



Copernicus Atmosphere Monitoring Service



# **Validation report for the CAMS global reanalyses of aerosols and reactive trace gases, years 2003 - 2020**

Issued by: KNMI

Date: 28 June 2021

Ref: CAMS84\_2018SC3\_D5.1.1-2020

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



# Validation report for the CAMS global reanalyses of aerosols and reactive trace gases, years 2003 - 2020

## **EDITORS:**

Q. Errera (BIRA-IASB), Y. Bennouna (CNRS-LA),  
M. Schulz (MetNo), H.J. Eskes (KNMI)

## **AUTHORS:**

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

## **VALIDATION REPORT OF THE COPERNICUS ATMOSPHERE MONITORING SERVICE, ISSUED BY KNMI.**

Available at: <http://atmosphere.copernicus.eu/>

## **CITATION:**

Errera, Q., Y. Bennouna, M. Schulz, H.J. Eskes, S. Basart, A. Benedictow, A.-M. Blechschmidt, S. Chabrillat, H. Clark, E. Cuevas, H. Flentje, K.M. Hansen, U. Im, J. Kapsomenakis, B. Langerock, K. Petersen, A. Richter, N. Sudarchikova, V. Thouret, A. Wagner, Y. Wang, T. Warneke, C. Zerefos, Validation report of the CAMS global Reanalysis of aerosols and reactive gases, years 2003-2020, Copernicus Atmosphere Monitoring Service (CAMS) report, CAMS84\_2018SC3\_D5.1.1-2020.pdf, June 2021, doi:10.24380/8gf9-k005.

## **STATUS:**

Version 1

## **DATE:**

28/06/2021



## Summary

The Copernicus Atmosphere Monitoring Service (<http://atmosphere.copernicus.eu>, CAMS) is a component of the European Earth Observation programme Copernicus. As one of the service products, CAMS is currently producing a global reanalysis of reactive trace gases, greenhouse gases and aerosol concentrations (Inness et al., 2019). The production of the reanalysis has started early 2017, and the period 2003-2019 was completed in January 2020. The CAMS reanalysis will be further extended in the years up to 2021.

This document presents the validation results for the 18-year reanalysis run, 2003 - 2020, focussing on aerosols and reactive gases. Earlier versions of this document have appeared during the production of the reanalysis: after completing year 1 (2003), year 5 (2003-2007), year 14 (2003-2016), year 15 (2003-2017), year 16 (2003-2018) and year 17 (2003-2019; Bennouna et al., 2020). The evaluation of the CAMS greenhouse gas reanalysis (CO<sub>2</sub>, CH<sub>4</sub>) is discussed in a separate report (Ramonet et al., 2021). The CAMS reanalysis is discussed in Inness et al., 2019. More recently, the validation of the reactive gases was documented in the paper by Wagner et al., 2021.

This extended 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 events. The CAMS reanalysis is compared to the previous MACC reanalysis, available for the period 2003-2012, and to a control run without atmospheric composition data assimilation.

In general, the CAMS reanalysis has proven to be very stable compared to independent observations, without major drifts. Changes in the satellite observation datasets entering the data assimilation are leading to small jumps in the bias, as expected. The last reanalysis year 2020 behaves similarly to the previous years, without major changes in performance.

### **Air quality and atmospheric composition**

#### ***Global Aerosol***

The validation (2003-2020) of the CAMS reanalysis aerosol optical depth and Ångström exponent shows a very consistent performance through the period. Two discontinuities can be found. Firstly, the agreement with Aeronet was worse in 2003, possibly because of an underdeveloped Aeronet network. The comparison between Aeronet and the CAMS reanalysis improved after 2003 (see Figure S1). The CAMS reanalysis shows improvements compared to the MACC reanalysis, in particular for AOD bias and correlation, and Ångström correlation. Secondly, after 2012, with the absence of AATSR data in the assimilation, the seasonal cycle in AOD is slightly reduced and a larger bias and FGE develops in North Africa where dust dominates. Taking the old MACC reanalysis as reference, the following changes with respect to aerosol optical depth (AOD) can be found in the CAMS reanalysis: Global reductions for the period 2003-2012 in optical depth are found for total aerosol (-20%), dust (-60%), sea-salt (-40%) and sulphate (-30%), while increases in organic aerosol are simulated (+55%).

