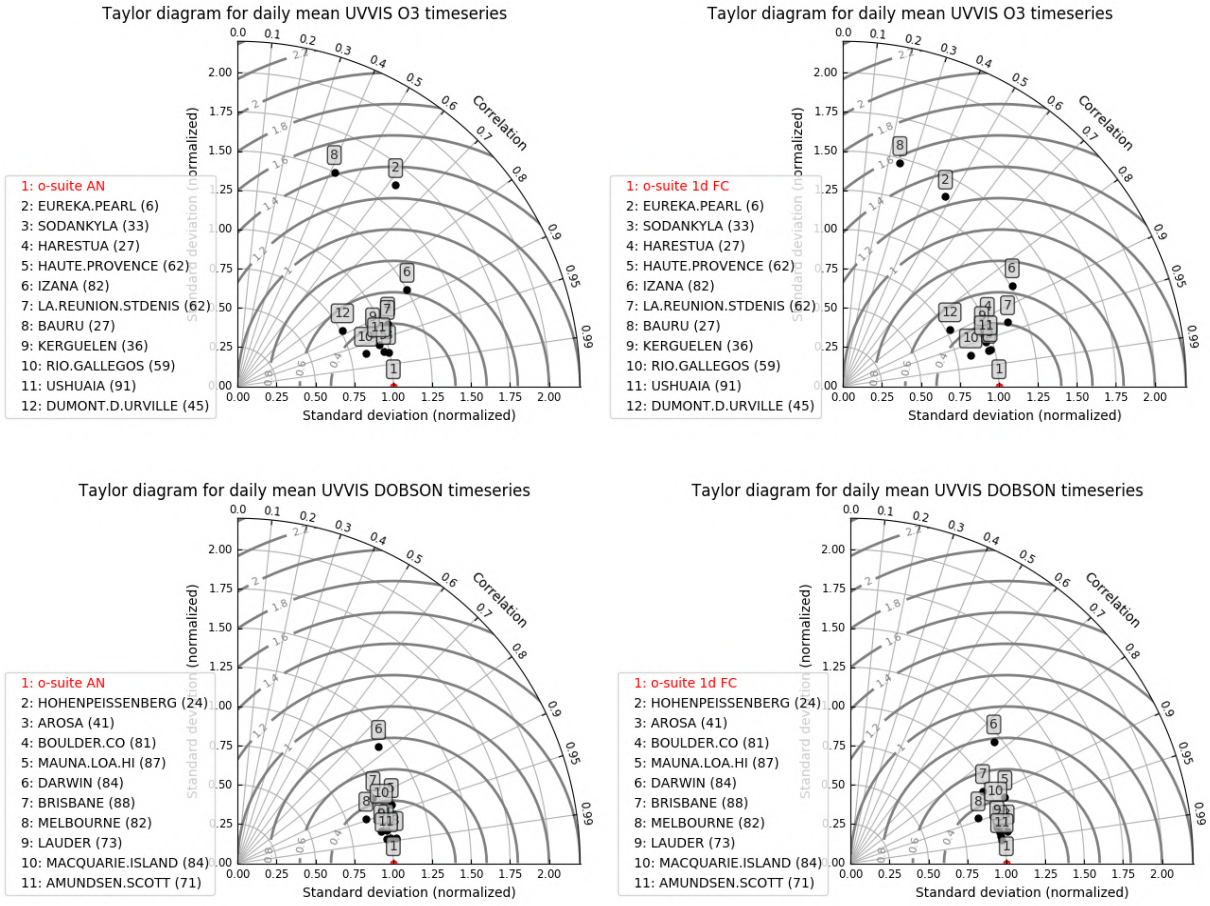


Figure 9.2.2. Taylor diagrams relating the standard deviations for the model and GB stratospheric column time series and their correlation for the period DJF 2021. All time-series are normalized such that the standard deviation of the model is 1. The performance for the o-suite is slightly better (averaged correlation is 0.97 for FTIR, 0.94 for DOBSON and 0.86 for UVVIS) compared to the 1-day forecast (averaged correlation is 0.96 for FTIR, 0.93 for DOBSON and 0.83 for UVVIS).

Table 9.2.1: Detailed statistics for stratospheric ozone column comparisons for UVVIS (zenith and Dobson) and FTIR measurements during DJF 2021. Standard deviations (std) are relative to the std of the o-suite.

	o-suite AN					o-suite 1d FC					
UVVIS ZENITH site	#	rel. std	corr	rel diff (%)	rel diff std(%)	#	rel. std	corr	rel diff (%)	rel diff std(%)	lat
EUREKA.PEARL	6	1.6	0.62	5.11	2.61	6	1.4	0.48	3.24	2.97	80.1
SODANKYLA	33	1	0.98	4.47	2.53	33	1	0.97	4.29	2.74	67.4
HARESTUA	27	1	0.92	1.95	3.53	27	1	0.92	1.23	3.69	60
HAUTE.PROVENCE	62	1	0.97	0.05	2.08	62	1	0.97	-0.14	2.2	43.9
IZANA	82	1.3	0.87	4.65	2.59	82	1.3	0.86	4.6	2.68	28.3
LA.REUNION.STDENIS	62	1	0.93	1.07	1.09	62	1.1	0.93	0.72	1.07	-20.9
BAURU	27	1.5	0.41	-4.52	2.19	27	1.5	0.25	-3.91	2.49	-22.3
KERGUELEN	36	0.9	0.93	0.76	2.12	36	1	0.93	0.77	2.09	-49.4
RIO.GALLEGOS	59	0.8	0.97	2.62	2.33	59	0.8	0.97	2.49	2.33	-51.6
USHUAIA	91	0.9	0.96	4.77	3.21	91	1	0.96	4.81	3.27	-54.8
DUMONT.D.URVILLE	45	0.8	0.88	1.3	4.36	45	0.8	0.88	1.14	4.23	-66.7
		1.1	0.86	2.02	2.6		1.1	0.83	1.75	2.71	

	o-suite AN					o-suite 1d FC					
UVVIS DOBSON site	#	rel. std	corr	rel diff (%)	rel diff std(%)	#	rel. std	corr	rel diff (%)	rel diff std(%)	lat
HOHENPEISSENBERG	24	1	0.99	4.33	2.06	24	1	0.98	3.95	2.18	46.8
AROSA	41	1	0.99	3.25	1.83	41	1	0.98	2.98	2.26	40
BOULDER.CO	81	1	0.98	1.8	2.21	81	1	0.98	2.07	2.12	19.5
MAUNA.LOA.HI	87	1.1	0.94	3.2	1.76	87	1.1	0.92	2.39	1.92	-12.4
DARWIN	84	1.2	0.77	3	1.69	84	1.2	0.76	2.89	1.72	-27.4
BRISBANE	88	1	0.9	-1.01	1.6	88	1	0.88	-1.16	1.79	-37.7
MELBOURNE	82	0.9	0.95	-0.03	1.88	82	0.9	0.94	0.16	1.92	-45
LAUDER	73	0.9	0.98	1.34	1.74	73	1	0.97	1.55	1.85	-54.5
MACQUARIE.ISLAND	84	1	0.94	2.24	3.32	84	1	0.94	2.25	3.38	-77.8
AMUNDSEN.SCOTT	71	1	0.99	-0.22	2.21	71	1	0.99	-0.01	2.03	-89.9
		1	0.94	1.79	2.03		1	0.93	1.71	2.12	

	o-suite AN					o-suite 1d FC					
FTIR site	#	rel. std	corr	rel diff (%)	rel diff std(%)	#	rel. std	corr	rel diff (%)	rel diff std(%)	lat
ST.PETERSBURG	6	1	1	-2.77	1.8	6	1	1	-2.19	1.74	59.9
TORONTO	23	1	0.95	-1.8	3.86	23	1	0.96	-1.91	3.85	43.6
RIKUBETSU	5	1.1	0.97	-4.15	0.86	5	0.9	0.97	-4.44	1.02	43.5
BOULDER.CO	25	0.9	0.92	-4.2	3.15	25	0.9	0.94	-4.11	2.68	40
IZANA	3	1.3	1	-2.2	1.09	3	1.5	0.94	-2.29	2.25	28.3
MAUNA.LOA.HI	28	1.3	0.96	-5.98	1.61	28	1.3	0.94	-7.04	1.81	19.5
LAUDER	47	1	0.97	-3.69	1.73	47	1	0.97	-3.44	1.72	-45
ARRIVAL.HEIGHTS	21	1	0.99	-2.72	1.21	21	1.1	0.99	-2.66	1.32	-77.8
		1.1	0.97	-3.44	1.91		1.1	0.96	-3.51	2.05	

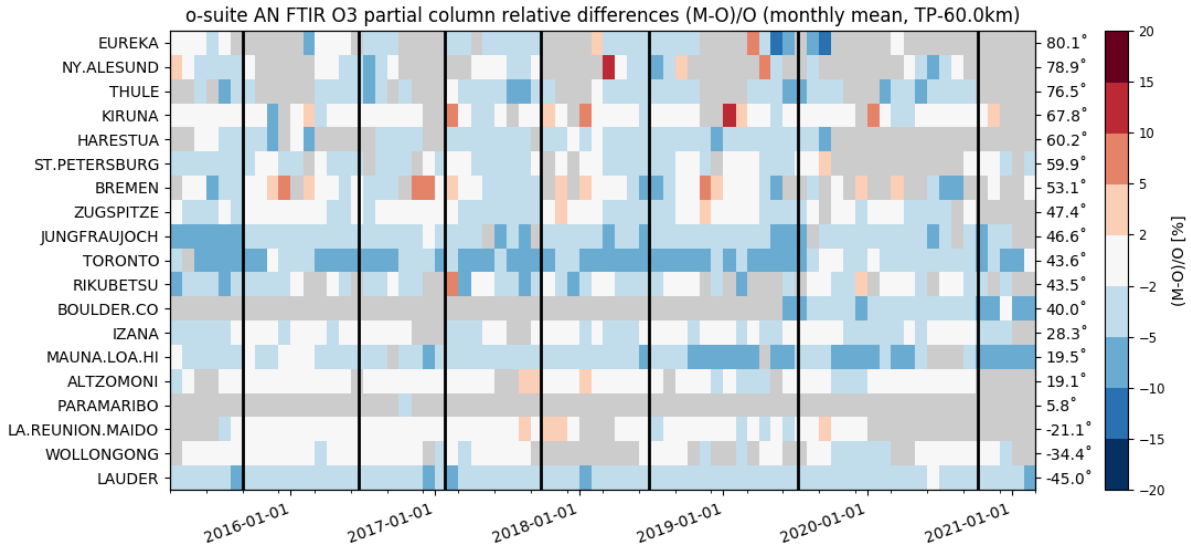
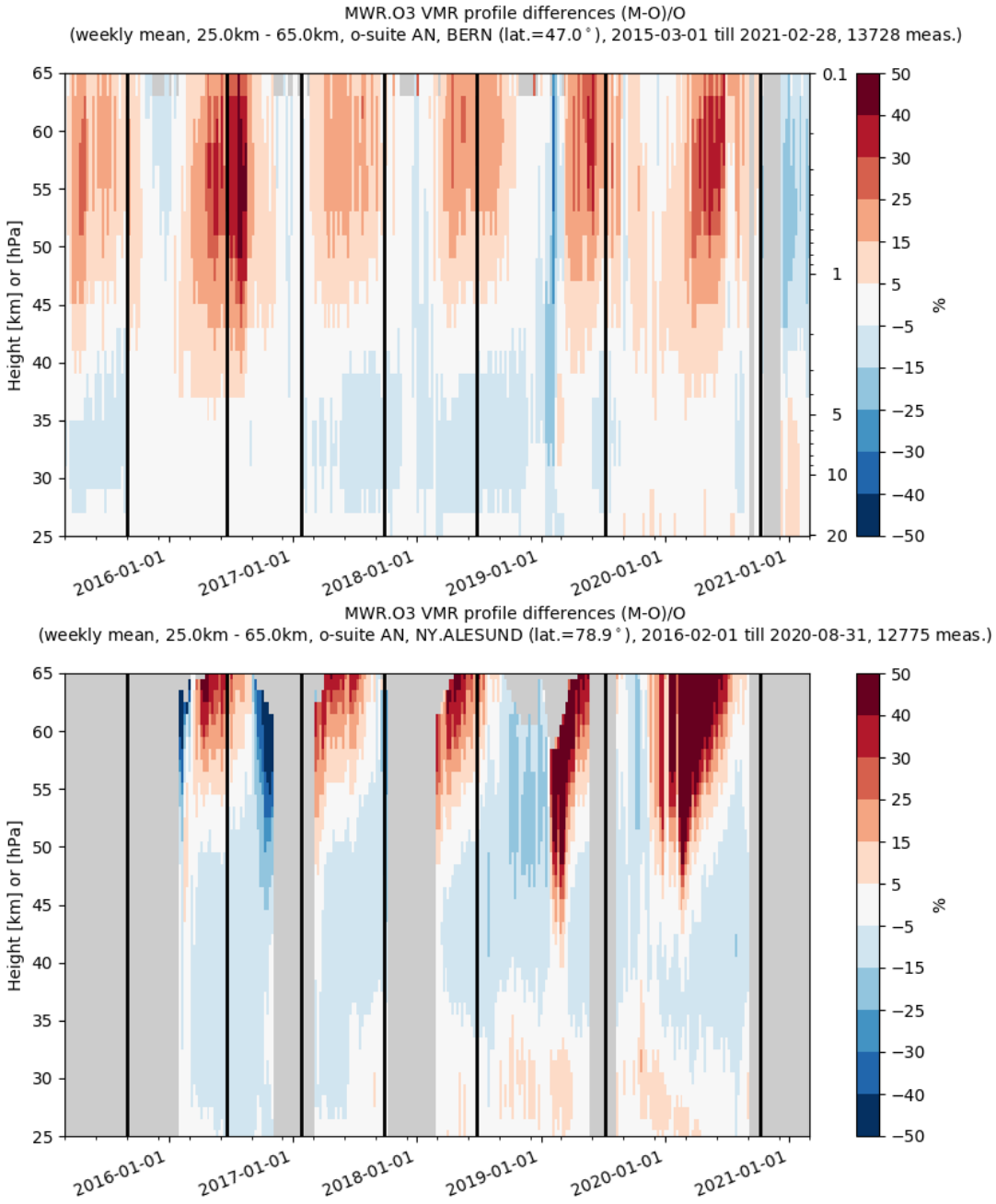


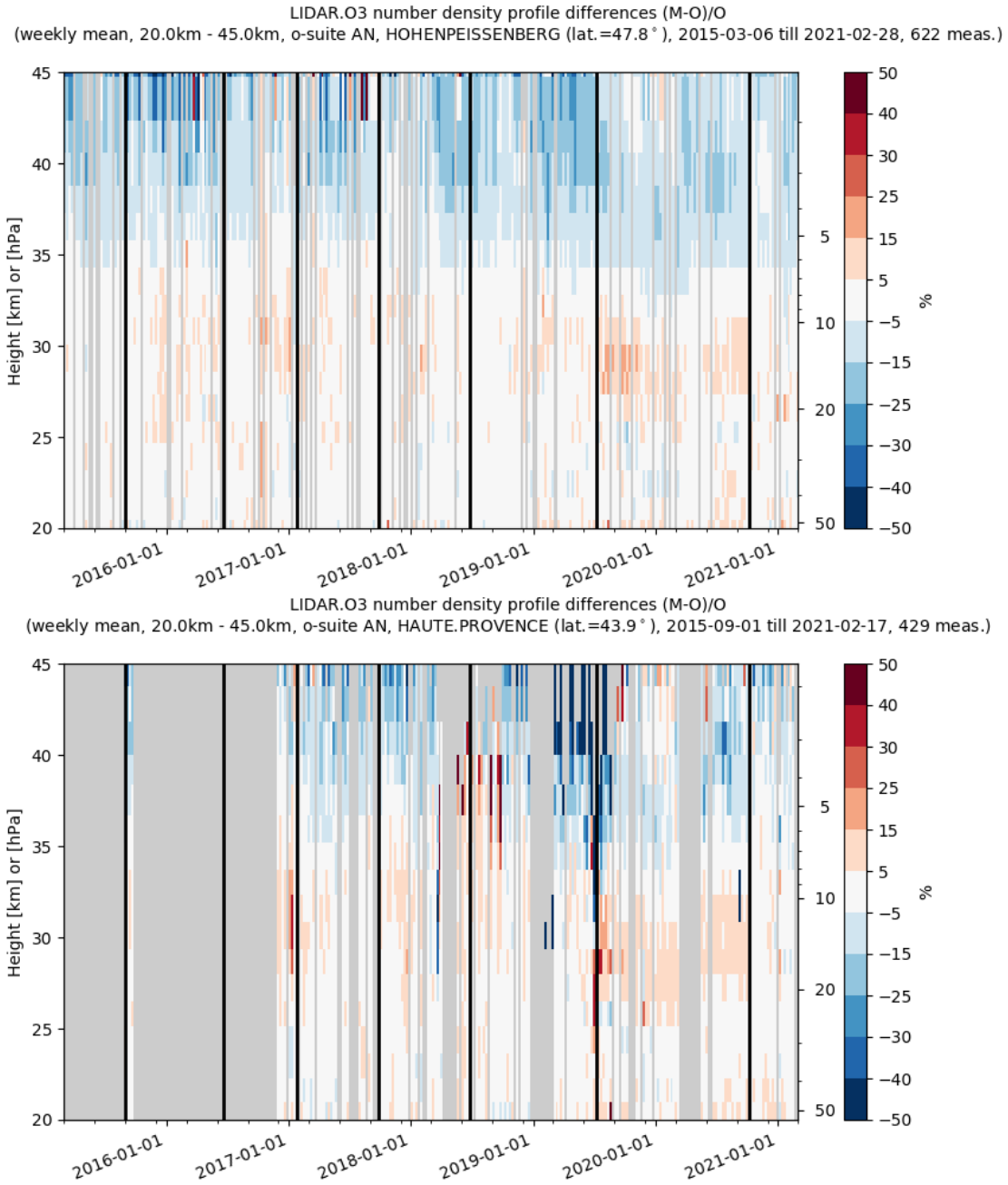
Figure 9.2.3 Time series of monthly mean relative differences for stratospheric FTIR columns along with CAMS cycle updates (black vertical lines show o-suite AN updates). Period March 2015 – February 2021.