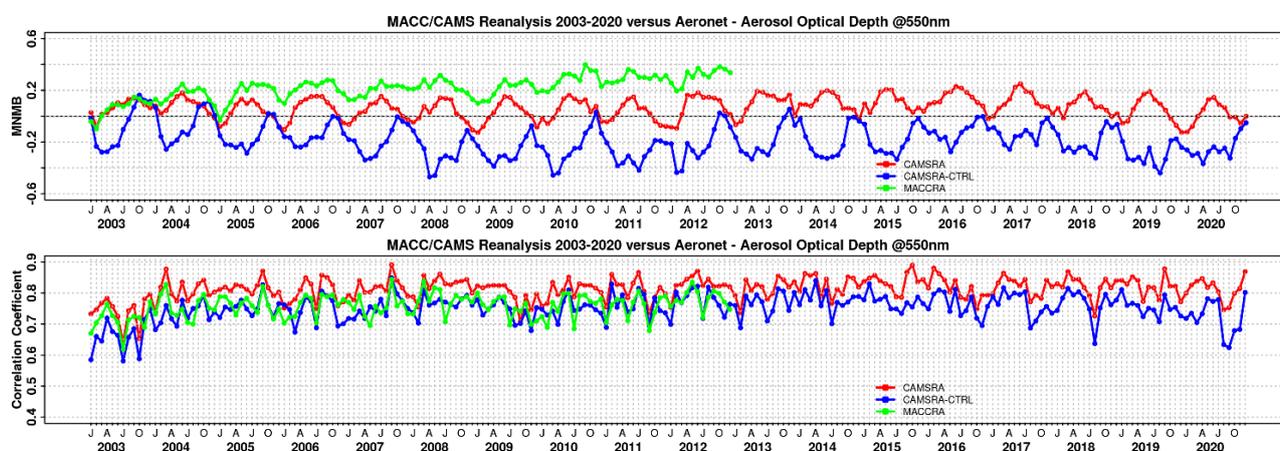


Figure S.1: a) Aerosol optical depth at 550nm in IFS 00Z model simulations for 2003–2020 against daily matching Aeronet Version3 level 1.5; (top) b) Modified normalized mean bias (MNMB); CAMS reanalysis (red) and control run (blue); MACC reanalysis (green); (bottom) Corresponding correlation coefficient.

With this considerable composition change the overall MNMB bias error against daily Aeronet is lower in the CAMS reanalysis. For the CAMS reanalysis the error changes little until 2013, but after 2013 a slight overestimation appears, especially in wintertime. For the MACC reanalysis the mean bias increases from 2004 onwards until 2013. The regional RMS is reduced against MACC in regions such as East Asia, North Africa, India for the MACC reanalysis period 2003–2012. Only in 2007 in North America a higher RMS is found for CAMS reanalysis compared to MACC.

The spatial distribution of the AOD bias has become more evenly distributed in the CAMS reanalysis with a few spots sticking out. Volcanic aerosol hot spots near Hawaii seem to be responsible for high model outliers. The fraction of dust AOD appears to be too small, or the dust size is still too small, showing up as a high bias of the Ångström coefficient in cases of low Ångström coefficient (near deserts). Reduced FGE and RMS and higher correlation of the Ångström coefficient indicate significant improvement in aerosol size from MACC to CAMS reanalysis. Compared to MACC the spatial dust distribution in the CAMS reanalysis appears more consistent with observations (very little dust far from deserts e.g., at the poles and in the western tropical Atlantic) and with a lower Ångström coefficient in dust areas.

### Dust

Over North Africa, Middle East and Europe, the dust optical depth (DOD) fields from the CAMS reanalysis show a distinct seasonal pattern linked to the spatial distribution of dust emissions and transport throughout the seasons along the study period (2003–2020), in good qualitative agreement with ground based (AERONET) and satellite (MODIS and MISR) observations. However, DOD appears underestimated by up to 46% in comparison with AERONET observations. Moreover, the DOD comparison with AERONET observations shows that the reanalysis reproduces the annual variability showing annual correlation coefficients between 0.84 and 0.87 on average for all the AERONET sites.

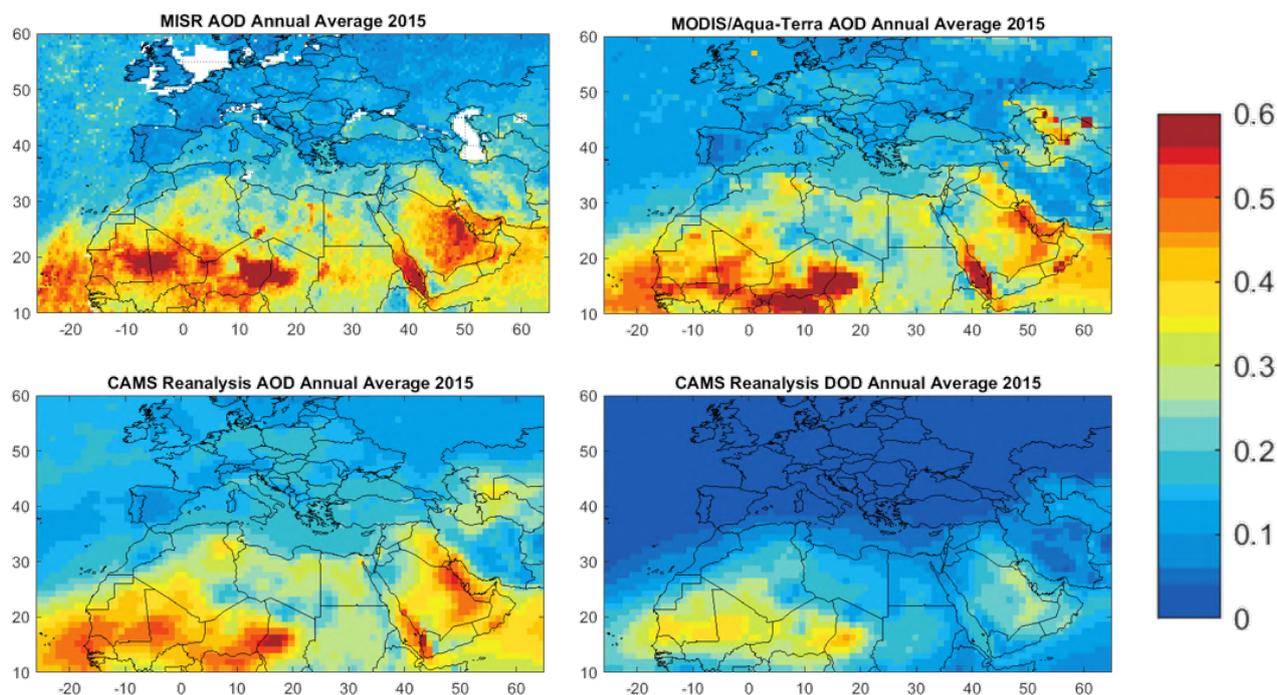


Figure S.2: Seasonal averaged AOD from MISR monthly Level 3 monthly  $0.5^\circ \times 0.5^\circ$  global product (top-left), MODIS Collection 6 Terra/Aqua merged Dark target and Deep Blue Level 3 daily  $1^\circ \times 1^\circ$  global product (top-right), as well as AOD (lower-left) and DOD (lower-right) from CAMS reanalysis at 12UTC for the year 2015.

Differences between AOD and DOD observed over desert dust source regions in North Africa and the Middle East (causing strong DOD underestimations) are linked to an overestimation of organic matter from biomass burning during wintertime and to overestimation of secondary organics over heavily populated areas during summertime. The CAMS AOD is in better agreement with observations than DOD over dust-dominated regions.

The year 2020 has witnessed extreme dust events like over Canary Islands in late-February 2020, and the ‘Godzilla’ dust event in June 2020. The transport and magnitude of these events were well captured by the CAMS reanalysis.

### ***Tropospheric ozone ( $O_3$ )***

For the free troposphere (Fig. S.3) comparisons between the CAMS reanalysis and GAW, ozone sonde, and MOZAIC-IAGOS observations show a good agreement, with a mean bias within  $\pm 10\%$  at midlatitudes. From 2013 onwards, MNMBs show a more pronounced seasonal behaviour, with negative values during winter and positive values in summer. This is due to a change in the assimilated SBUV/2 data. Before July 2013 a 13-layer product was assimilated, after 9 July 2013 a 21-layer product (see Table 2.2). Unfortunately, this change had an impact on the ozone bias in the troposphere. The same behaviour in tropospheric ozone is seen in the CAMS near-real-time analysis, which also assimilates the 21L data from May 2013 onwards.

In 2003 and 2004, for mid- and high-latitudes, negative MNMBs appear (Fig.S.3). After 2004, MNMBs remain within  $\pm 15\%$  for the free Antarctic troposphere, until September 2012. Thereafter

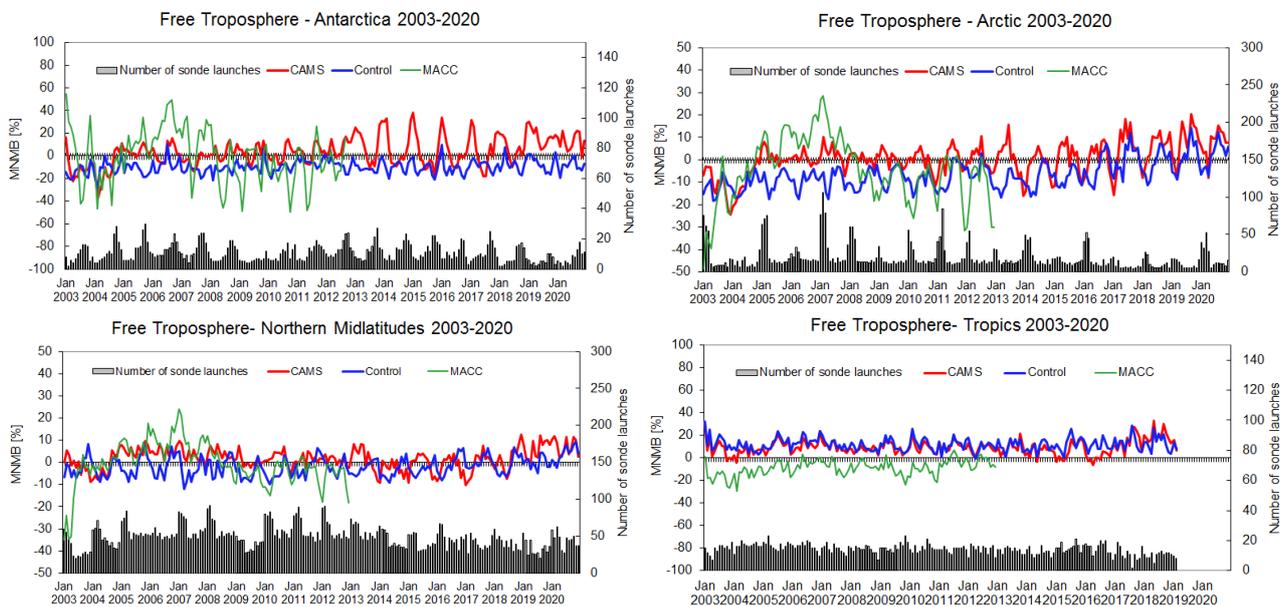


Figure S.3: Comparison with ozone sondes in the free troposphere, period 2003-2020: MNMBs for 4 regions (first row left: Antarctica, first row right: Arctic, second row left: Northern Midlatitudes, second row right: Tropics) in the free troposphere. Red: CAMS reanalysis, green: MACC reanalysis, blue: Control run.

MNMBs get larger up to  $\pm 38\%$ , with a more pronounced seasonal bias. A similar behaviour can be seen for the Arctic, however, MNMBs remain  $\pm 20\%$ . In tropical regions, ozone is overestimated by up to 33% in more recent years.

MNMBs for surface ozone for GAW stations in the Northern midlatitudes range between  $-38\%$  and  $26\%$ . For Antarctica, MNMBs are within  $\pm 44\%$ , respectively (except for 2004). Larger MNMBs (up to  $75\%$ ) appear for Arctic stations during spring (ozone depletion events). For the Tropics, monthly MNMBs are within  $-6$  and  $51\%$ . MNMBs for stations in Antarctica and the Arctic show inconsistencies in the time-series from 2012 onwards for the CAMS reanalysis due to the loss of ENVISAT data and the switch in SBUV/2 data. The control run remains stable, except during DJF towards the end of the period due to a limited number of stations in recent years. Starting from the year 2015, both model runs show a smaller bias and smaller seasonality of the bias compared to the observations in the northern midlatitudes. For the Antarctic, from the year 2019, the discrepancy between the reanalysis and control run increased. In the year 2020, the control run shows an increase of the negative bias in the spring season.