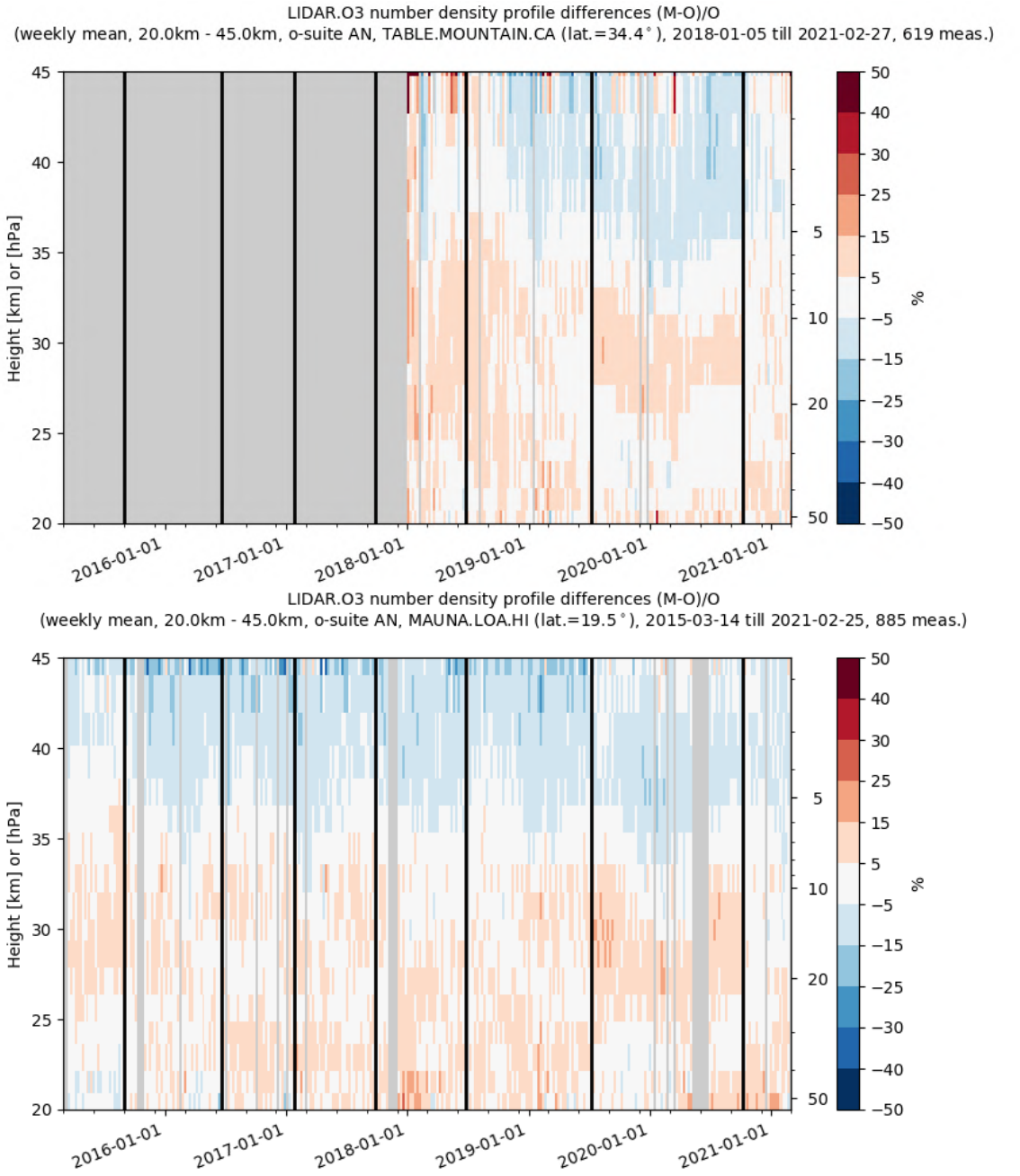
Profile comparison using LIDAR and MWR

In this section we present a comparison between the CAMS o-suite and control runs 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), Hohenpeissenberg, Germany (47°N, 11°E, altitude 1km), Table Mountain (34°N, 117.7°W, altitude 2.3km), Mauna Loa, Hawaii (19.5°N, 204°E, altitude 3.4km) and Lauder (45°S, 169.7°E),

For all LIDAR sites (see Figure 9.2.4) the o-suite slightly overestimated the observed ozone (<10%) between 25km and 35km. Since the latest osuite update, which introduced a different stratospheric ozone paramatrisation, this overestimation vanishes. 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).


MWR.O3 VMR profile differences (M-O)/O
(weekly mean, 25.0km - 65.0km, o-suite AN, BERN (lat.=47.0°), 2015-03-01 till 2021-02-28, 13728 meas.)
Height [km] or [hPa]
65
60
55
50
45
40
35
30
25
0.1
1
5
10
20
50
40
30
25
15
5
−5
−15
−25
−30
−40
−50
%
2016-01-01
2017-01-01
2018-01-01
2019-01-01
2020-01-01
2021-01-01
MWR.O3 VMR profile differences (M-O)/O
(weekly mean, 25.0km - 65.0km, o-suite AN, NY.ALESUND (lat.=78.9°), 2016-02-01 till 2020-08-31, 12775 meas.)
Height [km] or [hPa]
65
60
55
50
45
40
35
30
25
50
40
30
25
15
5
−5
−15
−25
−30
−40
−50
%
2016-01-01
2017-01-01
2018-01-01
2019-01-01
2020-01-01
2021-01-01


LIDAR.O3 number density profile differences (M-O)/O
(weekly mean, 20.0km - 45.0km, o-suite AN, HOHENPEISSENBERG (lat.=47.8°), 2015-03-06 till 2021-02-28, 622 meas.)
Height [km] or [hPa]
45
40
35
30
25
20
5
10
20
50
2016-01-01
2017-01-01
2018-01-01
2019-01-01
2020-01-01
2021-01-01
50
40
30
25
15
5
−5
−15
−25
−30
−40
−50
%
LIDAR.O3 number density profile differences (M-O)/O
(weekly mean, 20.0km - 45.0km, o-suite AN, HAUTE.PROVENCE (lat.=43.9°), 2015-09-01 till 2021-02-17, 429 meas.)
Height [km] or [hPa]
45
40
35
30
25
20
5
10
20
50
2016-01-01
2017-01-01
2018-01-01
2019-01-01
2020-01-01
2021-01-01
50
40
30
25
15
5
−5
−15
−25
−30
−40
−50
%


LIDAR.O3 number density profile differences (M-O)/O
(weekly mean, 20.0km - 45.0km, o-suite AN, TABLE.MOUNTAIN.CA (lat.=34.4°), 2018-01-05 till 2021-02-27, 619 meas.)
Height [km] or [hPa]
45
40
35
30
25
20
5
10
20
50
50
40
30
25
15
5
−5
−15
−25
−30
−40
−50
%
2016-01-01
2017-01-01
2018-01-01
2019-01-01
2020-01-01
2021-01-01
LIDAR.O3 number density profile differences (M-O)/O
(weekly mean, 20.0km - 45.0km, o-suite AN, MAUNA.LOA.HI (lat.=19.5°), 2015-03-14 till 2021-02-25, 885 meas.)
Height [km] or [hPa]
45
40
35
30
25
20
5
10
20
50
50
40
30
25
15
5
−5
−15
−25
−30
−40
−50
%
2016-01-01
2017-01-01
2018-01-01
2019-01-01
2020-01-01
2021-01-01

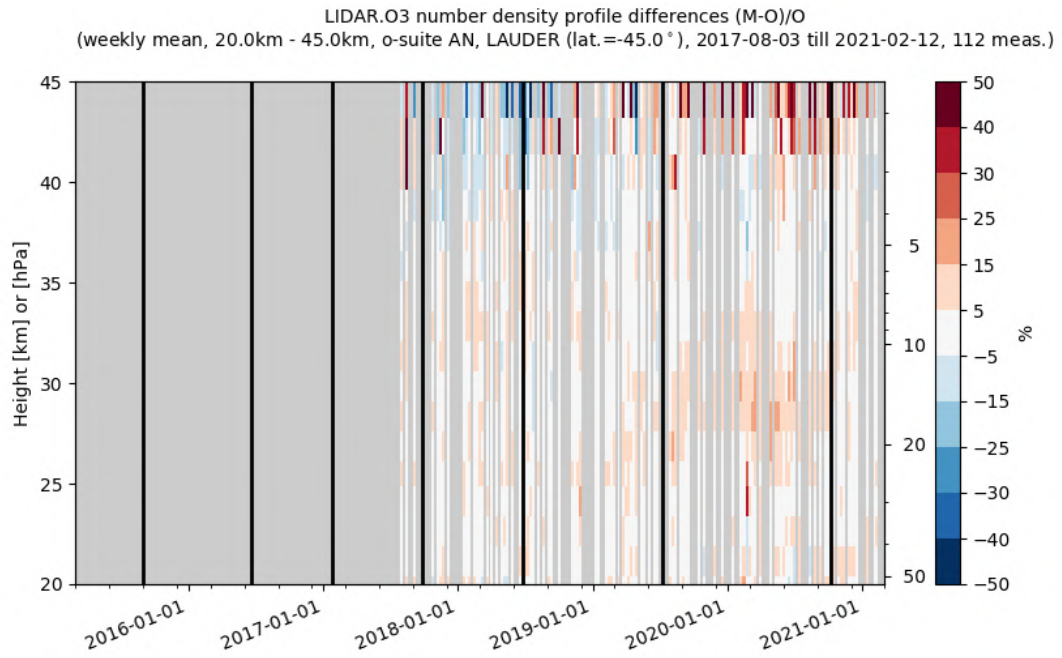


Figure 9.2.4: Comparison of the weekly mean profile bias between the O3 mixing ratios of o-suite AN and the NDACC station at Ny Alesund, Bern, Hohenpeissenberg, OHP, Mauna Loa, Table Mountain and Lauder. For the LIDAR stations, the measurement uncertainty above 35km is comparable to the observed profile bias. The latest model update changed the positive bias observed in the stratosphere (around 30km in the LIDAR comparisons). Period March 2015 – February 2021.

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

This section compares the output of the o-suite for the DJF 2020-21 period with observations by limb-sounding 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.

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.

The level-2 data from limb sounding instrument used in this section are:

- ACE-FTS version 3.6, on board SCISAT-1.
- SAGE-III version 5.1, on board the International Space Station (ISS); among the 3 different ozone profiles delivered by the solar occultation (denoted Mesospheric, MLR and AO3), we use the AO3 retrieval which is recommended by the mission science team.
- OMPS-LP version 2.5, on board NPP

For reference, we also compare these observations with BASCOE analyses which are constrained by the Aura MLS v4.2 offline profiles.

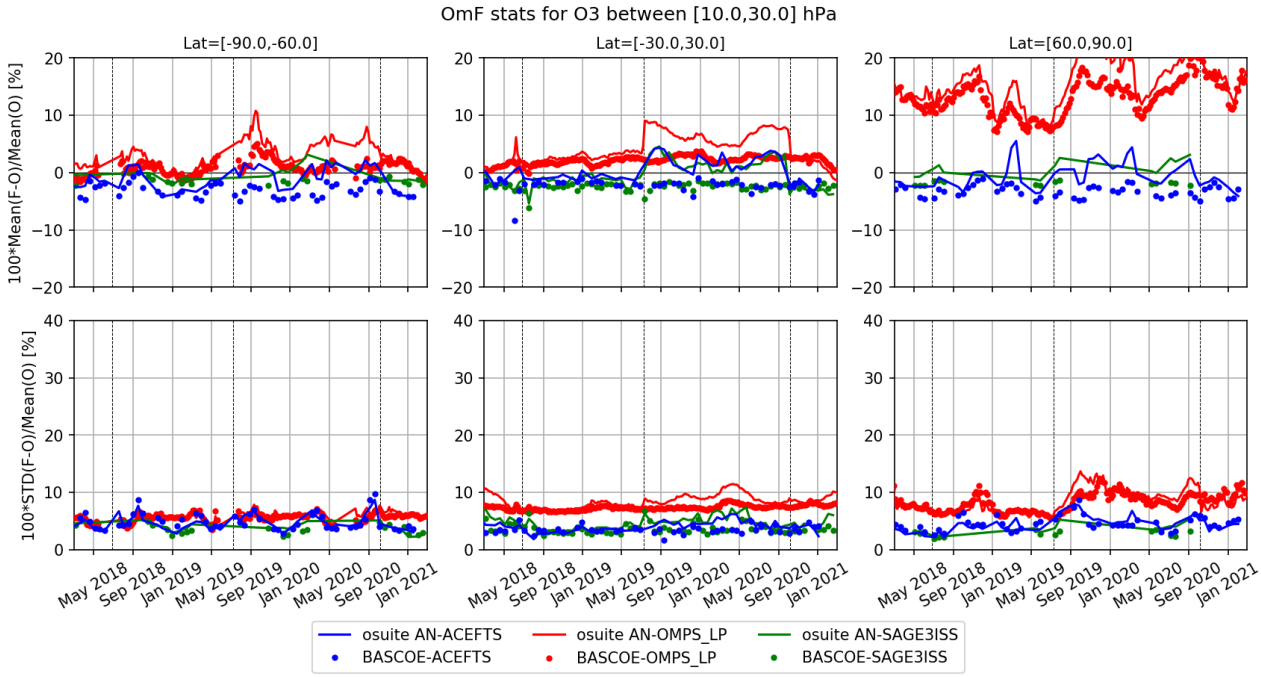


Figure 9.3.1: Time series comparing model runs to observations for the period 01-Mar-2018 to 01-Mar-2021 in the pressure range between 10 and 30 hPa: o-suite analyses (solid lines) and BASCOE (dotted lines) vs OMPS-LP (red), ACE-FTS (blue) and SAGE-III (green). Top row, normalized mean bias (model-obs)/obs (%); bottom row, standard deviation of relative differences (%). Vertical dashed lines indicate the date of CAMS model updates: CY45r1 (26 Jun 2018), CY46r1 (9 Jul 2019) and CY47r1 (6 Oct 2020).

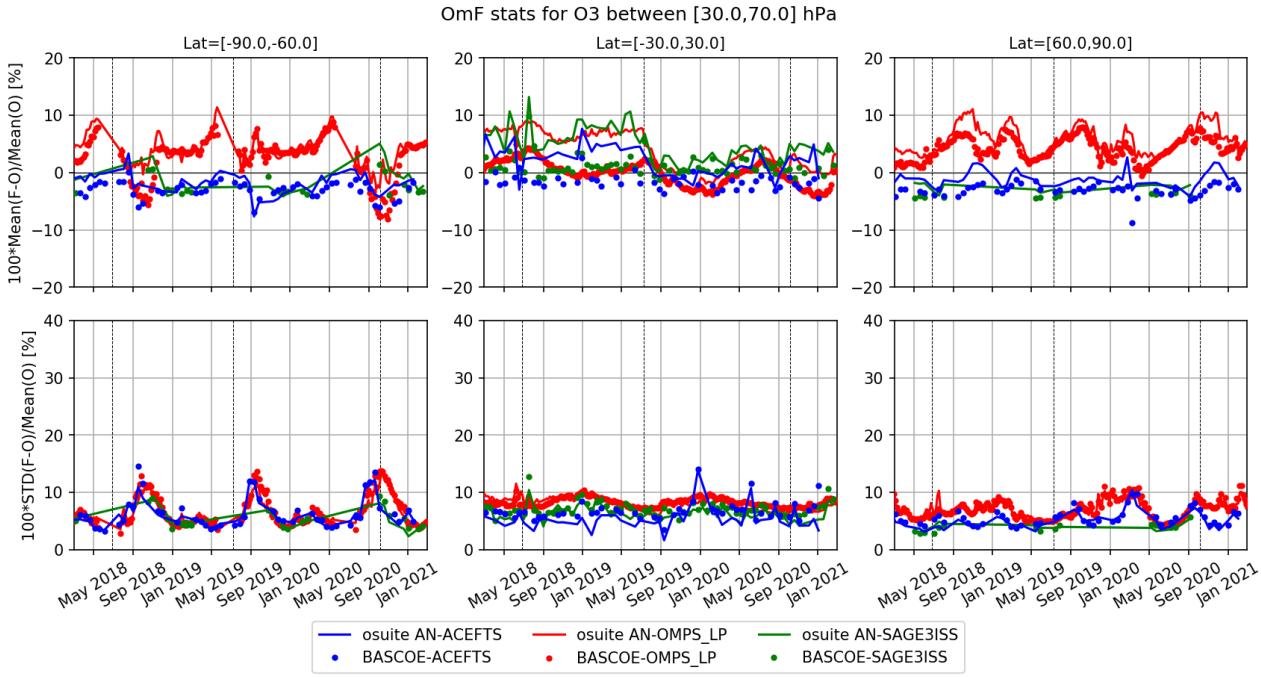


Figure 9.3.2: As Fig. 9.3.1 but in the pressure range between 30 and 70hPa.

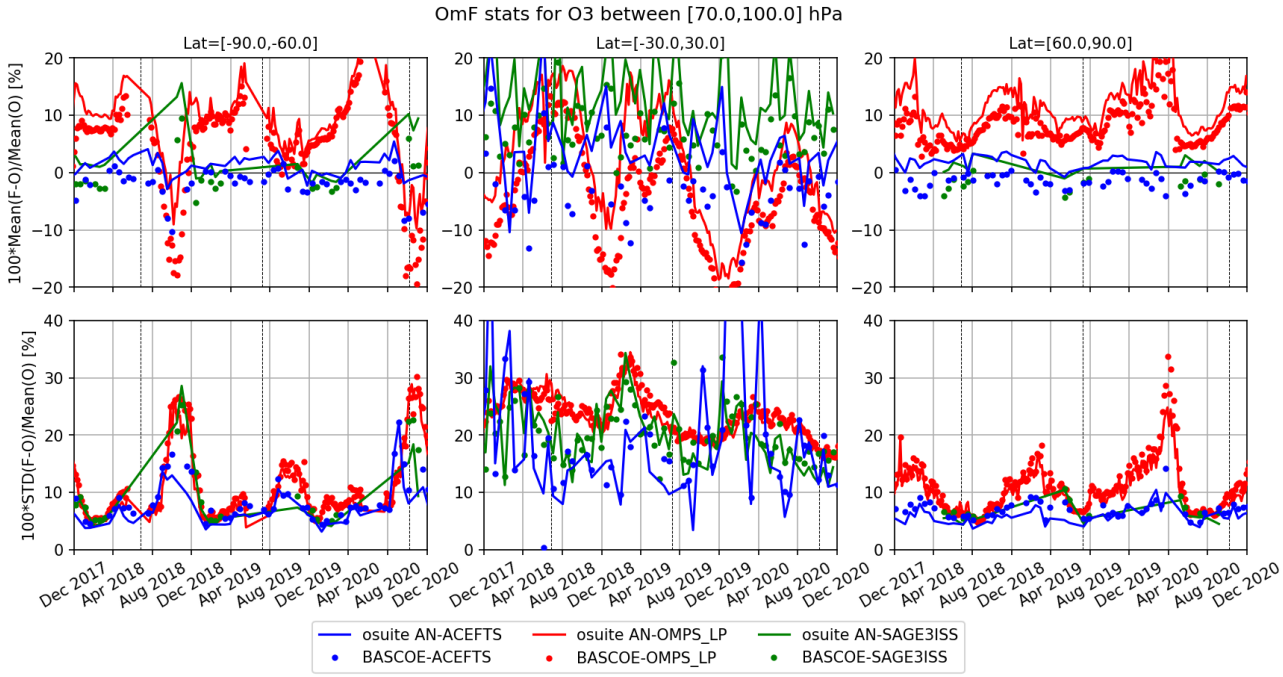


Figure 9.3.3: As Fig. 9.3.1 but in the pressure range between 70 and 100 hPa.

Figure 9.3.1 to 9.3.3 present, in the upper row, the time series over the last 36 months of the bias of the o-suite against the three satellite measurements for respectively three layers of the stratosphere (10-30hPa, 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.

Against ACE-FTS, the agreement of the o-suite and BASCOE is good, the bias is generally within ±5% except in the tropical lower stratosphere (TLS, between 70 and 100 hPa) where the bias is around ±10%.

The SAGE-III onboard ISS provide observations since June 2017. The latitudinal coverage is more limited than ACE-FTS; the polar regions are only covered during the polar summer. Where available, the agreement of the o-suite with SAGE-III is good, with biases similar to those observed against ACE-FTS, except in the TLS where they are more positive (3-13%).

Compared to OMPS-LP, there is an almost systematic overestimation by the o-suite; the biases are more variable and more marked than for the other instruments (10% to 15% in the north polar at 10-30hPa region, up to 10% at 30-70hPa and up to 20% at 70-100hPa).

However, the recent implementation of CY47r1 in early October 2020 show an overall improvement of the o-suite, in particular in the tropics between 10-30 hPa. Also, the o-suite ozone bias against these independent observations in this region is now similar to the one of BASCOE.

Figure 9.3.4 to 9.3.7 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., DJF 2020-21 unless for ACE-FTS where data for March 2021 were not yet available at the time of the writing.