The CAMS reanalysis, the control run and the MACC reanalysis were all compared to the surface EMEP and AirBase European background rural observations at the latitudinal zones  $30^{\circ}\text{N}$ - $40^{\circ}\text{N}$ ,  $40^{\circ}\text{N}$ - $50^{\circ}\text{N}$  and  $50^{\circ}\text{N}$ - $70^{\circ}\text{N}$  as well as using only the high altitude stations at elevations exceeding 1000 m above mean sea level for the period 2003-2020 (Fig. S.4). The description of surface ozone over Northern and central Europe has improved considerably compared with the MACC reanalysis. The bias has improved, especially between June to September. Additionally, data assimilation reduces the negative offset seen in the control run during January to April by about 50%. At high Northern latitudes in Europe a clear winter-spring negative bias is found in both the MACC and CAMS reanalysis. A positive trend in bias over Europe during the period 2003-2007, observed in

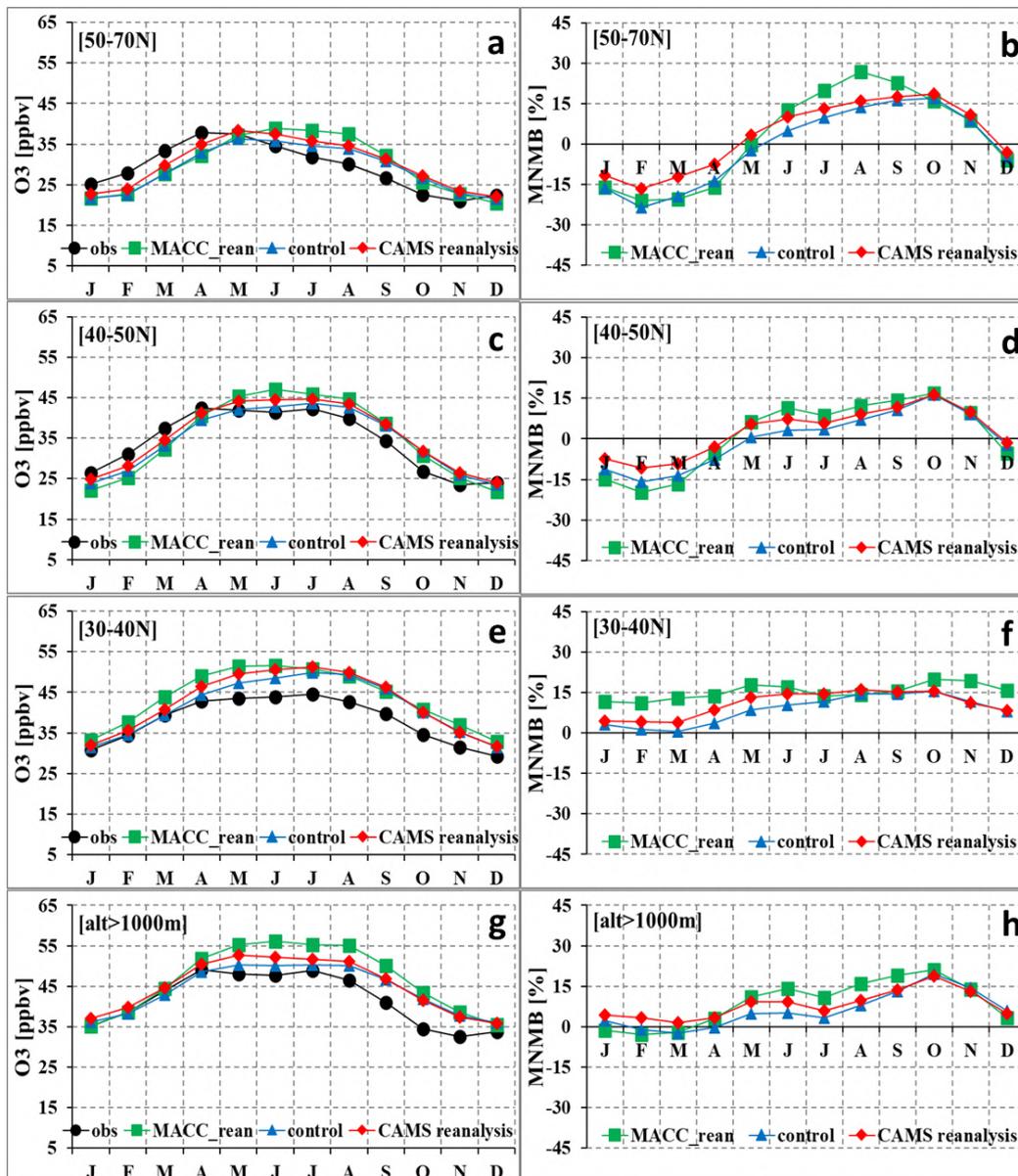


Figure S.4: Mean monthly ozone values for the period 2003-2020 (left) and the MNMBs (right) of the CAMS reanalysis (red robs), the associated CAMS control run (blue triangles), the MACC reanalysis (green squares), and the EMEP observations (black circles) over Northern Europe (1st row, a and b), Central Europe (2<sup>nd</sup> row, c and d), Southern Europe (3<sup>rd</sup> row, e and f) as well as for stations with altitude greater than 1000m above sea level (4<sup>th</sup> row, g and h).

MACC are not observed in the new CAMS reanalysis and time series have improved compared to MACC, although the first two years show a difference of about 10% in bias which may reflect the use of Aura observations after the end of 2004. After 2008 the performance of CAMS and MACC reanalysis are almost equal in terms of annual MNMBs. However, the CAMS reanalysis shows a lower seasonal dependence of MNMBs (this is an improvement against the MACC reanalysis). It should be noted also that the CAMS reanalysis reproduces better the monthly surface ozone variability over all three European latitudinal belts, especially when the seasonal cycle is removed from the time-series.

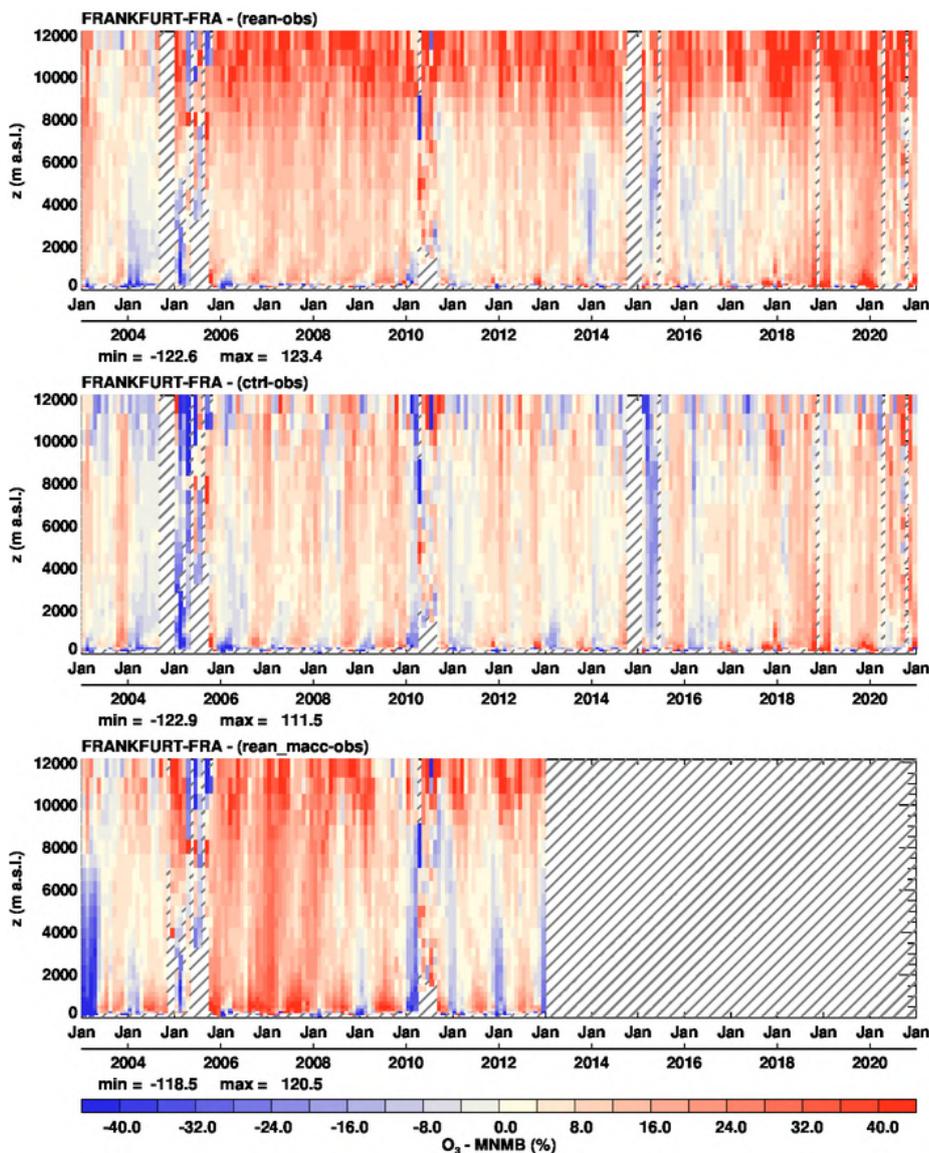


Figure S.5: Time series of the monthly scores (MNMB) for ozone at Frankfurt during the period 2003-2020 by comparing with MOZAIC/IAGOS aircraft landing and take-off profiles. From top to bottom panel, the scores are for: CAMS reanalysis, control run and MACC reanalysis. Units: %.

The CAMS reanalysis is compared with MOZAIC-IAGOS take-off and landing profiles, see e.g. Fig. S.5. For the lower troposphere, in Europe the scores show a seasonal pattern with a lower bias in winter and spring, and overestimation in summer and fall. Before 2007 winter underestimations close to the surface are found (reaching MNMB of -40%) and MNMB start to be slightly positive afterwards in the CAMS reanalysis. In North America, India and East Asia both positive and negative biases are found, with sometimes a pronounced seasonality, depending on the airport.