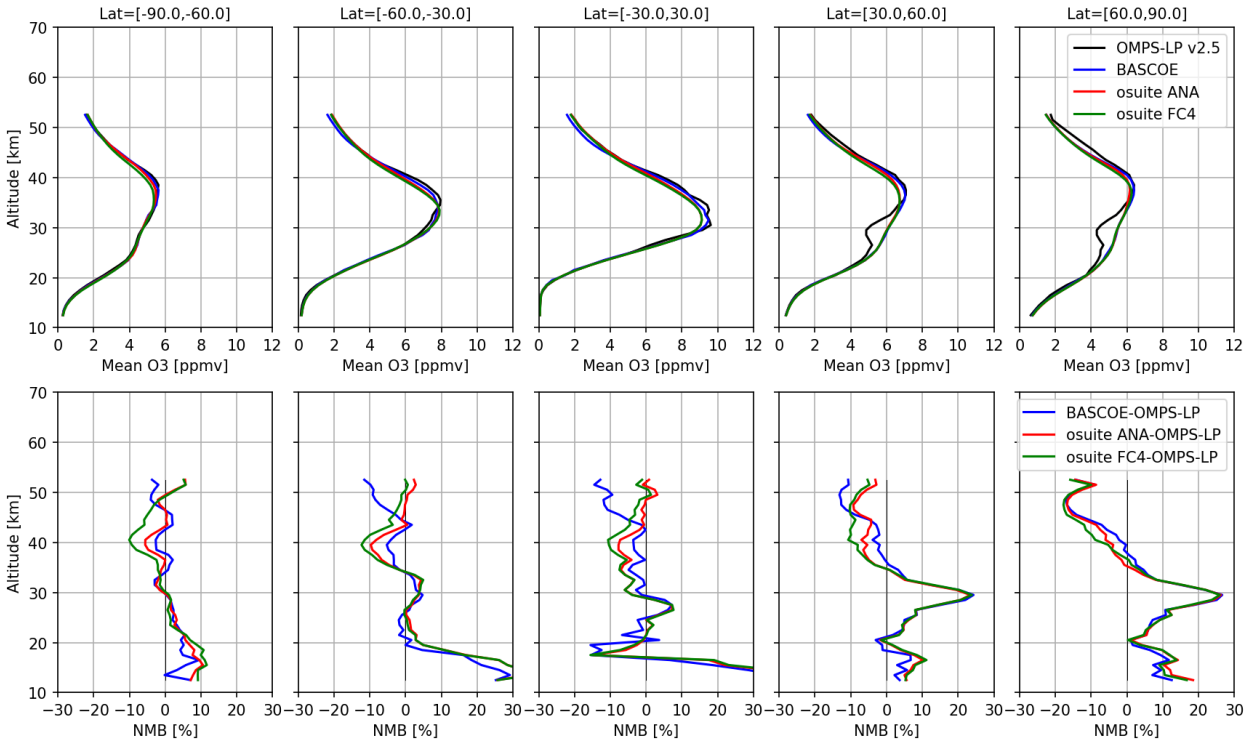


Figure 9.3.4: Mean value (top) and normalized mean bias (bottom) of the ozone profile between o-suite analyses (red), o-suite forecasts 4th day (green) and BASCOE (blue) with OMPS-LP v2.5 observations for the period DJF 2020-21.

All o-suite profiles present a common feature of a slight overestimation at around 30km, followed by a stronger underestimation at around 40km, which is evidenced in the 4th day forecast.

The profiles of OMPS-LP in the northern hemisphere present irregularities (mainly in the part contributed by the sensor in the visible), which are not found in the other instruments nor in the o-suite or the BASCOE models; hence they should be disregarded for this validation.

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-LP and Aura MLS provide daily data with more than 40000 valid profiles (while OMPS-LP being blind in the polar night), while ACE-FTS provides around 700 profiles in the polar region and 200 profiles in the tropics, and SAGE-III around 800 profiles in each latitude band except the south polar region (none).

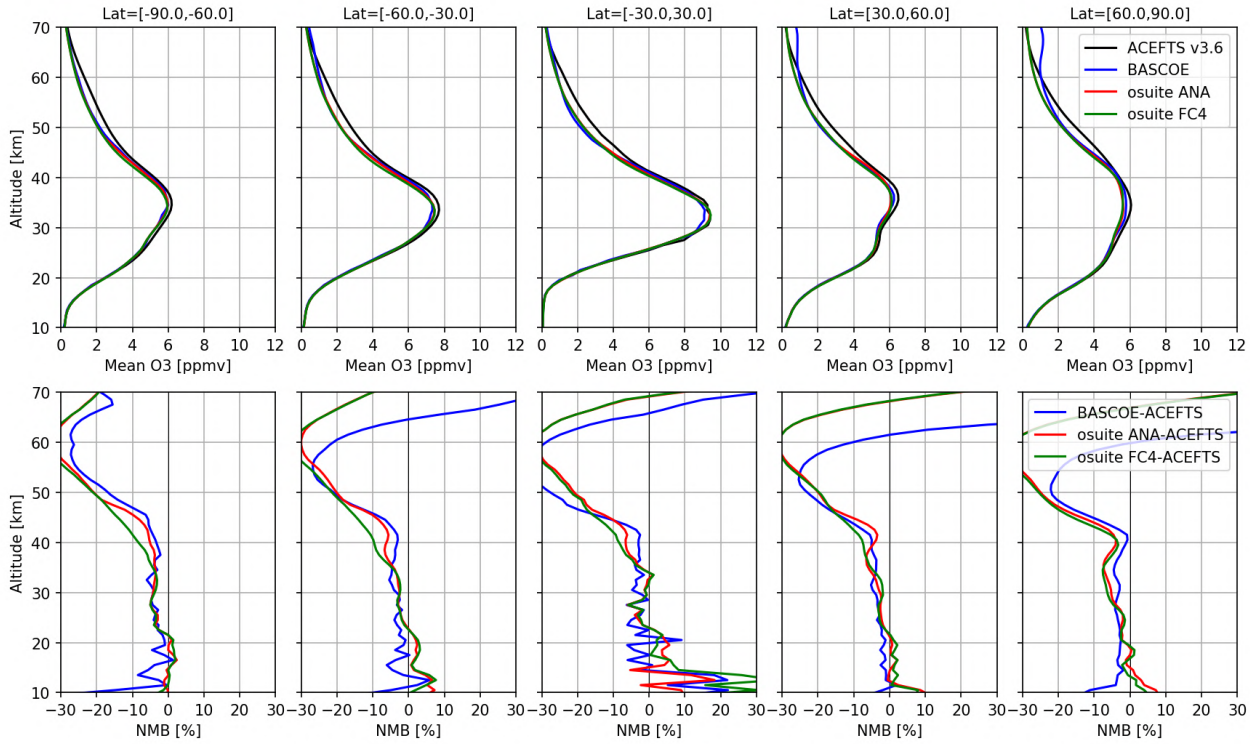


Figure 9.3.5: As Fig. 9.3.4 but for comparison against ACE-FTS observations (here for the period Dec 2020-Jan 2021 since Feb 2021 where not yet available at the time of writing).

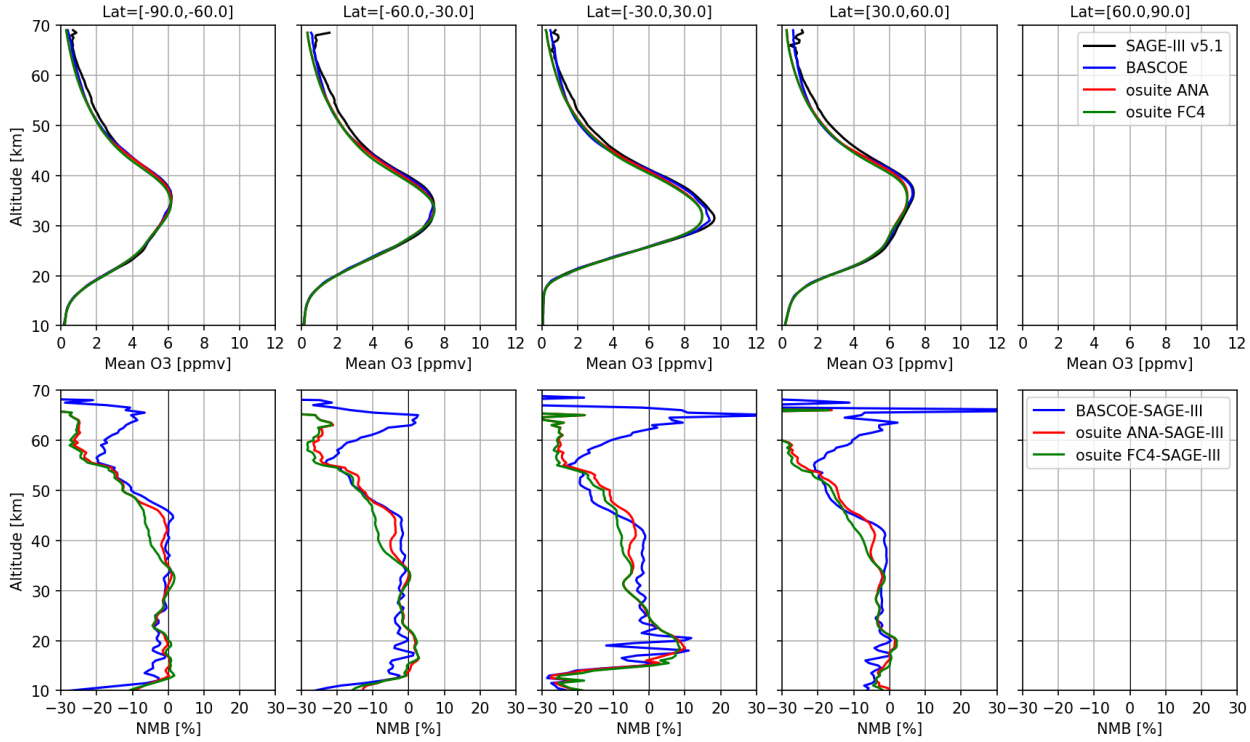


Figure 9.3.6: As Fig. 9.3.4 but for comparison against SAGE-III observations.

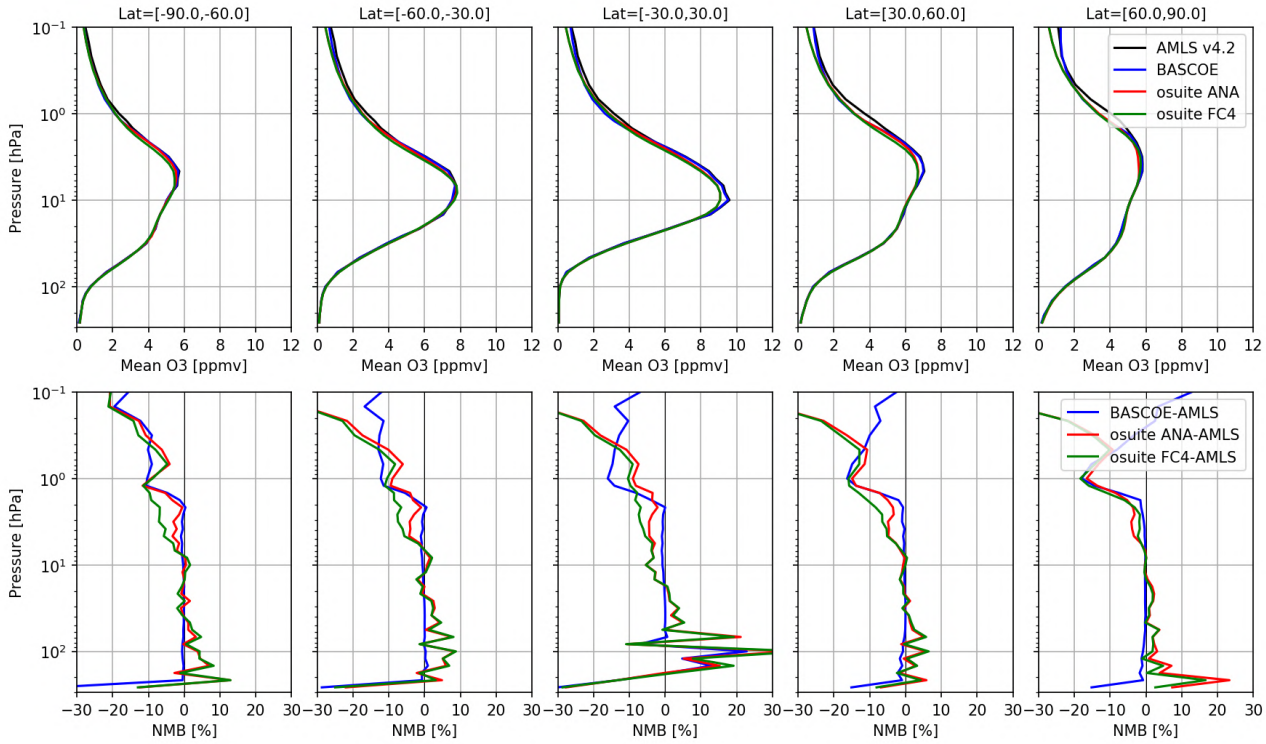


Figure 9.3.7: As Fig 9.3.4 but for comparison against MLS offline observations.

9.4 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 NO_2, 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 compared to modelled stratospheric NO_2 columns (note that the comparisons against GOME-2A have ended in Nov 2020, comparisons against TROPOMI and GOME-2C are not available yet and will be implemented with the next reports). 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.

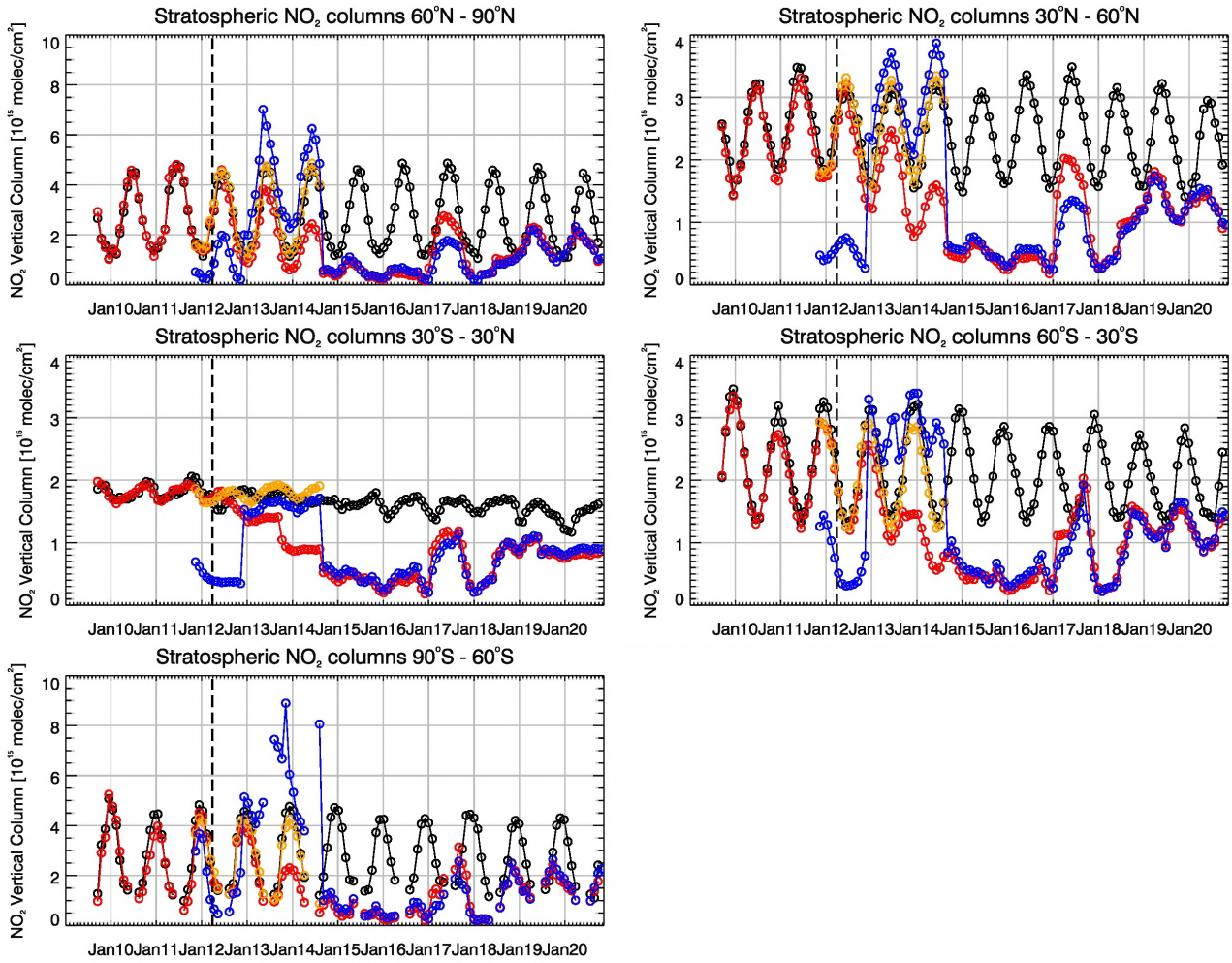


Figure 9.4.1: Time series of average stratospheric NO2 columns [1015 molec cm-2] from SCIAMACHY (up to March 2012) and GOME-2 (from April 2012, black) compared to model results (red: o-suite, blue: MACC fcnrt TM5/MACC CIFS TM5/control, orange: MACC fcnrt MOZ) for different latitude bands (time perido: Sep 2009 – Nov 2020). 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 9.4.1, 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.

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.

10. 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 March 2019 to 1st March 2021 based on observations from 26 surface stations, located in Western Europe: 10 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 experiments (*h9sp, Tco1279L137; 9x9 km*), coupled to the analysis experiment (*h72g, Tco399L137, 25x25 km*). The new experiments (*he9h* for the forecast; *hd7v* for the analysis) have been used since 1st November 2020.

10.1 CH_4 and CO_2 validation against ICOS observations

The CO_2 and CH_4 simulations from the analysis and high-resolution forecast have been compared to 26 ICOS stations. The near–real time data processing of the in-situ measurements is ensured by the Atmospheric Thematic Center (Hazan et al., 2016). All stations follow the ICOS Atmospheric Station specification (Laurent et al., 2017), which is a requirement in the labelling process (Yver Kwok et al., 2020). Among the 26 stations we can distinguish four sites located on top of mountains (PUY, JFJ, CMN, PRS), two background sites (PAL, ZEP), two coastal sites (UTO, MHD) and 17 tall towers. In addition, there is one site in South Hemisphere at La Réunion Island (RUN). For the tall towers we consider only in this report the highest sampling levels which are at least at 100m above the ground.

The figure 10.1.1 shows the time varying biases (CAMS runs minus observations), averaged on a weekly basis, for ICOS stations. The CO_2 biases are characterized by a clear seasonal cycle at most sites with maximum values in Spring/Summer, and minimum in Autumn/Winter. Ispra (IPR), a tall tower located in the Po valley, appears as an outlier probably due to the complex orography in the surroundings. We also observed higher weekly biases at Lutjewad (LUT), Karlsruhe (KIT) and Saclay (SAC) due to the vicinity of emission hot spots or urban areas. Four examples are detailed on Figures 10.1.2 for Monte Cimone (CMN, Italy) and Puy de dôme (PUY, France), and on Figure 10.1.3 for Norunda (NOR, Sweden) and Svarberget (SVB, Sweden). The two mountain sites show quite similar features: positive biases of 1 to 2% in Spring-Summer, but for the rest of the year the mean biases are close to 0. The main difference can be found in the timing of the maximum bias, which occurs earlier (May) in PUY compared to CMN (July). A major synoptic event at PUY in January 2021 is represented with a very good timing in both CAMS runs, with an amplitude underestimated by 1%. The two Nordic site (Figure 10.1.3) of Norunda and Svarberget, also show similar patterns. For the two sites the biases remain close to zero till spring. In June the CO_2 drawdown observed at both sites in underestimated by the CAMS experiments, similarly to PUY. Then the highest biases are observed in autumn, when the CAMS runs cannot reproduce the CO_2 enhancements observed at Nordic sites. Since the beginning of the new experiment starting in November 2020, the biases are close to zero at both sites.