In the free troposphere, after 2004, the CAMS reanalysis is generally more stable than the MACC reanalysis, which developed a positive bias, especially over Europe, for the years 2006-2008. Overall, the biases are improved in the CAMS reanalysis compared to the MACC reanalysis. We find mainly positive biases in the Northern hemisphere free troposphere of 10-20% of CAMS compared

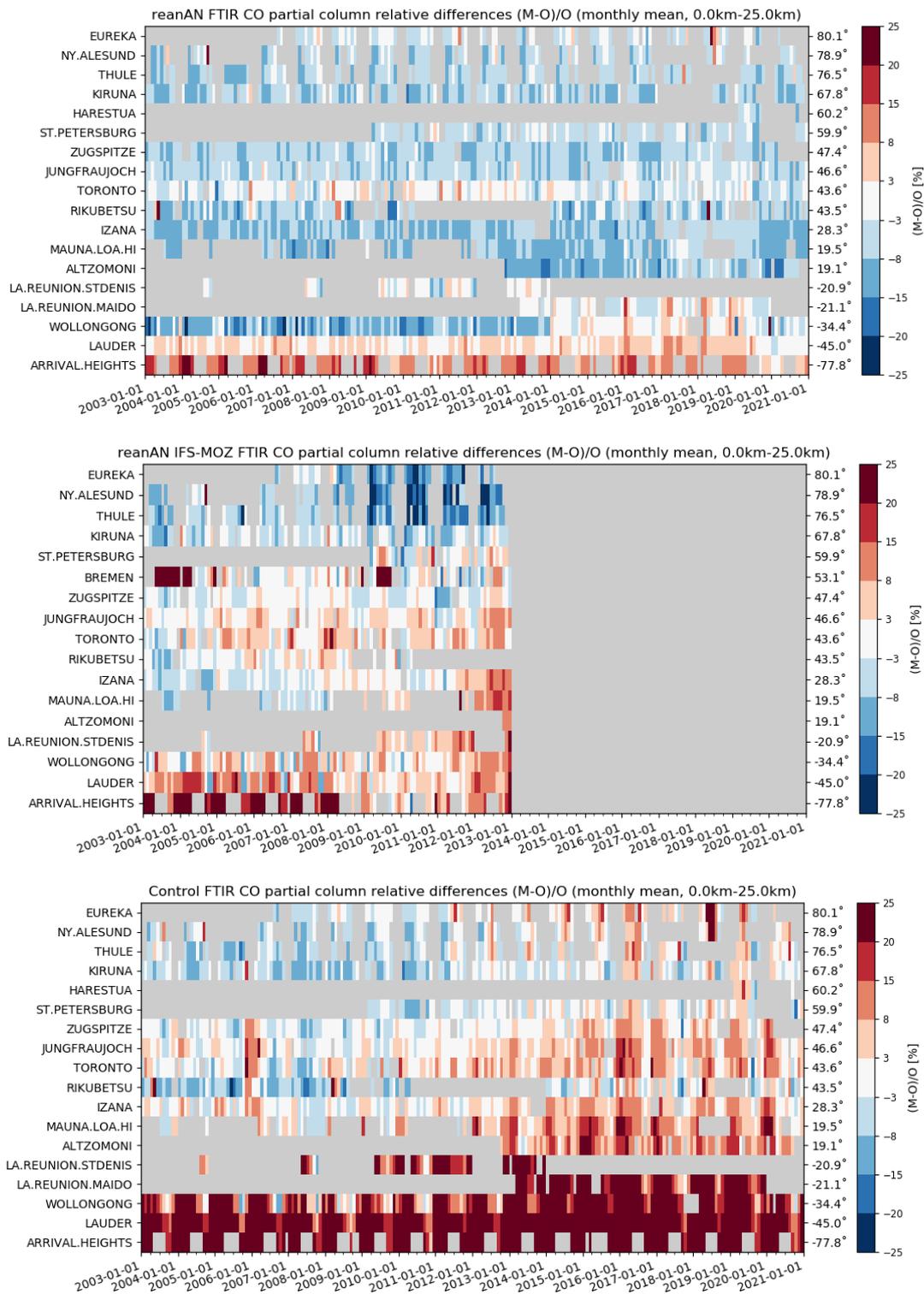


Figure S.6: Mosaic plot of seasonal relative biases in CO ( $[\text{reanalysis-observation}]/\text{observation}$ ) at all NDACC FTIR sites for the period 2003-2020. The CAMS reanalysis (bottom) is more stable compared to both the MACC reanalysis (middle) and CAMS control run (bottom). A slight negative bias is found in the NH for the CAMS reanalysis. At the tropical sites in the northern hemisphere the negative bias reaches values up to 10% (Izana, Mauna Loa, Altzomoni). The control run shows a strong positive bias in the southern hemisphere efficiently corrected by the assimilation.



to the aircraft observations with some variability depending on the airport location. The CAMS values are often higher than for the MACC reanalysis. In India, the seasonality driven by the monsoon regime is well reproduced but there are some differences in behaviour between CAMS and MACC, with again positive, but rather uniform biases in the CAMS reanalysis.

In the southern hemisphere, at the southwest African airport of Windhoek, the seasonality is generally well reproduced. In the lower troposphere, the agreement is good during the dry season with slightly positive MNMB in the range of  $\pm 16\%$ , and worse during the wet season with significant underestimations (MNMB between  $-40\%$  and  $-20\%$ ). The CAMS reanalysis shows an increasing positive bias with altitude, different from the MACC reanalysis.

In the upper troposphere and lower stratosphere (UTLS), prior to 2005, MNMB values are small within  $\pm 15\%$  on average. From 2005 on, UTLS ozone is mostly overestimated by both CAMS and MACC in the Northern Hemisphere, but more so by CAMS. In Windhoek large differences are found between the two reanalyses in the UTLS, where ozone is generally overestimated by CAMS and underestimated by MACC, with no clear difference in the overall performance in the upper layers.

CAMS reanalysis  $O_3$  total column data have been compared with IASI satellite observations. The reanalysis  $O_3$  total columns are in good agreement with the satellite data, showing general biases within 5%. The control run shows an overestimation over the Northern Hemisphere high- and midlatitudes up to 20%, and an underestimation over other parts of the Globe (within 10%) with some regional and temporal exceptions, which can reach 20% over the Southern Hemisphere high- and midlatitudes during the autumn season.

### ***Tropospheric Carbon Monoxide (CO)***

Overall, a good agreement is found between the CAMS reanalysis and independent CO observations.

CO surface mixing ratios are slightly underestimated for European and Asian GAW stations with mean monthly MNMBs within  $\pm 18\%$  to  $\pm 20\%$  for Asia and Europe, respectively. For North America, surface CO is mostly overestimated with mean monthly MNMBs of up to 30%. Correlation coefficients for European stations are mostly around 0.6, for North American stations around 0.7, and for stations in Asia around 0.8. Starting from the year 2014 the difference between the reanalysis and control runs increased, especially in Asia.

For comparisons against NDACC and TCCON FTIR observations the CAMS reanalysis captures well the day-to-day variations of the tropospheric columns of CO and is more stable than the CAMS control run and the MACC reanalysis (Fig. S.6). The relative difference of the control run shows a strong latitudinal dependence in the southern hemisphere and a positive trend for the stations in the northern hemisphere. The correlation coefficients of the CAMS reanalysis have also improved with respect to the former reanalysis, with a coefficient value of 0.9 for most stations against 0.8 for MACC. In the northern hemisphere CAMS reanalysis underestimates the CO abundance by 5% which increases towards the tropics where the model underestimates the CO abundance by values up to 10%. In the southern hemisphere, the bias of the CAMS reanalysis at Reunion is negligible. It should be noted that the CAMS reanalysis shows a bias at the sites in the northern tropics well above the reported measurement uncertainty. As the three locations concerned are mountain sites,

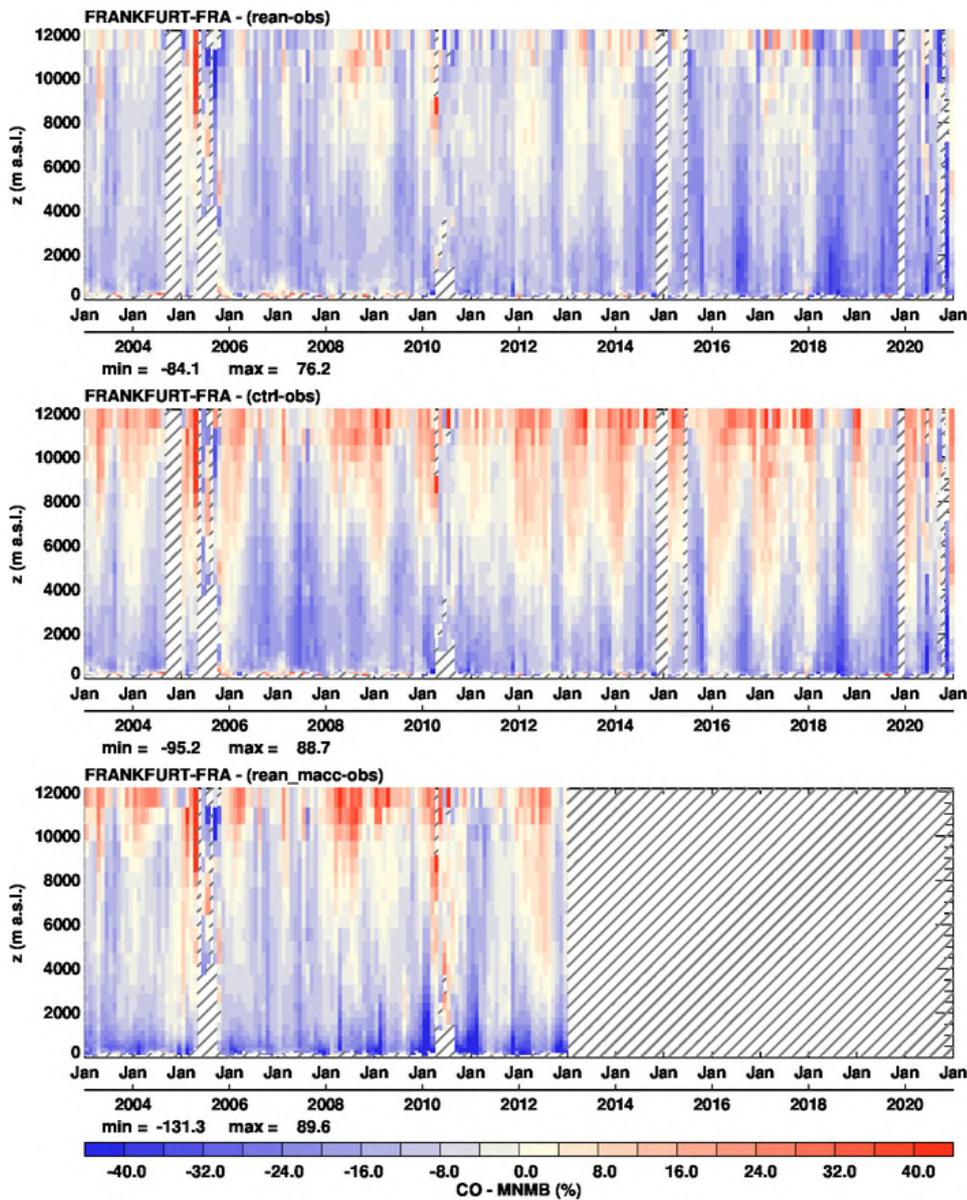


Figure S.7: Time series of the monthly scores (MNMB) for CO at Frankfurt during the period 2003-2020 by comparing with MOZAIC/IAGOS aircraft landing and take-off profiles. From top to bottom panel, the scores are for: CAMS reanalysis, control run and MACC reanalysis. Units: %.

co-location issues might partly explain this issue. Meanwhile, this bias was not observed in the MACC reanalysis. Changes in the evolution of the bias of the CAMS reanalysis can also be related to changes in the assimilated MOPITT CO product.