The seasonal cycles of the CO_2 biases calculated for all ICOS sites are shown on figure 10.1.6. Except for Ispra (IPR), and Karlsruhe (KIT) in Autumn, the model/observations differences generally range

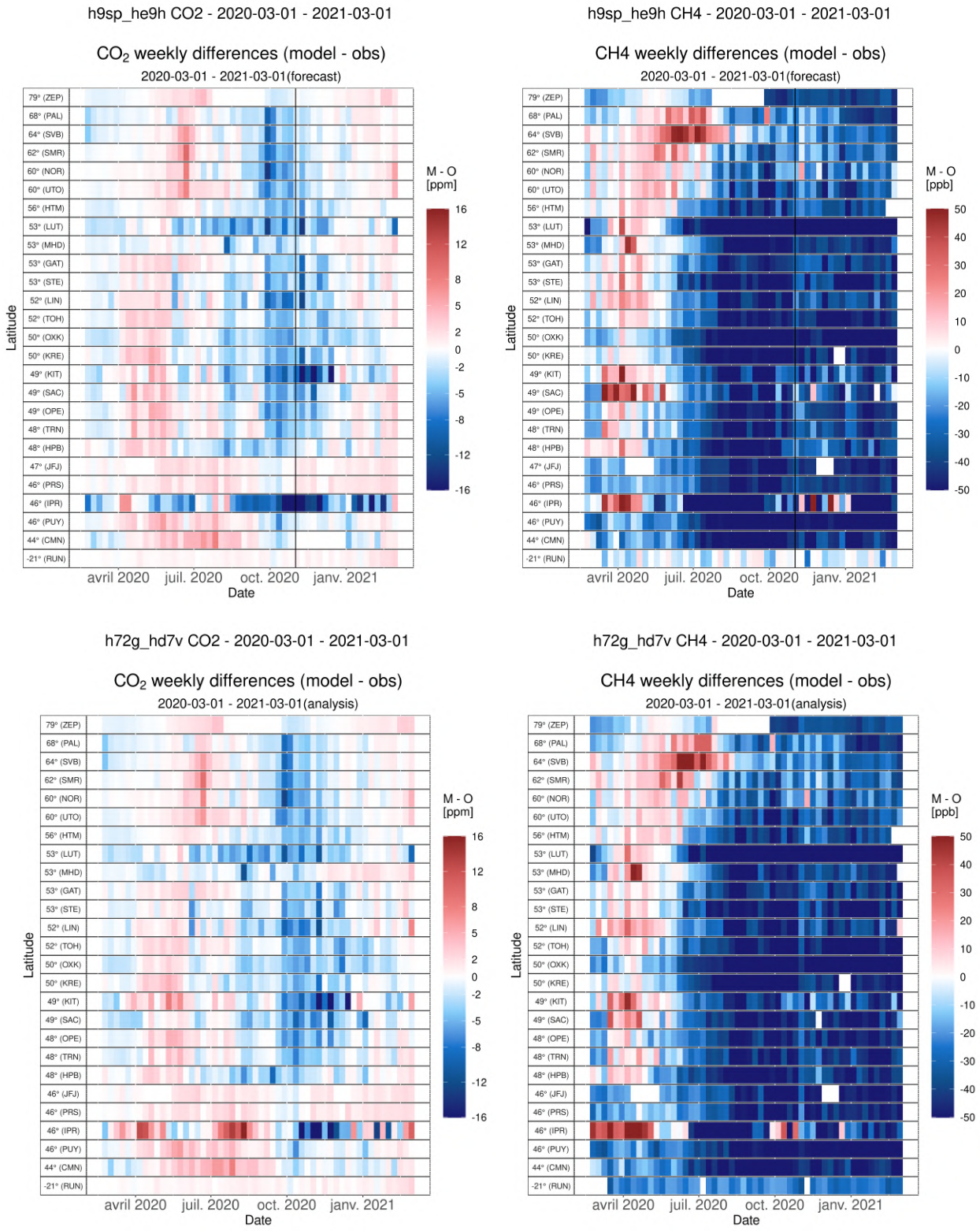


Figure 10.1.1: Mosaic plot of CO_2 (left, in ppm) and CH_4 (right, in ppb) biases of the CAMS high resolution forecast (top panel) and analysis (bottom panel), compared to surface station observations. Each vertical colored line represents a weekly mean.

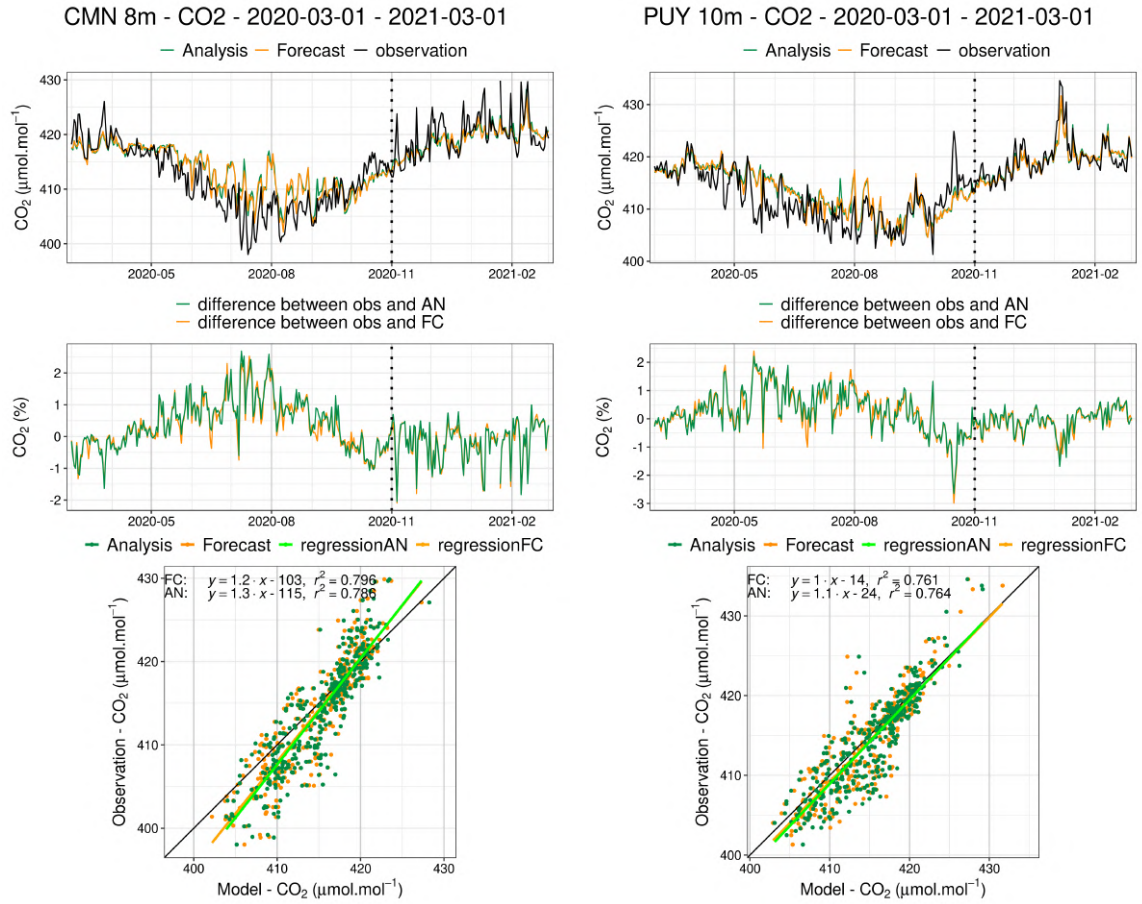


Figure 10.1.2: Comparison of CO_2 daily means observed (black) with the analysis run (green) and the high-resolution forecast (orange) at Monte Cimone (left) and Puy de dôme (right). Middle: differences of the observations minus the simulations. Below: Linear fit between observations and simulations. The dashed vertical line represents the change of experiments in November 2020.

within ±5 ppm, with a maximum value occurring between May and July. Higher negative biases are observed in the peri-urban sits of KIT and SAC. The lowest biases are observed on average at Mace Head, La Réunion Island, and the mountain sites (JFJ, PRS).

For CH_4 the biases show seasonal and latitudinal patterns, quite similar to the ones observed for CO_2 (Figure 2.1.1). At the Scandinavian sites the CAMS runs overestimates (up to 50 ppb) the observations from spring to summer. The example of Svarberget station (Figure 10.1.4) clearly shows the overestimation of the baseline and the amplitude of the CH_4 spikes at this season by the CAMS runs, which could indicate that the wetland emissions are overestimated. Then starting in August, the CAMS experiments underestimate the CH_4 concentrations down to -2%. The more we go at lower latitudes, the more we observe negative biases. In Germany and Northern France, the bias is positive only during a short period in spring. At Ochsenkopf (Figure 2.1.5) the bias is close to zero in May but remains negative all the year by -2% before spring and -3% afterward. The CAMS runs at Zeppelin site (figure 2.1.5) show a similar comparison to observation. At this site it is worth to note a step increase of the CH_4 concentrations by about 30 ppb in early July. This CH_4 enhancement corresponds to a change in the air masses origins, from the North Sea to Siberia on

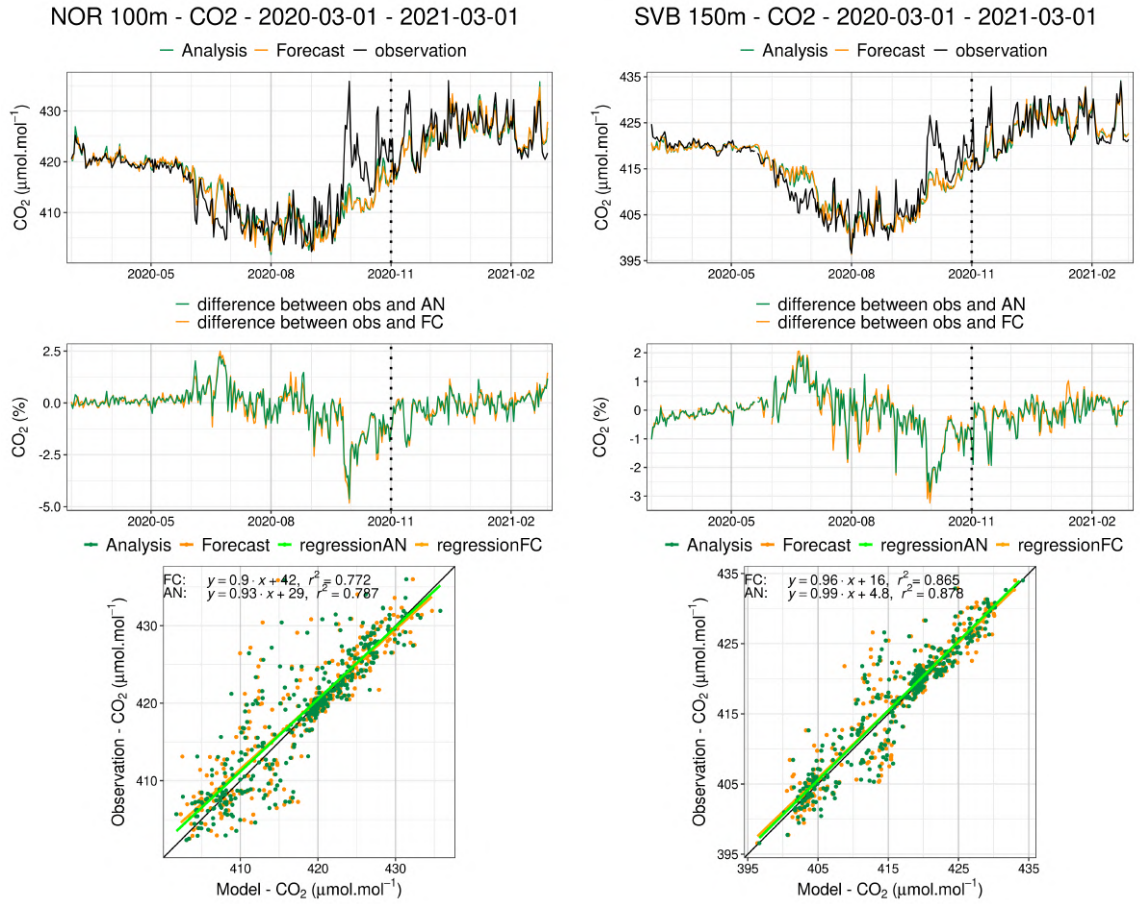


Figure 10.1.3: Same as Figure 10.1.2 for Norunda (NOR) and Svarberget (SVB).

July 2nd. The CAMS runs represent this change, associated to the CH_4 emissions in Siberia, with the good timing and amplitude. At the southern hemisphere site in La Reunion, we observe a more systematic bias over the year. There we can note that the bias of the forecast run is smaller by about 1%, but the correlation is improved in the analysis thanks to a decrease of the amplitude of synoptic events. The seasonal cycle of the CH_4 biases (Figure 10.1.6) generally present a maximum value in March and a minimum in September.

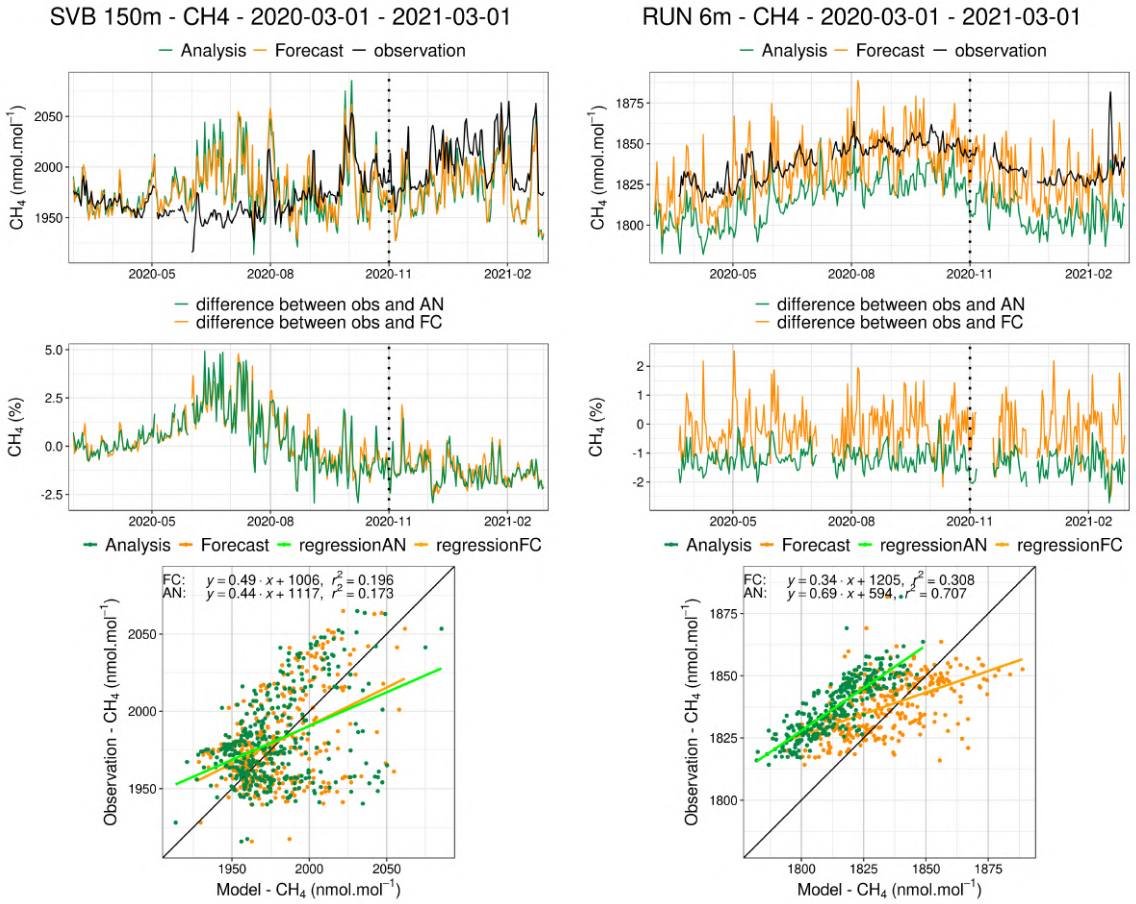


Figure 10.1.4: Comparison of CH4 daily means observed (black) with the analysis run (green) and the high-resolution forecast (orange) at Svarberget (left) and La Réunion (right). Middle: differences of the observations minus the simulations. Below: Linear fit between observations and simulations. The dashed vertical line represents the change of experiments in November 2020.

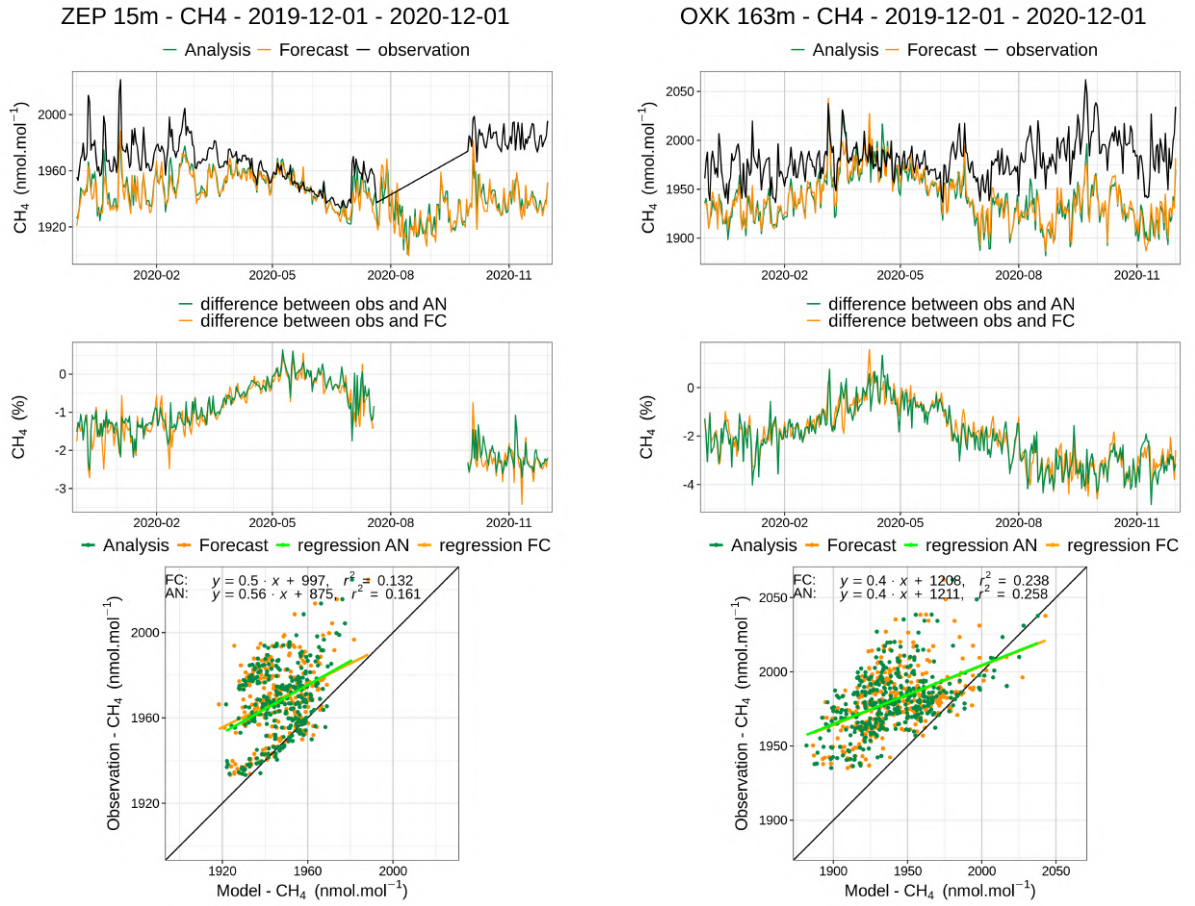


Figure 10.1.5: Same as Figure 10.1.4 for Zeppelin (left) and Ochsenkopf (right).

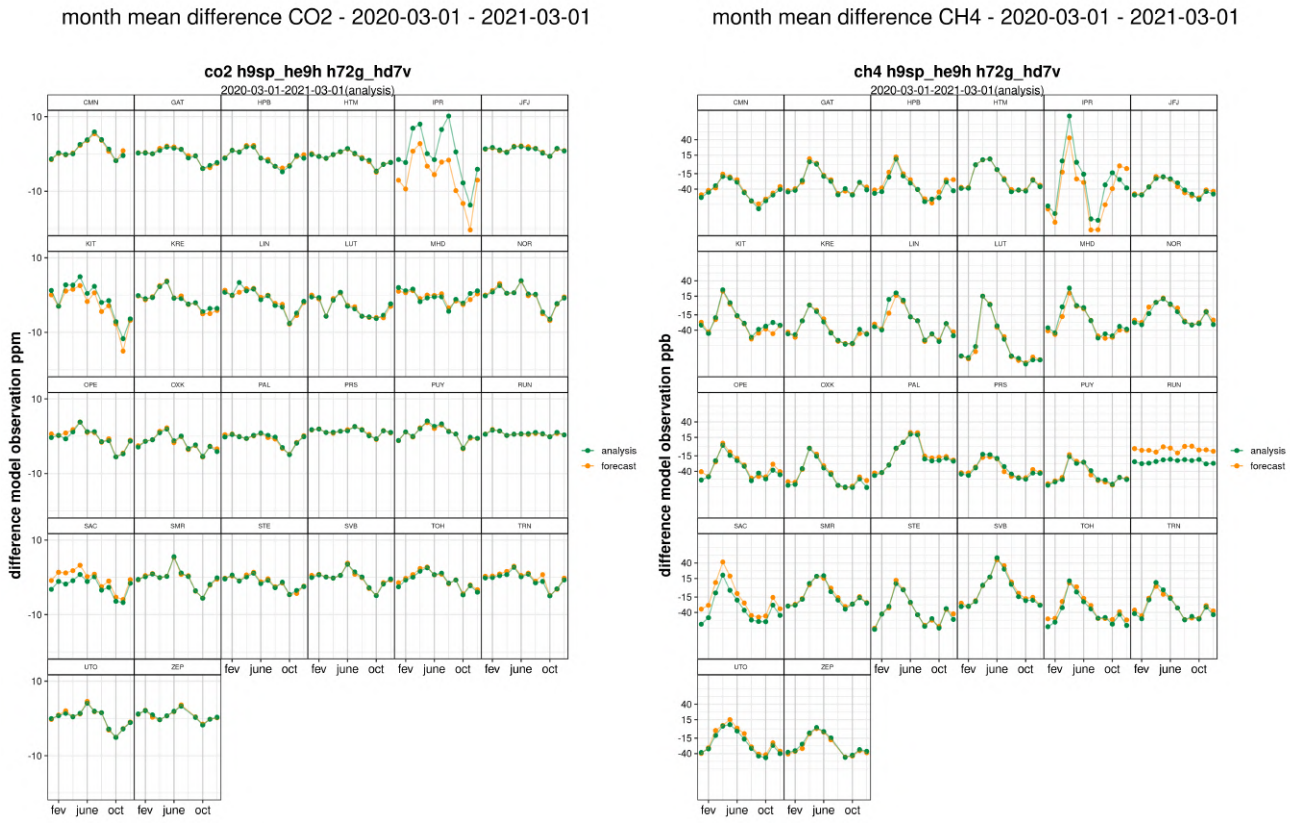


Figure 10.1.6: Mean seasonal cycles of the biases for CO_2 (left, in ppm) and CH_4 (right, in ppb) at ICOS stations. The forecast experiment is shown in orange, and the analysis in green.

10.2 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. The validation routines used for TCCON data are the same as used for the NDACC network and are documented in Langerock et al. (2015). In this section, we compare column averaged mole fractions of CH_4 and CO_2 of the CAMS models with TCCON retrievals. Data from the following TCCON sites has been used:

Izana (Blumenstock et al., 2017), Reunion (De Mazière et al., 2017), Bialystok (Deutscher et al., 2017), Manaus (Dubey et al., 2017), Four Corners (Dubey et al., 2017), Ascension (Feist et al., 2017), Anmeyondo (Goo et al., 2017), Darwin (Griffith et al., 2017), Wollongong (Griffith et al., 2017), Karlsruhe (Hase et al., 2017), Edwards (Iraci et al., 2017), Indianapolis (Iraci et al., 2017), Saga (Kawakami et al., 2017), Sodankyla (Kivi et al., 2017), Hefei (Liu et al., 2018), Tsukuba (Morino et al., 2017), Burgos (Morino et al., 2018), Rikubetsu (Morino et al., 2017), Bremen (Notholt et al., 2017), Spitsbergen (Notholt et al., 2017), Lauder (Sherlock et al., 2017, Pollard et al., 2019), Eureka (Strong et al., 2018), Garmisch (Sussmann et al., 2017), Zugspitze (Sussmann et al., 2018), Paris (Te et al., 2017), Orleans (Warneke et al., 2017), Park Falls (Wennberg et al., 2017), Caltech (Wennberg et al., 2017), Lamont (Wennberg et al., 2017), Jet Propulsion Laboratory (Wennberg et al., 2017), East Trout Lake (Wunch et al., 2017), Nicosia (Petri et al., 2020)

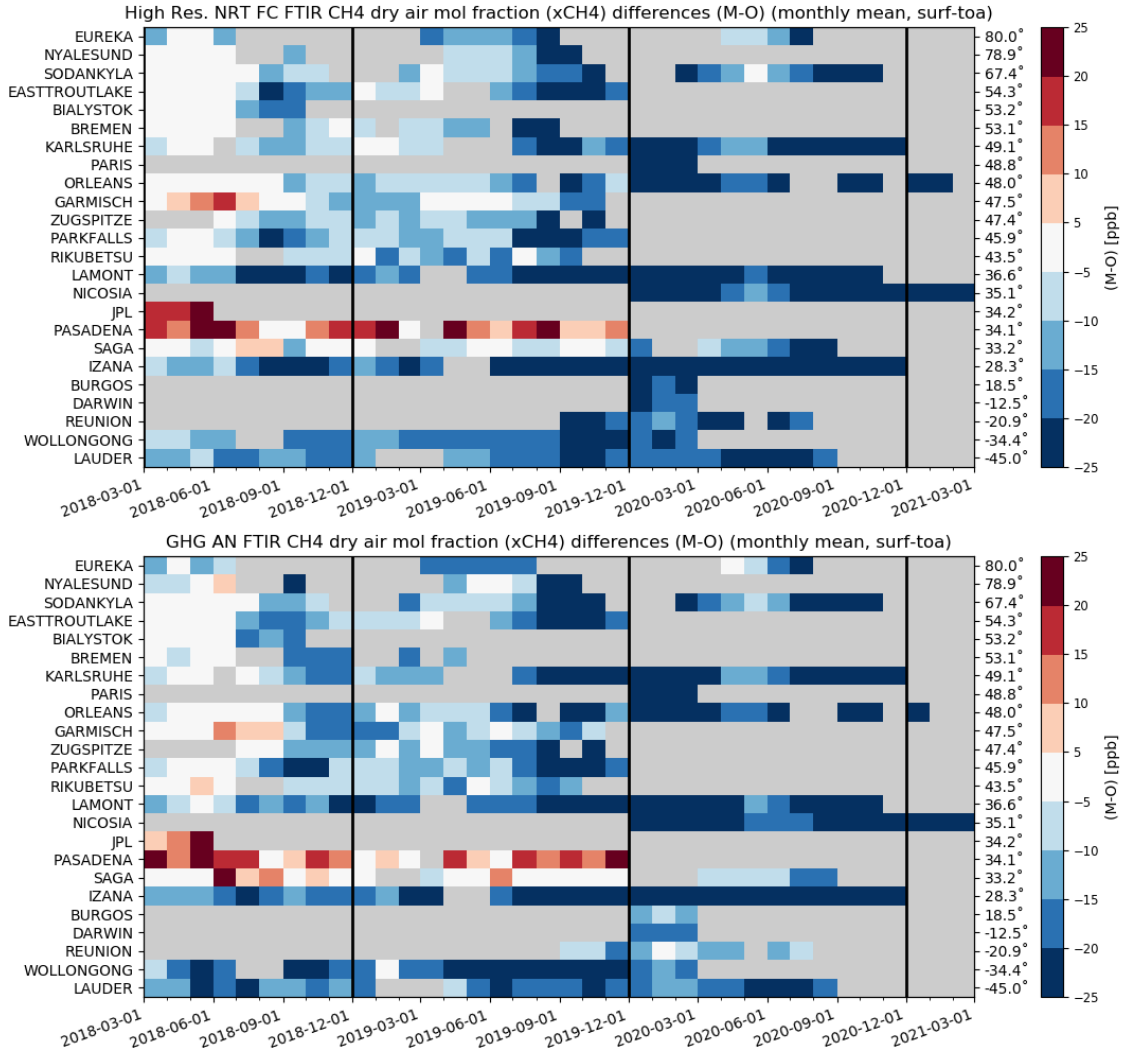


Figure 10.2.1: Monthly differences for the last 2.5 years (upper plot: high res NRT, lower plot: GHG AN). The stations are sorted by latitude (northern to southern hemisphere).

For the validation of the models in December, January and February TCCON data was available only from the sites Orleans and Nicosia. However, there are only a few data points from Orleans. The reason is that problems with the TCCON instrument at Orleans occurred and it was difficult to resolve these problems during the lockdown. Currently the site at Orleans is not operational. The data from the stations Orleans and Nicosia used for the comparison period is the rapid delivery data, which has not undergone all TCCON quality checks.

Methane (CH_4)

Figure 10.2.1 shows the data for the last 2.5 years. The comparison is shown for the sites Orleans (Fig. 10.2.2) and Nicosia (Fig. 10.2.3). The data from these stations show that the model data underestimates the CH_4 for these stations by up to 50 ppb.

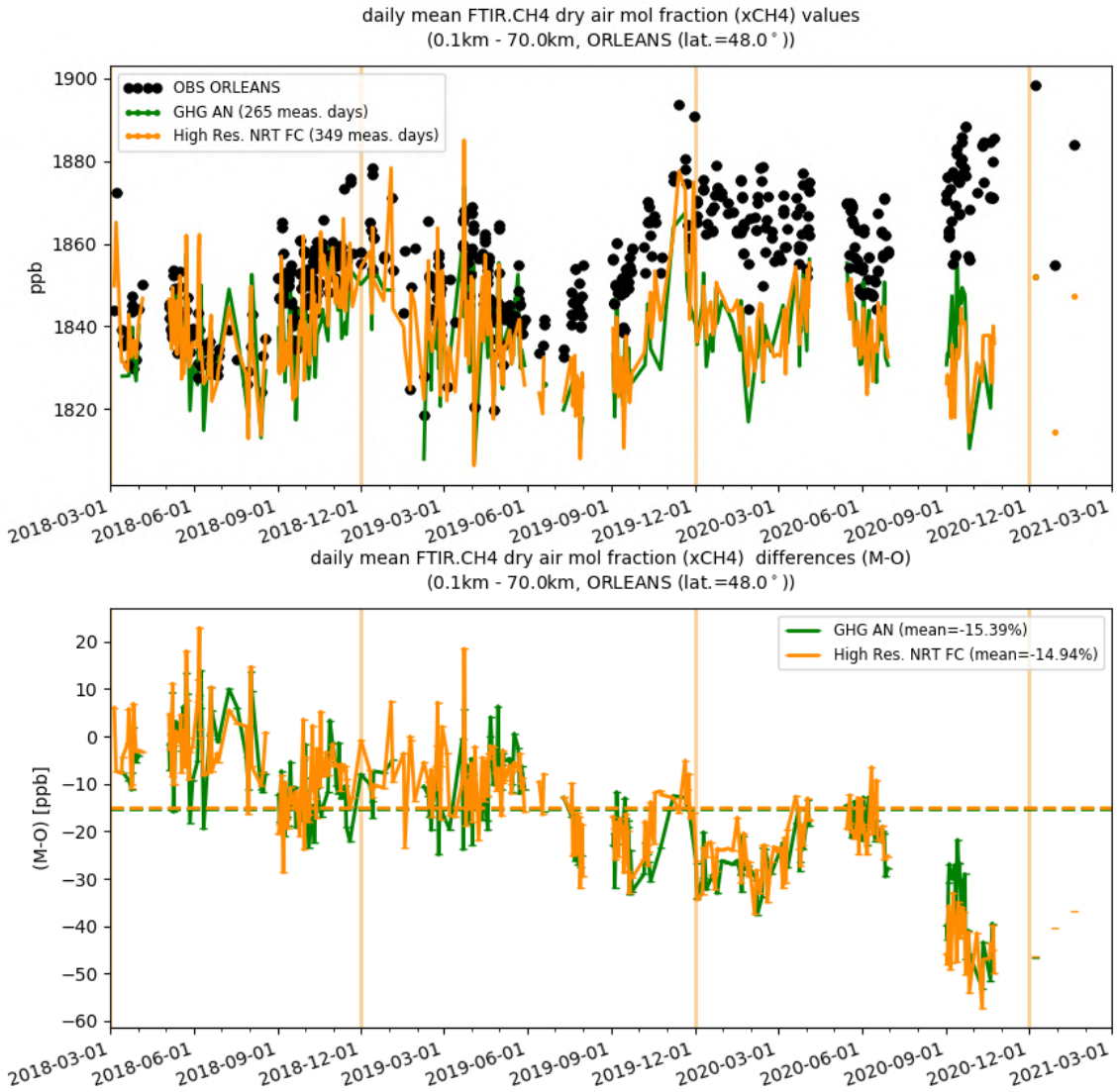


Figure 10.2.2: Comparison of the CH_4 model data with TCCON CH_4 at Orleans.

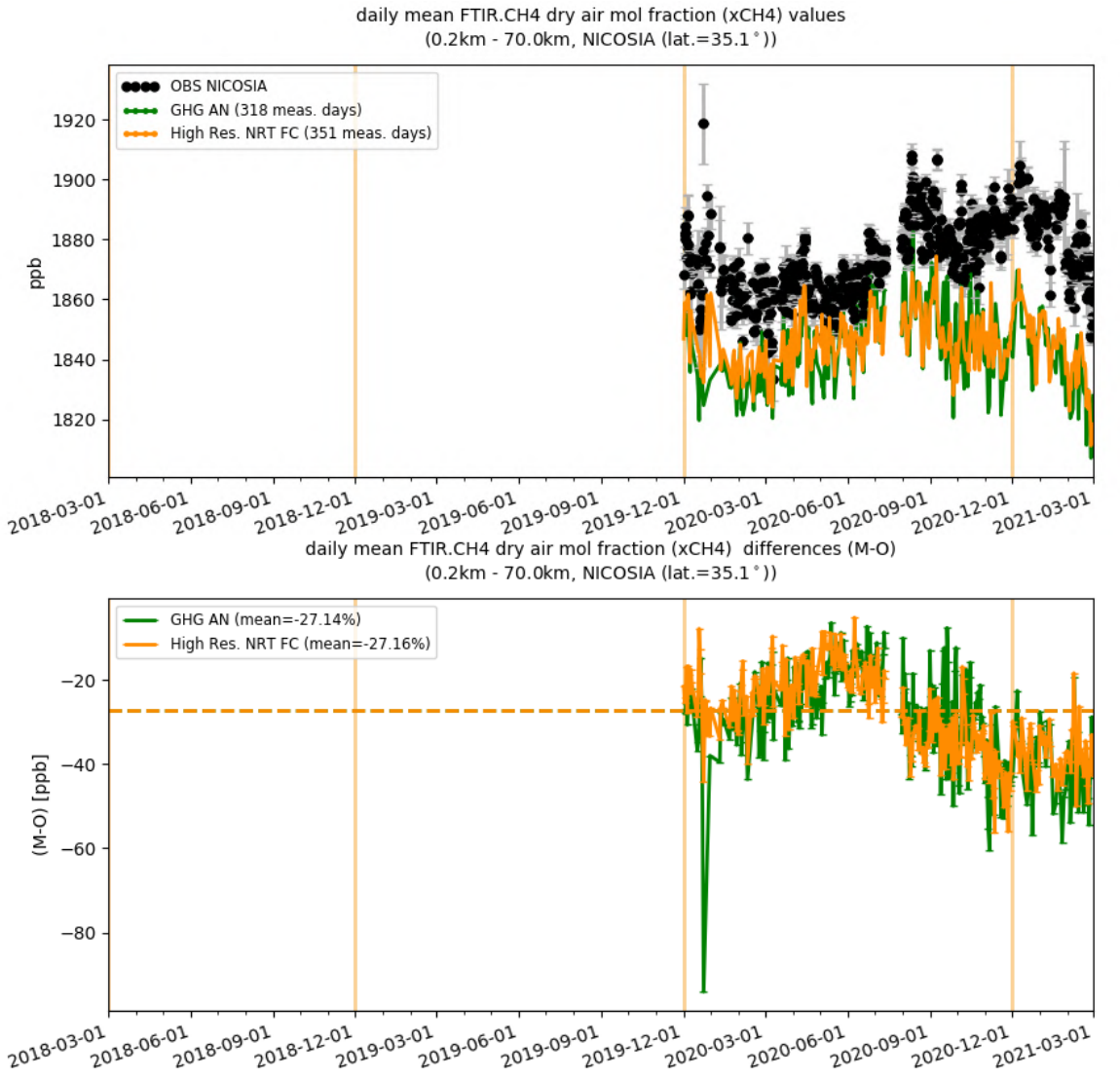


Figure 10.2.3: Comparison of the CH_4 model data with TCCON CH_4 at Nicosia.

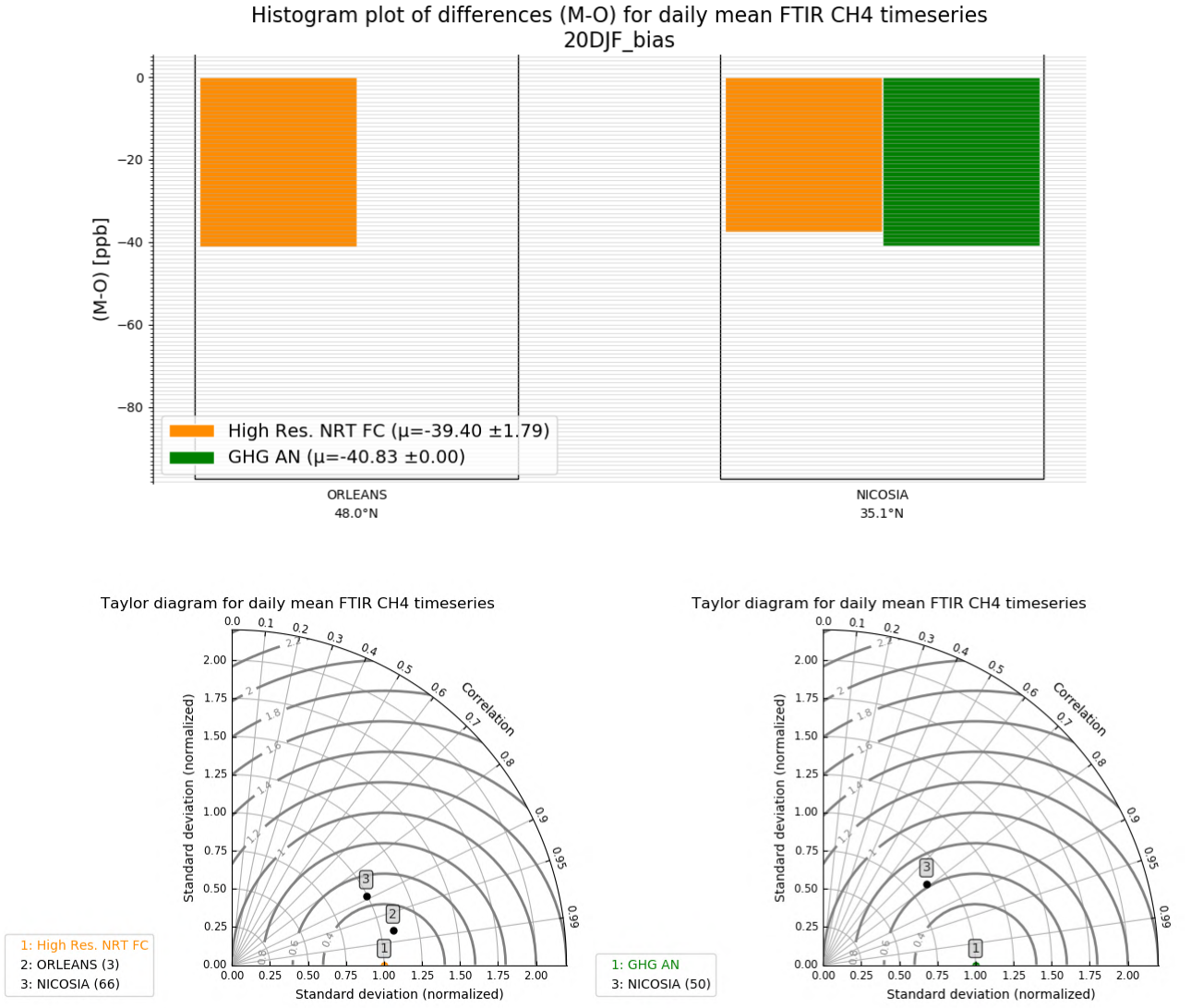


Figure 10.2.4: Histogram plots and Taylor diagrams for comparisons against NDACC FTIR CH_4 for the comparison period. Orange: CAMS GHG high resolution forecast; green: GHG analysis.

Carbon dioxide (CO_2)

Figure 10.2.5 shows the comparisons for the last 2.5 years. For the reporting period the models compare with the measurements within 1%. The comparison is shown for the sites Orleans (Fig. 10.2.6) and Nicosia (Fig. 10.2.7). At Nicosia it obvious that the seasonality is not well captured by the models.

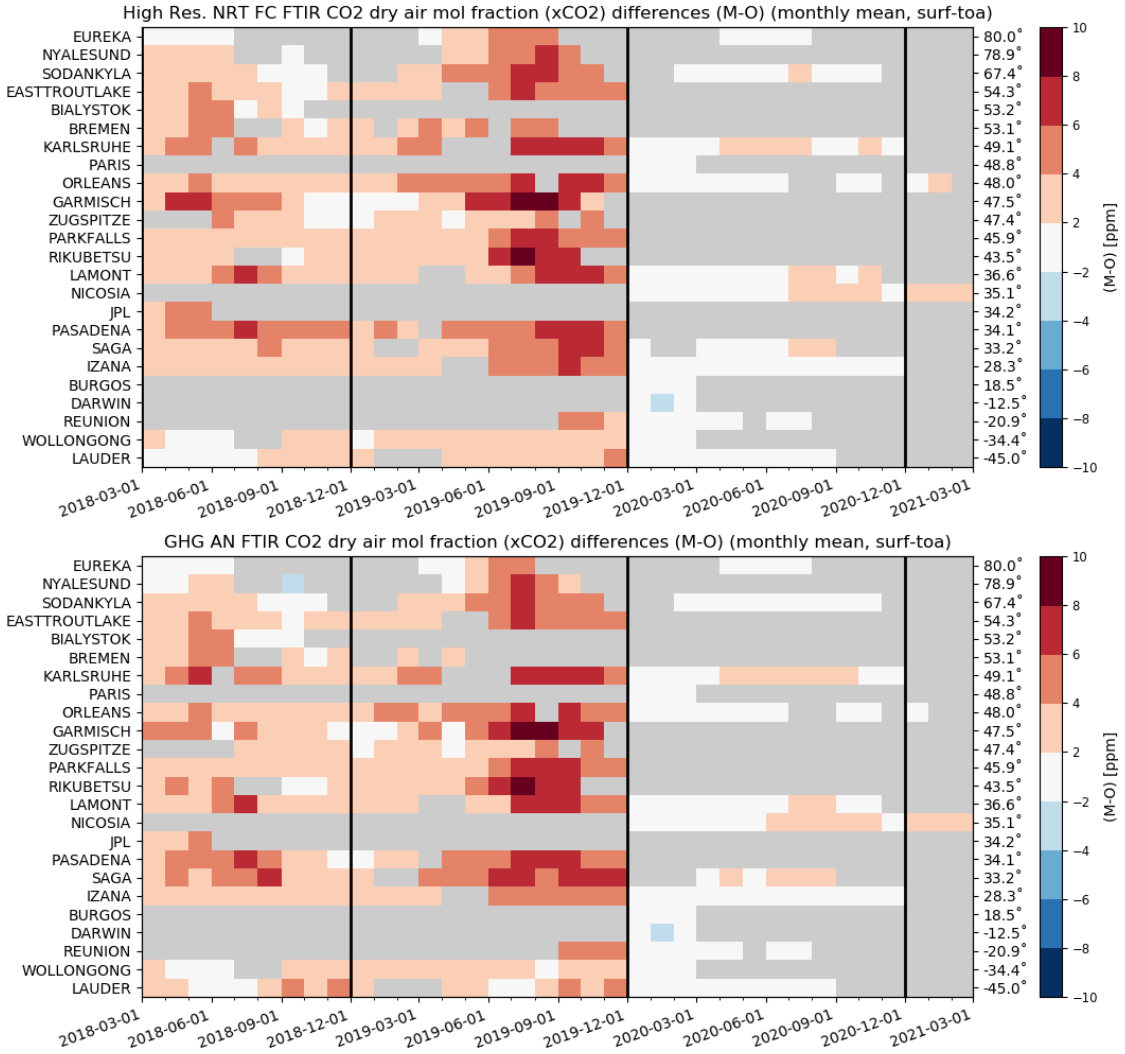


Figure 10.2.5: Monthly differences between FTIR CO_2 and the CAMS results for the last 4 years (upper plot: high resolution NRT, lower plot: GHG AN). The stations are sorted by latitude (northern to southern hemisphere).

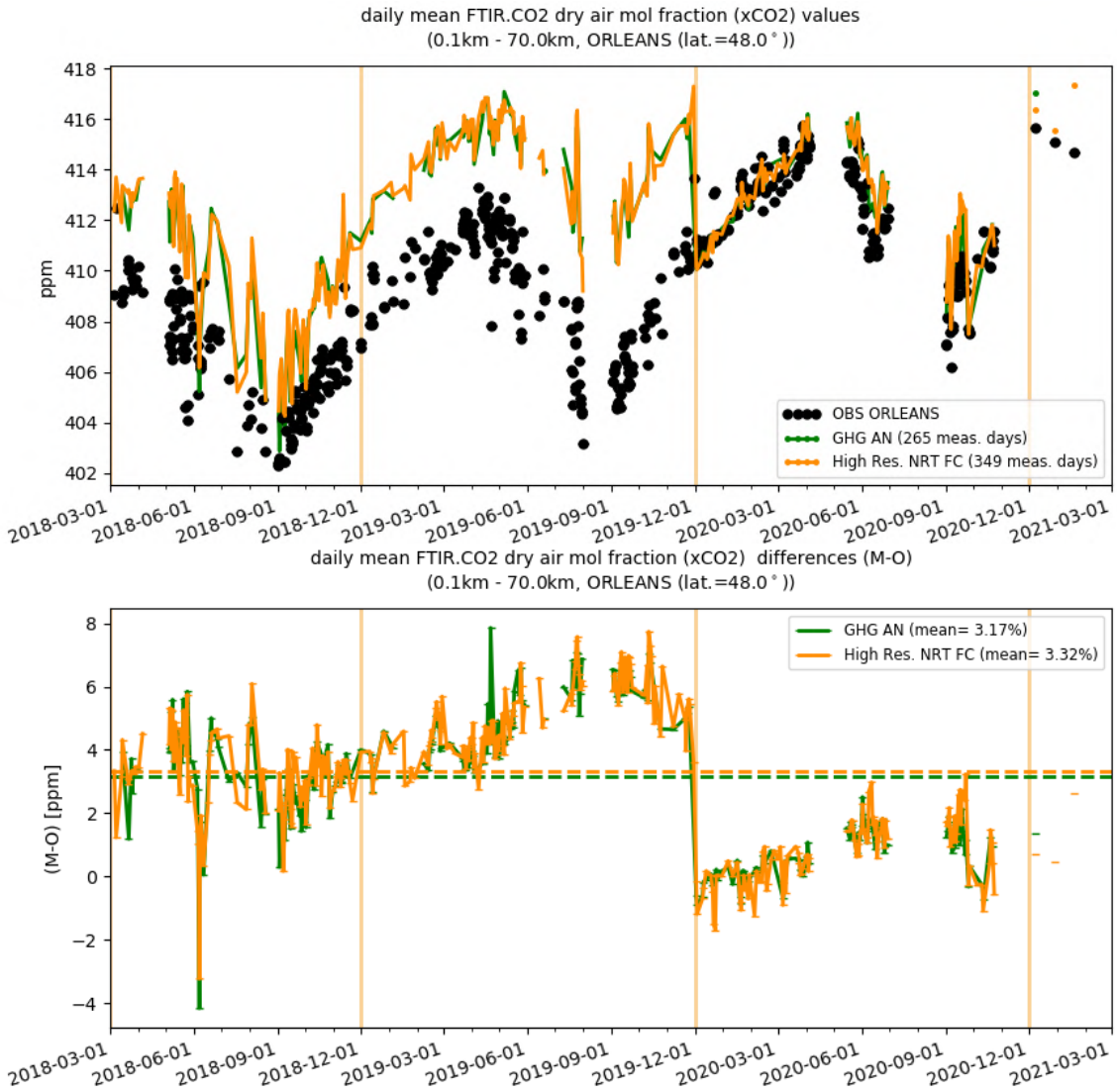


Figure 10.2.6: Comparison of the CO_2 CAMS data with TCCON CO_2 at Orleans. Green: analysis; orange: high-resolution forecast.

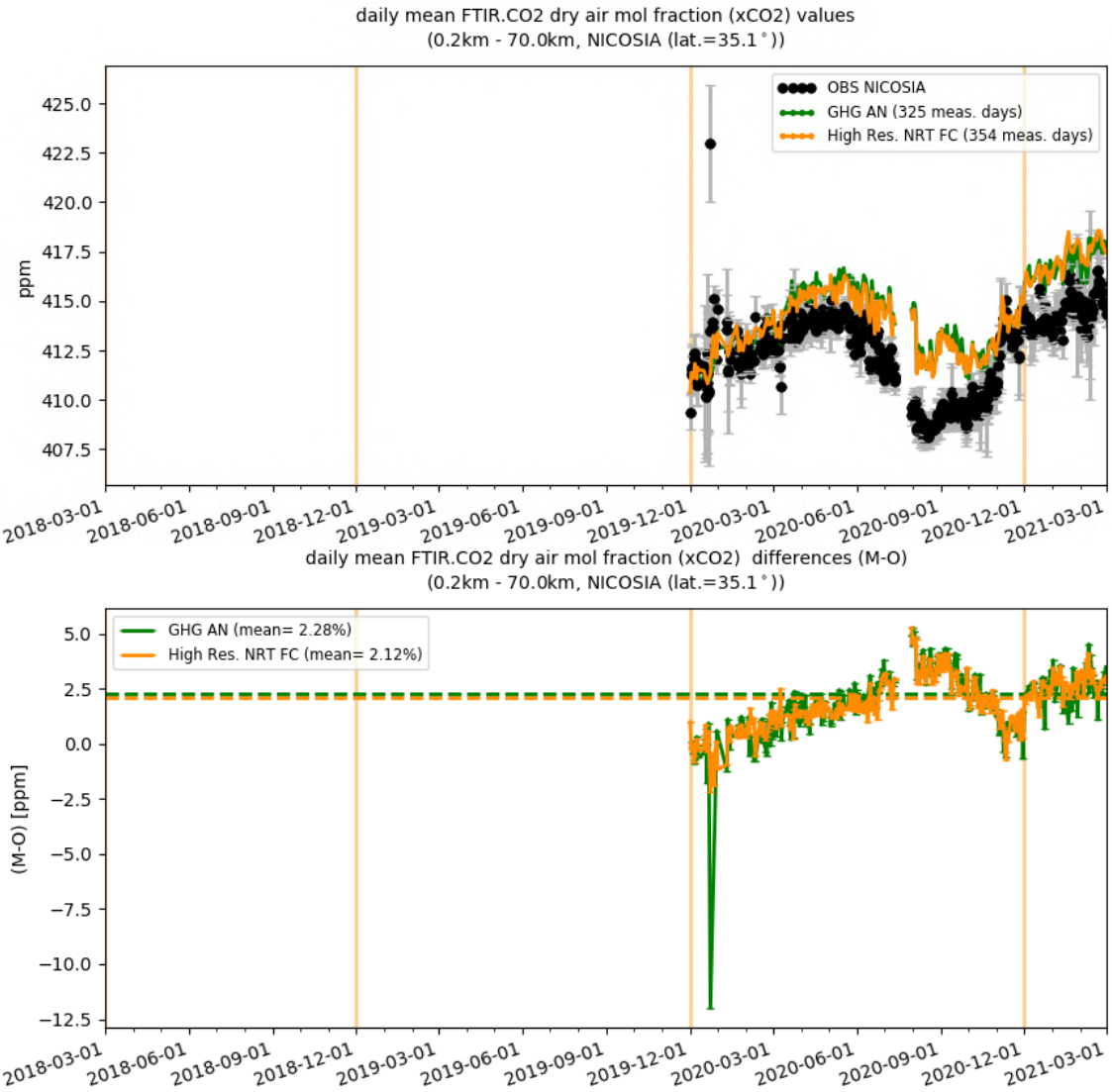


Figure 10.2.7: Comparison of the CO_2 CAMS data with TCCON CO_2 at Nicosia.

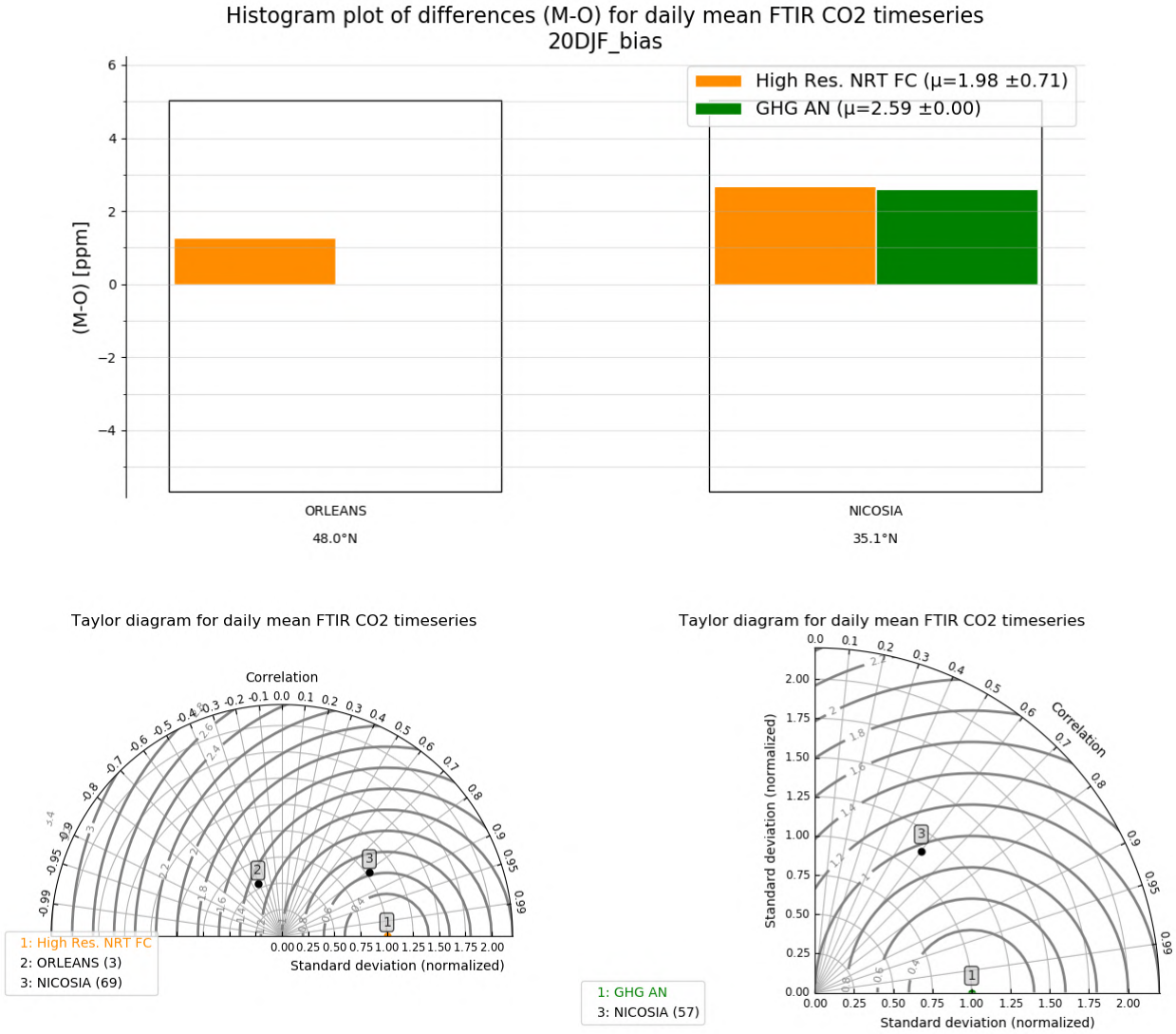


Figure 10.2.8 Histogram plots and Taylor diagrams for CO_2 for the comparison period.

10.3 Validation against FTIR observations from the NDACC network

In this section, we compare the CH4 profiles of the CAMS GHG products 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 CAMS 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 2%, while the uncertainty on the stratospheric column is 7.5%, adding together to a 3% uncertainty on the total column. The systematic uncertainty is large for the NDACC methane product mostly due to higher spectroscopic uncertainties.

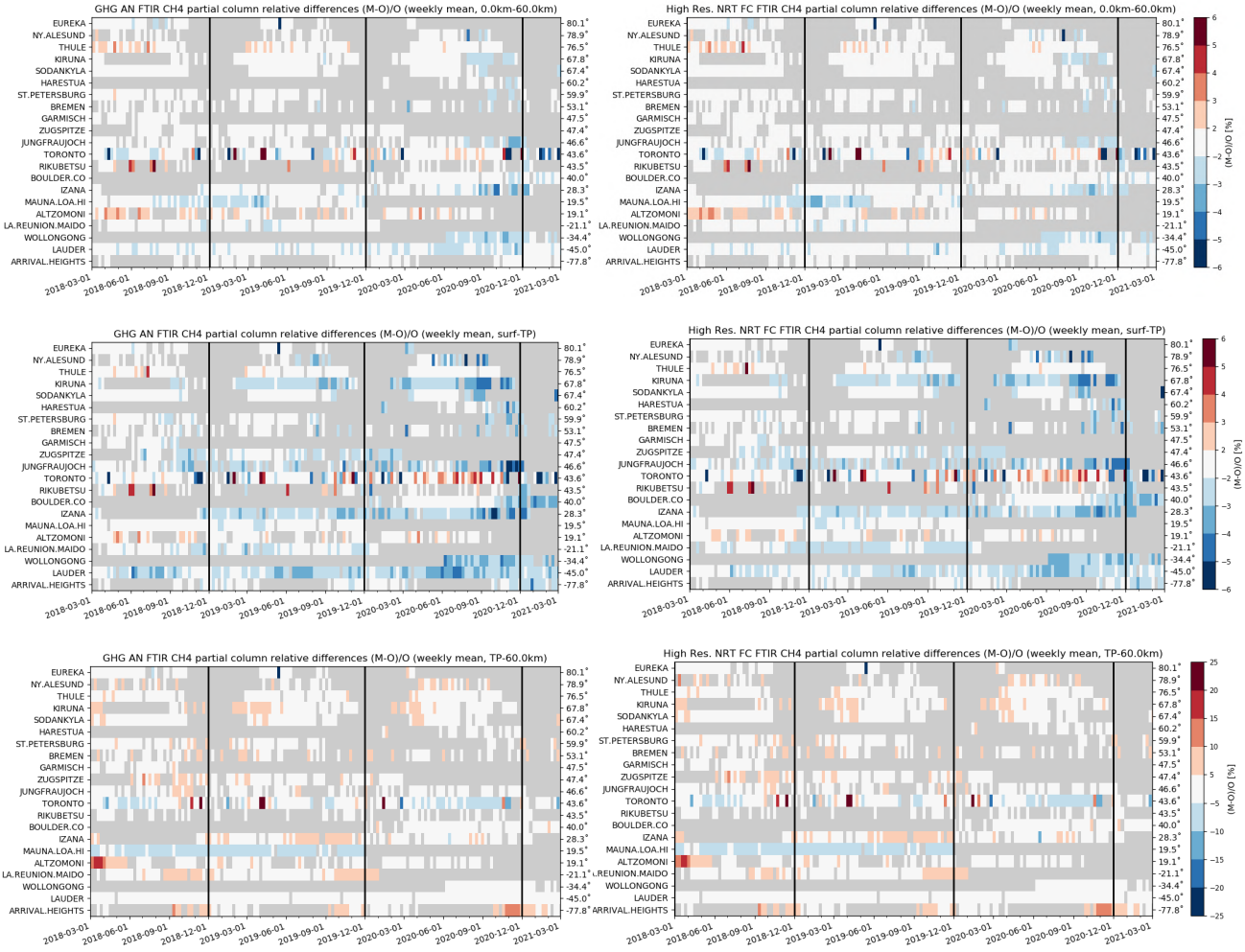


Figure 10.3.1: Weekly mean relative bias for total (top row), tropospheric (middle row) and stratospheric CH_4 columns (bottom row) for the period March 2018 – February 2021 for high resolution forecast (left) and the analysis (right). The overall uncertainty for the CH4 total column measurements is approximately 4%. The overall uncertainty for the CH_4 total/tropospheric column measurements is approximately 2%, while the stratospheric uncertainty is 7.5% (the colour scale for the mosaic plots follows the uncertainty scale)

Figure 10.3.1 (middle row) shows that the tropospheric columns of CH_4 agree well and only small differences appear between the analysis and the high-resolution run. In comparison with the measurement uncertainty, a slight underestimation is observed in the tropospheric columns which agrees with the TCCON results. This underestimation has increased during the last 6 months. The Paramaribo measurements have reduced sensitivity and the tropospheric/stratospheric split is not valid in this case and the Toronto FTIR time series suffers from low outlying values.

The stratospheric columns (Figure 10.3.1, bottom row) show a slight overestimation compared to the measurement uncertainty.

Figure 10.3.2 shows Taylor diagrams for the DJF-2021 period and for a selected number of sites: some stations have limited observations and should be treated with care. Assimilation has a small effect on the correlation coefficients for most sites: the average correlation for 9 stations is 0.83 for the analysis and 0.87 for the high-resolution forecast. Table 10.3.1 contains detailed statistics per site.

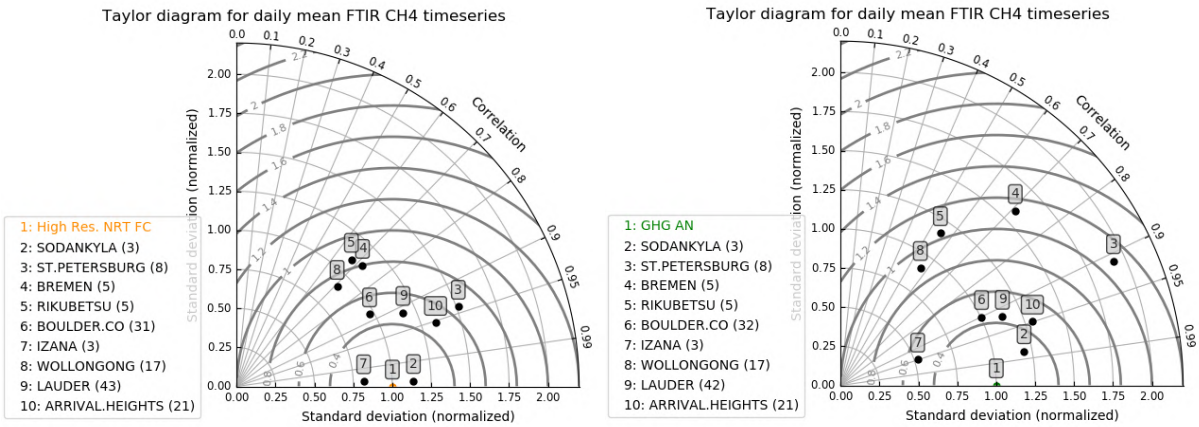


Figure 10.3.2: Taylor diagrams relating the standard deviations for the model /GB time series of daily mean total CH_4 column data and their correlation for the period DJF 2021 (the stations with a limited number of measurements should be ignored, Toronto is left out). All time-series are normalized such that the standard deviation of the CAMS column time series is 1.

Table 10.3.1: Detailed statistics for total CH_4 column comparisons against FTIR measurements during SON 2020. Both analysis and highres forecast behave similar.

	Highres Forecast total column					ghg AN total column					
FTIR site	#	rel. std	corr	rel diff (%)	rel diff std(%)	#	rel. std	corr	rel diff (%)	rel diff std(%)	lat
SODANKYLA	3	1.1	1	-3.02	0.18	3	1.2	0.98	-2.76	0.39	67.4
ST.PETERSBURG	8	1.5	0.94	-0.82	0.46	8	1.9	0.91	-0.89	0.59	59.9
BREMEN	5	1.1	0.72	-1.53	0.63	5	1.6	0.71	-1.54	0.62	53.1
RIKUBETSU	5	1.1	0.67	-1.87	0.45	5	1.2	0.55	-2.01	0.52	43.5
BOULDER.CO	31	1	0.88	-1.8	0.46	32	1	0.9	-1.83	0.45	40
IZANA	3	0.8	1	-2.62	0.16	3	0.5	0.95	-2.75	0.67	28.3
WOLLONGONG	17	0.9	0.71	-2.45	0.55	17	0.9	0.56	-2.46	0.67	-34.4
LAUDER	43	1.2	0.91	-1.35	0.83	42	1.1	0.92	-1.6	0.82	-45
ARRIVAL.HEIGHTS	21	1.3	0.95	-0.35	0.5	21	1.3	0.95	-0.39	0.49	-77.8
		1.1	0.87	-1.76	0.47		1.2	0.83	-1.8	0.58	

11. Event studies

11.1 Dusty February 2021 in Europe and the North Tropical Atlantic

During February 2021, several intense Africa dust events arrived in Europe and the Canary Islands (see MODIS/Terra images in Figure 11.1). On 5-6 February, the African dust was transported from northeast Algeria caused reddish skies in large parts of Europe. This event coincided with snow and rain giving the Pyrenees and the Alps a brownish color. On 15-19 February, a Saharan dust outbreak severely affected the Canary Islands before heading toward continental Europe, reaching as far north as Scandinavia. Finally, on 21 February a new dust event with origin in Algeria hit the Western Mediterranean.

The CAMS o-suite (see Figure 11.2) shows AOD values above 1.2 during the most intense moments of each of the selected dust outbreaks: in the western Mediterranean and central Europe on 5th February, in North Tropical Atlantic on 17th February and in the western Mediterranean, Spain and France in 21st February. The comparison with MODIS shows how the o-suite can reproduce the spatial and temporal distribution of the observed AOD plumes. However, DOD appears underestimated, particularly in long-range transport regions in Europe. The comparison over the OHP Observatory and Dakar AERONET sites (Figure 11.3) shows an enhancement of AOD compared to the control run that is not observed in the natural contribution. This emphasizes the effect of the data assimilation during extreme aerosol events. PM10 concentration above 100 $\mu g/m^3$ are predicted in during the most intense peaks (see Figure 11.4). These PM10 values are a little bit overestimated as shows the comparison with an EEA European sites in Southern Europe (in Figure 8.3.5) and Central Europe (FR26012 in Figure 11.4).

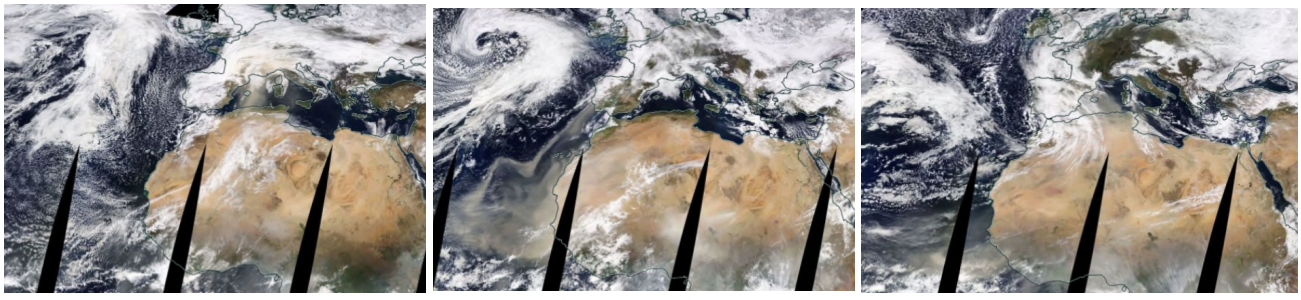


Figure 11.1: Daily visible composite of NASA MODIS Terra on 6, 18 and 21 February 2021 over Europe and Canary Islands. These images are a zoom of the images included in the comparison with CAMS o-suite in Figure 11.2.

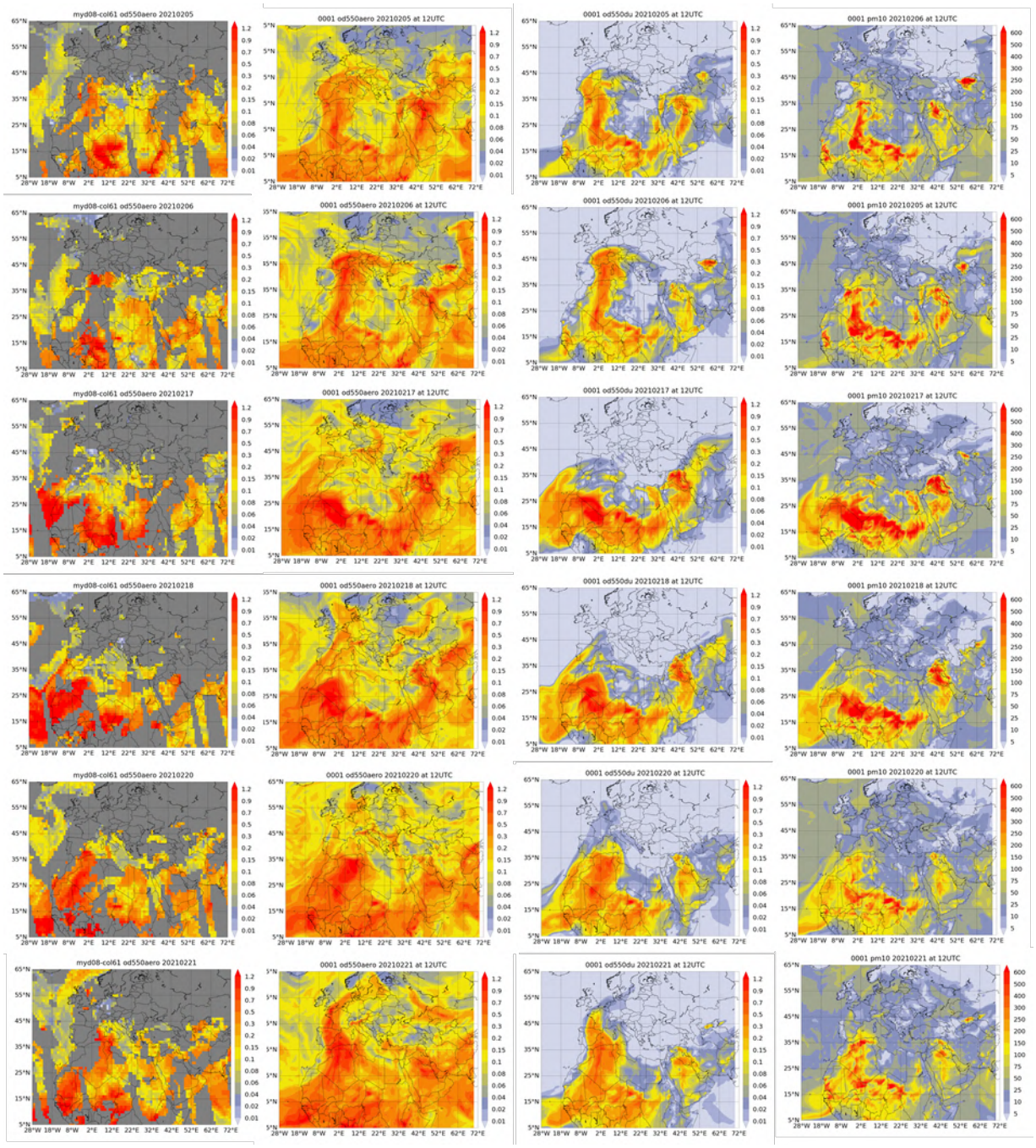


Figure 11.2: Observed AOD from MODIS/Aqua Level 3 and modelled o-suite AOD, DOD and PM10 (in µg/m^3) at 12UTC from o-suite for 5-6, 17-18 and 20-21 February 2021.

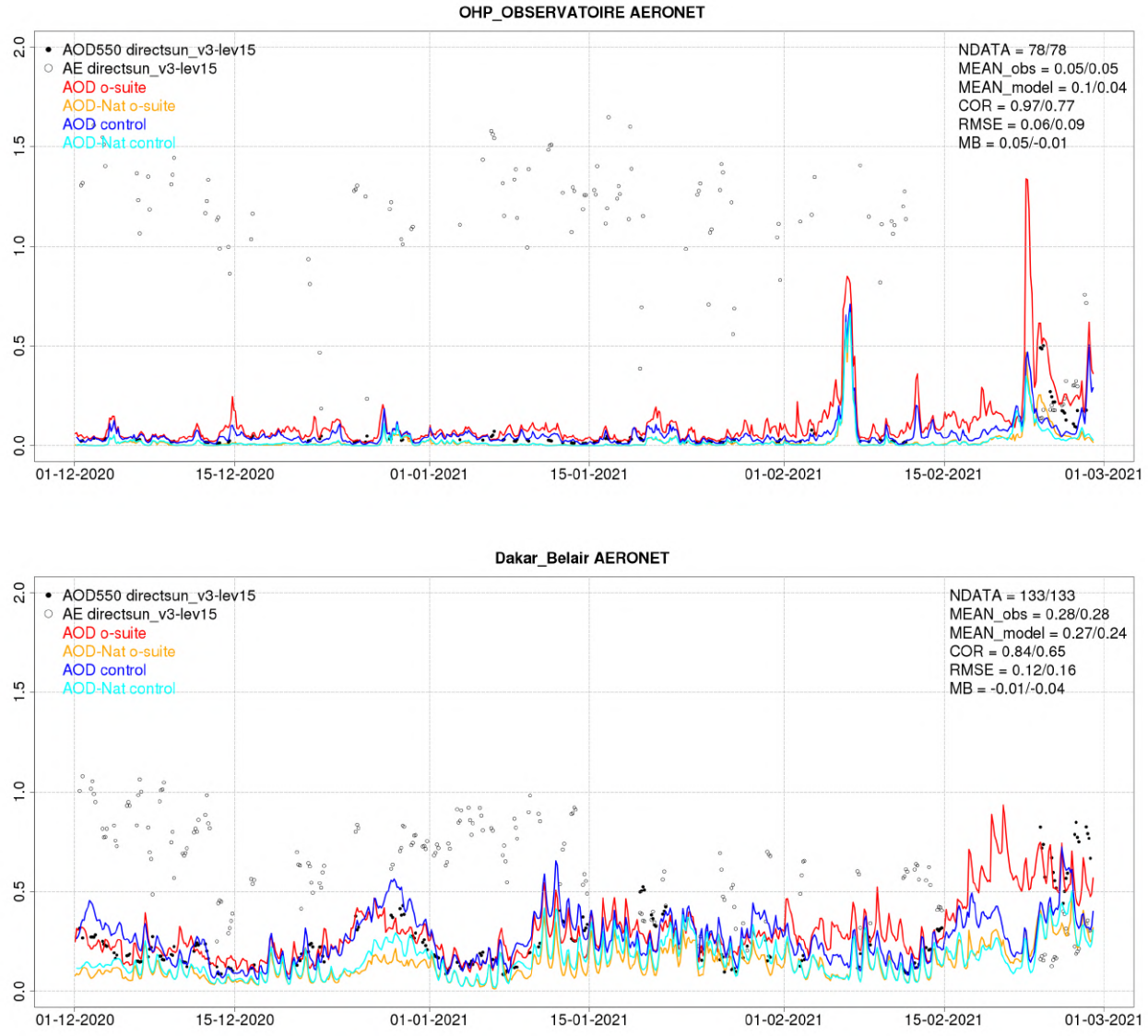


Figure 11.3: AOD from AERONET (black dot), AOD o-suite (red line), AOD control (blue line), AOD-Nat o-suite (orange line), AOD-Nat control (cyan line), from December 2020 to February 2021 over OHP Observatory (France) and Santa Cruz de Tenerife (Canary Islands). AOD-Nat corresponds to the natural aerosol optical depth that includes dust and sea-salt. Skill scores per 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).

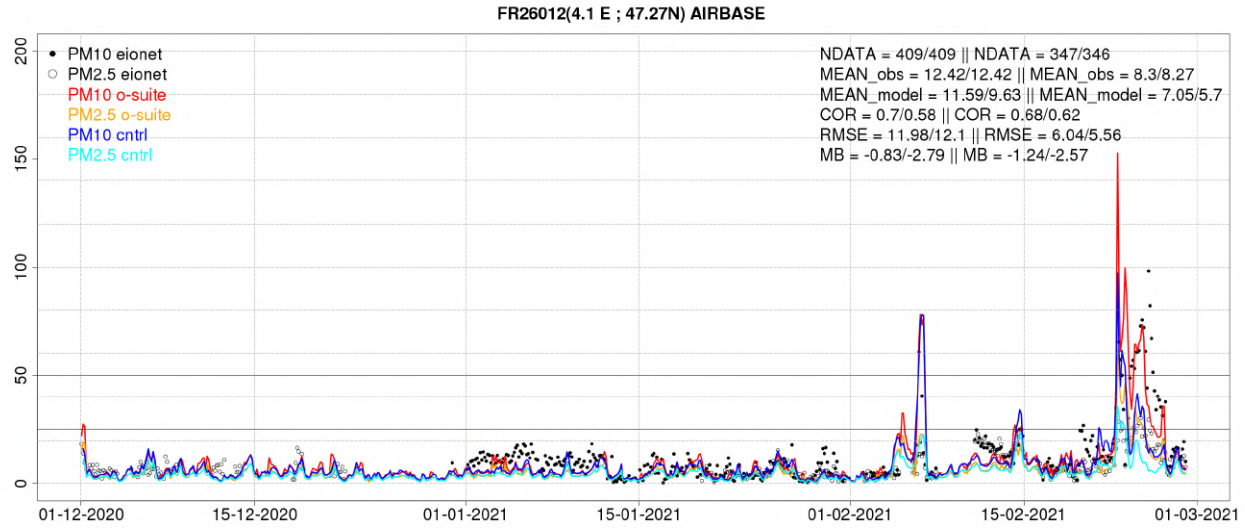


Figure 11.4: PM10 and PM2.5 Airbase observations (black and grey dots, respectively), PM10 and PM2.5 o-suite (red and orange lines, respectively) and PM10 and PM2.5 control (blue and cyan lines, respectively) from December 2020 to February 2021 over FR26012 (France). Skill scores per site and model (o—suite/control) are shown in the upper right corner (NDATA: available 3-hourly values used for the calculations, MEAN observations, MEAN model, COR, RMSE, MB).

12. References

Albert, M. F. M. A., Anguelova, M. D., Manders, A. M. M., Schaap, M., and de Leeuw, G.: Parameterization of oceanic whitecap fraction based on satellite observations, Atmos. Chem. Phys., 16, 13725–13751, https://doi.org/10.5194/acp-16-13725-2016, 2016.

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

Bai, K., Li, K., Guo, J., Yang, Y., & Chang, N.-B., Filling the gaps of in situ hourly PM2.5 concentration data with the aid of empirical orthogonal function analysis constrained by diurnal cycles, Atmospheric Measurement Techniques, 13(3), 1213–1226. https://doi.org/ 10.5194/amt-13-1213-2020, 2020.

Basart, S, A. Benedictow, Y. Bennouna, A.-M. Blechschmidt, S. Chabrillat, Y. Christophe, E. Cuevas, H. J. Eskes, K. M. Hansen, O. Jorba, J. Kapsomenakis, B. Langerock, T. Pay, A. Richter, N. Sudarchikova, M. Schulz, A. Wagner, C. Zerefos, Upgrade verification note for the CAMS near-real time global atmospheric composition service: Evaluation of the e-suite for the CAMS upgrade of July 2019, Copernicus Atmosphere Monitoring Service (CAMS) report, CAMS84_2018SC1_D3.2.1-201907_esuite_v1.pdf, July 2019.

Beirle, S., Hörmann, C., Jöckel, P., Liu, S., Penning de Vries, M., Pozzer, A., Sihler, H., Valks, P., and Wagner, T.: The STRatospheric Estimation Algorithm from Mainz (STREAM): estimating stratospheric NO_2 from nadir-viewing satellites by weighted convolution, Atmos. Meas. Tech., 9, 2753–2779, https://doi.org/10.5194/amt-9-2753-2016, 2016.

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.

Birmili, W., Schepanski, K., Ansmann, A., Spindler, G., Tegen, I., Wehner, B., Nowak, A., Reimer, E., Mattis, I., Müller, K., Brüggemann, E., Gnauk, T., Herrmann, H., Wiedensohler, A., Althausen, D., Schladitz, A., Tuch, T., and Löschau, G.: A case of extreme particulate matter concentrations over Central Europe caused by dust emitted over the southern Ukraine, Atmos. Chem. Phys., 8, 997–1016, https://doi.org/10.5194/acp-8-997-2008, 2008.

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.

Carn, S. A., V. E. Fioletov, C. A. McLinden, C. Li & N. A. Krotkov, A decade of global volcanic SO2 emissions measured from space, Scientific Reports volume 7, Article number: 44095 (2017).

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, http://www.atmos-chem-phys.net/6/4321/2006/acp-6-4321-2006.pdf.

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

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. J., S. Basart, A. Benedictow, Y. Bennouna, A.-M. Blechschmidt, S. Chabrillat, Y. Christophe, K. M. Hansen, J. Kapsomenakis, B. Langerock, M. Pitkänen, M. Ramonet, A. Richter, N. Sudarchikova, M. Schulz, A. Wagner, T. Warneke (UBC), C. Zerefos, Upgrade verification note for the CAMS near-real time global atmospheric composition service: Evaluation of the e-suite for the CAMS 47R1 upgrade of October 2020, Copernicus Atmosphere Monitoring Service (CAMS) report, CAMS84_2018SC2_D3.2.1- 202009_esuite.pdf, 2 October 2020, doi:10.24380/fzdx-j890.

Eskes, H.J., S. Basart, A. Benedictow, Y. Bennouna, A.-M. Blechschmidt, S. Chabrillat, Y. Christophe, E. Cuevas, H. Flentje, K. M. Hansen, J. Kapsomenakis, B. Langerock, M. Ramonet, A. Richter, M. Schulz, N. Sudarchikova, A. Wagner, T. Warneke, C. Zerefos, Observation characterisation and validation methods document,

Copernicus Atmosphere Monitoring Service (CAMS) report, CAMS84_2018SC2_D6.1.1-2020_observations_v5.pdf, January 2021. Available from: http://atmosphere.copernicus.eu/user-support/validation/verification-global-services

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 (2018b)

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 (2018c).

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.

Flentje, Harald, Ina Mattis, Zak Kipling, Samuel Rémy, and Werner Thomas, Evaluation of ECMWF IFS-AER (CAMS) operational forecasts during cycle 41r1–46r1 with calibrated ceilometer profiles over Germany, Geosci. Model Dev., 14, 1721–1751, 2021, https://doi.org/10.5194/gmd-14-1721-2021

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.

Joly, M. and Peuch, V. H., Objective classification of air quality monitoring sites over Europe, Atmospheric Environment, 47, 111-123, 2012.

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.

Karion, A., C. Sweeney, P. Tans and T. Newberger (2010). AirCore: An Innovative Atmospheric Sampling System. Journal of Atmospheric and Oceanic Technology 27(11): 1839-1853.

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

Membrive, O., C. Crevoisier, C. Sweeney, F. Danis, A. Hertzog, A. Engel, H. Bönisch and L. Picon (2017). AirCore-HR: a high-resolution column sampling to enhance the vertical description of CH4 and CO2. Atmos. Meas. Tech. 10(6): 2163-2181

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.

Rémy, S., Kipling, Z., Flemming, J., Boucher, O., Nabat, P., Michou, M., Bozzo, A., Ades, M., Huijnen, V., Benedetti, A., Engelen, R., Peuch, V.-H., and Morcrette, J.-J.: Description and evaluation of the tropospheric

aerosol scheme in the European Centre for Medium-Range Weather Forecasts (ECMWF) Integrated Forecasting System (IFS-AER, cycle 45R1), Geosci. Model Dev., 12, 4627–4659, https://doi.org/10.5194/gmd-12-4627-2019, 2019.

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.

Sudarchikova, N., M. Schulz, Q. Errera, M. Ramonet, H. J. Eskes, S. Basart, A. Benedictow, Y. Bennouna, A.-M. Blechschmidt, S. Chabrillat, Christophe, Y., E. Cuevas, A. El-Yazidi, H. Flentje, P. Fritzsche, K.M. Hansen, U. Im, J. Kapsomenakis, B. Langerock, A. Richter, V. Thouret, A. Wagner, T. Warneke, C. Zerefos, Validation report of the CAMS near-real-time global atmospheric composition service: Period September - November 2020, Copernicus Atmosphere Monitoring Service (CAMS) report, CAMS84_2018SC3_D1.1.1_SON2020.pdf, March 2021, doi: 10.24380/rysv-7371.

Stauffer, R. M., Thompson, A. M., Kollonige, D. E., Witte, J. C., Tarasick, D. W., and Davies, J., Voemel, H., Morris, G. A., VanMalderen, R., Johnson, B. J. J., Querel, R.R., Selkirk, H.B., Stübi, R., Smit, H.G.J: A Post-2013 Drop-off in Total Ozone at Half of Global Ozonesonde Stations: ECC Instrument Artifacts? Earth and Space Science Open Archive JF M310.1002/essoar.10501543.1, 2020.

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.

Thompson, A. M., Witte, J. C., Sterling, C.,Jordan, A., Johnson, B. J., Oltmans, S. J.,Thiongo, K. (2017). First reprocessing of Southern Hemisphere Additional Ozonesondes (SHADOZ) ozone profiles (1998\u20132016): 2. Comparisons with satellites and ground-based instruments. Journal of Geophysical Research: Atmospheres, 122, 13,000\u201313,025. https://doi.org/10.1002/2017JD027406, 2017.

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.

van Geffen, J., Boersma, K. F., Eskes, H., Sneep, M., ter Linden, M., Zara, M., and Veefkind, J. P.: S5P TROPOMI NO2 slant column retrieval: method, stability, uncertainties and comparisons with OMI, Atmos. Meas. Tech., 13, 1315–1335, https://doi.org/10.5194/amt-13-1315-2020, 2020.

*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

Witte, J. C., Thompson, A. M., Smit, H. G. J., Vömel, H., Posny, F., & Stübi, R. : First reprocessing of Southern Hemisphere ADditional OZonesondes profile records: 3. Uncertainty in ozone profile and total column. Journal of Geophysical Research: Atmospheres, 123, 3243\u20133268. https://doi.org/10.1002/2017JD027791, 2018.

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

Natalia Sudarchikova, MPG (editor, satellite IR observations)
Annette Wagner, DWD (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)
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)
Abdelhadi El-Yazidi 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.

S5P lv1 radiances and NO_2 operational were provided by EU Copernicus.

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), the SAGE III-ISS ozone data https://eosweb.larc.nasa.gov/project/sageiii-iss/sageiii-iss_table 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 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.

We are grateful to the IAGOS operators from the various institutes which are members of IAGOS-AISBL (http://www.iagos.org). The authors also acknowledge the strong support of the European Commission, Airbus, and the airlines (Lufthansa, Air France, Austrian, Air Namibia, Cathay Pacific, Iberia, China Airlines and Hawaiian Airlines so far) which have carried the MOZAIC or IAGOS equipment and undertaken maintenance since 1994. In the last 10 years of operation, MOZAIC has been funded by INSU-CNRS (France), Météo-France, Université Paul Sabatier (Toulouse, France) and Research Center Jülich (FZJ, Jülich, Germany). IAGOS has been additionally funded by the EU projects IAGOS-DS and IAGOS-ERI. The MOZAIC–IAGOS database (http://www.iagos-data.fr) is supported by AERIS (CNES and INSU-CNRS). Data are also available via AERIS web site www.aeris-data.fr.

We acknowledge the contribution of the ICOS Atmospheric Thematic Center (Lynn Hazan, Amara Abbaris, and Leonard Rivier) for the near real time data processing of surface CO_2 and CH_4 concentrations. The ICOS monitoring sites are maintained by the national networks: ICOS-Belgium (Martine De Mazière , Mahesh Kumar Sha , Nicolas Kumps), ICOS-Czech Rep. (Michal Marek, Katerina Komínková, Gabriela Vítková), ICOS-Finland (Olli Peltola, Janne Levula, Tuomas Laurila, Juha Hatakka, Ari Leskinen, Ivan Mammarella, Petri Keronen), ICOS-France (Michel Ramonet, Marc Delmotte, Sebastien Conil, Laurent Langrene , Morgan Lopez, Victor Kazan, Aurélie Colomb, Jean Marc Pichon, Olivier Laurent, Camille Yver-Kwok, Zineb Mandrick, Jean-Marc Metzger), ICOS Germany (Matthias Lindauer, Dagmar Kubistin, Christian Plass-Duelmer, Dietmar Weyrauch, Jennifer Müller-Williams), ICOS-Italy (Paolo Cristofanelli, Michela Maione, Francesco Apadula, Pamela Trisolino), ICOS-Netherlands (Huilin Chen, Bert Scheeren), ICOS-Norway (Cathrine Lund Myhre, Ove Hermansen, Chris Lunder), ICOS-Sweden (Jutta Holst, Michal Heliasz, Meelis Molder, Mikaell Ottosson Lofvenius, Irene Lehner, Per Marklund, Paul Smith), ICOS-Switzerland (Martin Steinbacher, Markus Leuenberger), European Commission, Joint Research Centre, Directorate for Energy, Transport and Climate (Peter Bergamaschi, Giovanni Manca).

In collaboration with LSCE and the French Aircore program, the Cyprus Institute has coordinated the first AIRCORE campaign in Cyprus within the framework of the EMME-CARE (Eastern Mediterranean and Middle East – Climate and Atmosphere Research). Thanks to C. Rousogenous, P.Y. Quéhé, Th. Laemmel, C. Keleshis, P.Antoniou, and J.Sciare. The AIRCORE campaign at Trainou was coordinated by Th. Laemmel and J.Moyé from LSCE, with funding from CEA, OVSQ and CNES.

References (ICOS):

Hazan, L., Tarniewicz, J., Ramonet, M., Laurent, O., and Abbaris, A.: Automatic processing of atmospheric CO2 and CH4 mole fractions at the ICOS Atmosphere Thematic Centre, Atmos. Meas. Tech., 9, 4719–4736, https://doi.org/10.5194/amt-9-4719-2016, 2016.

Laurent, O., ICOS Atmosphere Monitoring Station Assembly, and ICOS Atmosphere Thematic Centre (ATC): ICOS Atmospheric Station Specifications v1.3, ICOS ERIC, doi:10.18160/SDW6-BX90, 2017.

Yver-Kwok, C., C. Philippon, P. Bergamaschi, T. Biermann, F. Calzolari, H. Chen, S. Conil, P. Cristofanelli, M. Delmotte, J. Hatakka, M. Heliasz, O. Hermansen, K. Komínková, D. Kubistin, N. Kumps, O. Laurent, T. Laurila, I. Lehner, J. Levula, M. Lindauer, M. Lopez, I. Mammarella, G. Manca, P. Marklund, J. M. Metzger, M. Mölder, S. M. Platt, M. Ramonet, L. Rivier, B. Scheeren, M. K. Sha, P. Smith, M. Steinbacher, G. Vítková and S. Wyss (2021). Evaluation and optimization of ICOS atmosphere station data as part of the labeling process. Atmos. Meas. Tech. 14(1): 89-116.

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.

References (TCCON):

Blumenstock, T., F. Hase, M. Schneider, O. E. García, and E. Sepúlveda. 2017. "TCCON data from Izana (ES), Release GGG2014.R1." CaltechDATA. doi:10.14291/tccon.ggg2014.izana01.r1.

De Mazière, M., M. K. Sha, F. Desmet, C. Hermans, F. Scolas, N. Kumps, J.-M. Metzger, V. Duflot, and J.-P. Cammas. 2017. "TCCON data from Réunion Island (RE), Release GGG2014.R1." CaltechDATA. doi:10.14291/tccon.ggg2014.reunion01.r1.

Deutscher, N. M., J. Notholt, J. Messerschmidt, C. Weinzierl, T. Warneke, C. Petri, and P. Grupe. 2017. "TCCON data from Bialystok (PL), Release GGG2014.R1." CaltechDATA. doi:10.14291/tccon.ggg2014.bialystok01.r1/1183984.

Dubey, M. K., B. G. Henderson, D. Green, Z. T. Butterfield, G. Keppel-Aleks, N. T. Allen, J.-F. Blavier, C. M. Roehl, D. Wunch, and R. Lindenmaier. 2017. "TCCON data from Manaus (BR), Release GGG2014.R0." CaltechDATA. doi:10.14291/tccon.ggg2014.manaus01.r0/1149274.

Dubey, M. K., R. Lindenmaier, B. G. Henderson, D. Green, N. T. Allen, C. M. Roehl, J.-F. Blavier, et al. 2017. "TCCON data from Four Corners (US), Release GGG2014.R0." CaltechDATA. doi:10.14291/tccon.ggg2014.fourcorners01.r0/1149272.

Feist, D. G., S. G. Arnold, N. John, and M. C. Geibel. 2017. "TCCON data from Ascension Island (SH), Release GGG2014.R0." CaltechDATA. doi:10.14291/tccon.ggg2014.ascension01.r0/1149285.

Goo, T.-Y., Y.-S. Oh, and V. A. Velazco. 2017. "TCCON data from Anmeyondo (KR), Release GGG2014.R0." CaltechDATA. doi:10.14291/tccon.ggg2014.anmeyondo01.r0/1149284.

Griffith, D. W. T., N. M. Deutscher, V. A. Velazco, P. O. Wennberg, Y. Yavin, G. Keppel-Aleks, R. A. Washenfelder, et al. 2017. "TCCON data from Darwin (AU), Release GGG2014.R0." CaltechDATA. doi:10.14291/tccon.ggg2014.darwin01.r0/1149290.

Griffith, D. W. T., V. A. Velazco, N. M. Deutscher, C. Paton-Walsh, N. B. Jones, S. R. Wilson, R. C. Macatangay, G. C. Kettlewell, R. R. Buchholz, and M. O. Riggenbach. 2017. "TCCON data from Wollongong (AU), Release GGG2014.R0." CaltechDATA. doi:10.14291/tccon.ggg2014.wollongong01.r0/1149291.

Hase, F., T. Blumenstock, S. Dohe, J. Groß, and M.ä. Kiel. 2017. "TCCON data from Karlsruhe (DE), Release GGG2014.R1." CaltechDATA. doi:10.14291/tccon.ggg2014.karlsruhe01.r1/1182416.

Iraci, L. T., J. R. Podolske, P. W. Hillyard, C. Roehl, P. O. Wennberg, J.-F. Blavier, J. Landeros, et al. 2017. "TCCON data from Edwards (US), Release GGG2014.R1." CaltechDATA. doi:10.14291/tccon.ggg2014.edwards01.r1/1255068.

—. 2017. "TCCON data from Indianapolis (US), Release GGG2014.R1." CaltechDATA. doi:10.14291/tccon.ggg2014.indianapolis01.r1/1330094.

Kawakami, S., H. Ohyama, K. Arai, H. Okumura, C. Taura, T. Fukamachi, and M. Sakashita. 2017. "TCCON data from Saga (JP), Release GGG2014.R0." CaltechDATA. doi:10.14291/tccon.ggg2014.saga01.r0/1149283.

Kivi, R., P. Heikkinen, and E. Kyrö. 2017. "TCCON data from Sodankylä (FI), Release GGG2014.R0." CaltechDATA. doi:10.14291/tccon.ggg2014.sodankyla01.r0/1149280.

Liu, Cheng, Wei Wang, and Youwen Sun. 2018. "TCCON data from Hefei (PRC), Release GGG2014.R0." CaltechDATA. doi:10.14291/tccon.ggg2014.hefei01.r0.

Morino, I., N. Yokozeki, T. Matsuzaki, and M. Horikawa. 2017. "TCCON data from Rikubetsu (JP), Release GGG2014.R2." CaltechDATA. doi:10.14291/tccon.ggg2014.rikubetsu01.r2.

Morino, I., T. Matsuzaki, and M. Horikawa. 2017. "TCCON data from Tsukuba (JP), 125HR, Release GGG2014.R2." CaltechDATA. doi:10.14291/tccon.ggg2014.tsukuba02.r2.

Morino, Isamu, Voltaire A. Velazco, Akihiro Hori, Osamu Uchino, and David W. T. Griffith. 2018. "TCCON data from Burgos, Ilocos Norte (PH), Release GGG2014.R0." CaltechDATA. doi:10.14291/tccon.ggg2014.burgos01.r0.

Notholt, J., C. Petri, T. Warneke, N. M. Deutscher, M. Palm, M. Buschmann, C. Weinzierl, R. C. Macatangay, and P. Grupe. 2017. "TCCON data from Bremen (DE), Release GGG2014.R0." CaltechDATA. doi:10.14291/tccon.ggg2014.bremen01.r0/1149275.

Notholt, J., T. Warneke, C. Petri, N. M. Deutscher, C. Weinzierl, M. Palm, and M. Buschmann. 2017. "TCCON data from Ny Ålesund, Spitsbergen (NO), Release GGG2014.R0." CaltechDATA. doi:10.14291/tccon.ggg2014.nyalesund01.r0/1149278.

Pollard, David Frank, John Robinson, and Hisako Shiona. 2019. "TCCON data from Lauder (NZ), Release GGG2014.R0." CaltechDATA. doi:10.14291/tccon.ggg2014.lauder03.r0.

Sherlock, V., B. Connor, J. Robinson, H. Shiona, D. Smale, and D. F. Pollard. 2017. "TCCON data from Lauder (NZ), 120HR, Release GGG2014.R0." CaltechDATA. doi:10.14291/tccon.ggg2014.lauder01.r0/1149293.

—. 2017. "TCCON data from Lauder (NZ), 125HR, Release GGG2014.R0." CaltechDATA. doi:10.14291/tccon.ggg2014.lauder02.r0/1149298.

Strong, K., S. Roche, J. E. Franklin, J. Mendonca, E. Lutsch, D. Weaver, P. F. Fogal, J. R. Drummond, R. Batchelor, and R. Lindenmaier. 2018. "TCCON data from Eureka (CA), Release GGG2014.R3." CaltechDATA. doi:10.14291/tccon.ggg2014.eureka01.r3.

Sussmann, R., and M. Rettinger. 2017. "TCCON data from Garmisch (DE), Release GGG2014.R2." CaltechDATA. doi:10.14291/tccon.ggg2014.garmisch01.r2.

—. 2018. "TCCON data from Zugspitze (DE), Release GGG2014.R1." CaltechDATA. doi:10.14291/tccon.ggg2014.zugspitze01.r1.

Té, Y., P. Jeseck, and C. Janssen. 2017. "TCCON data from Paris (FR), Release GGG2014.R0." CaltechDATA. doi:10.14291/tccon.ggg2014.paris01.r0/1149279.

Warneke, T., J. Messerschmidt, J. Notholt, C. Weinzierl, N. M. Deutscher, C. Petri, and P. Grupe. 2017. "TCCON data from Orléans (FR), Release GGG2014.R0." CaltechDATA. doi:10.14291/tccon.ggg2014.orleans01.r0/1149276.

Wennberg, P. O., C. M. Roehl, D. Wunch, G. C. Toon, J.-F. Blavier, R. Washenfelder, G. Keppel-Aleks, N. T. Allen, and J. Ayers. 2017. "TCCON data from Park Falls (US), Release GGG2014.R1." CaltechDATA. doi:10.14291/tccon.ggg2014.parkfalls01.r1.

Wennberg, P. O., C. M. Roehl, J.-F. Blavier, D. Wunch, and N. T. Allen. 2017. "TCCON data from Jet Propulsion Laboratory (US), 2011, Release GGG2014.R1." CaltechDATA. doi:10.14291/tccon.ggg2014.jpl02.r1/1330096.

Wennberg, P. O., D. Wunch, C. M. Roehl, J.-F. Blavier, G. C. Toon, and N. T. Allen. 2017. "TCCON data from Caltech (US), Release GGG2014.R1." CaltechDATA. doi:10.14291/tccon.ggg2014.pasadena01.r1/1182415.

—. 2017. "TCCON data from Lamont (US), Release GGG2014.R1." CaltechDATA. doi:10.14291/tccon.ggg2014.lamont01.r1/1255070.

Wennberg, P. O., D. Wunch, Y. Yavin, G. C. Toon, J.-F. Blavier, N. T. Allen, and G. Keppel-Aleks. 2017. "TCCON data from Jet Propulsion Laboratory (US), 2007, Release GGG2014.R0." CaltechDATA. doi:10.14291/tccon.ggg2014.jpl01.r0/1149163.

Wunch, D., J. Mendonca, O. Colebatch, N. T. Allen, J.-F. Blavier, S. Roche, J. Hedelius, et al. 2017. "TCCON data from East Trout Lake, SK (CA), Release GGG2014.R1." CaltechDATA. doi:10.14291/tccon.ggg2014.easttroutlake01.r1.

Wunch, D., Toon, G. C., Sherlock, V., Deutscher, N. M., Liu, C., Feist, D. G., & Wennberg, P. O. (2015). The Total Carbon Column Observing Network's GGG2014 Data Version. Tech. rep., California Institute of Technology, Pasadena. doi:10.14291/tccon.ggg2014.documentation.R0/1221662

Copernicus Atmosphere Monitoring Service

ECMWF - Shinfield Park, Reading RG2 9AX, UK

Contact: info@copernicus-atmosphere.eu

atmosphere.copernicus.eu copernicus.eu ecmwf.int