In the comparisons with aircraft observations from MOZAIC-IAGOS, the results obtained with the CAMS reanalysis give a similar picture at most airports, with Frankfurt (Fig. S.7) as example. In general, the shapes of the CO profiles are in good agreement with observations, without a clear seasonal bias variability, but the concentrations are in general somewhat smaller than observed. The best results are generally found in the middle troposphere. In most airports the CAMS reanalysis is an improvement over MACC, showing less seasonal variation and smaller biases, see

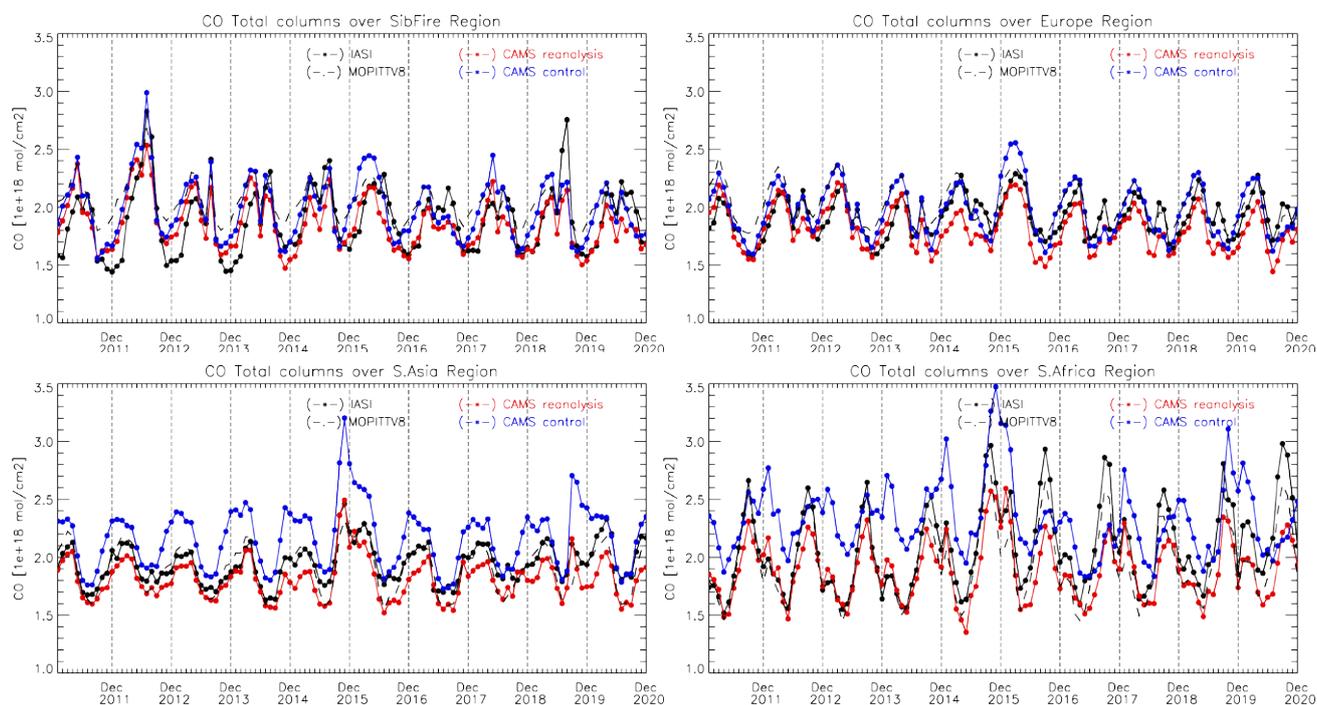


Figure S.8: CO total columns for satellite retrievals from IASI on Metop-B and MOPITT V8 (black) and reanalysis data (red) and control run (blue) over Europe (top left), the Siberian fire (top right), South Asia (bottom left) and South Africa (bottom right) regions for the years from 2011 to 2020. The modelled CO concentrations were transformed using averaging kernels IASI.

for instance Fig. S.7. Usually, the largest bias is observed in the low troposphere, in the range of -15% to -25% (Frankfurt) up to -40% in Shenyang in China. Some locations also present slight overestimations, e.g., Vancouver, in particular during winter months. At Hyderabad in India, for both reanalyses, MNMB exhibits a clear seasonal feature in the lowest layers, with overestimations in the winter monsoon with larger bias for CAMS than MACC (frequent MNMB greater than 40%) and underestimations during the summer monsoon (FGE<30%).

At the south-west African airport of Windhoek, the results from MACC show CO overestimations in the lowest layers during the dry season, and large underestimation from the mid-troposphere to UTLS, especially before 2008. The CAMS reanalysis performs significantly better at this airport.

In the UTLS, at the European airports, the CAMS reanalysis usually performs better than MACC. At the other locations the results from the two reanalyses usually differ less.

CO surface mixing ratios are slightly underestimated for European and Asian GAW stations (median MNMBs -6% for European Stations, -4% for Asian stations), whereas for North American stations an overestimation of surface CO with MNMBs of around 7% (North America) is observed. Correlation coefficients for European stations are mostly greater than 0.7, for North American stations mostly greater than 0.6, and for stations in Asia greater than 0.86.

At the northern hemisphere IAGOS airports the reanalysis and control run results differ mainly from the mid-troposphere to the UTLS (e.g., Fig. S.7). In the mid-troposphere the reanalysis mostly underestimates independently from the season while the bias from control run almost always

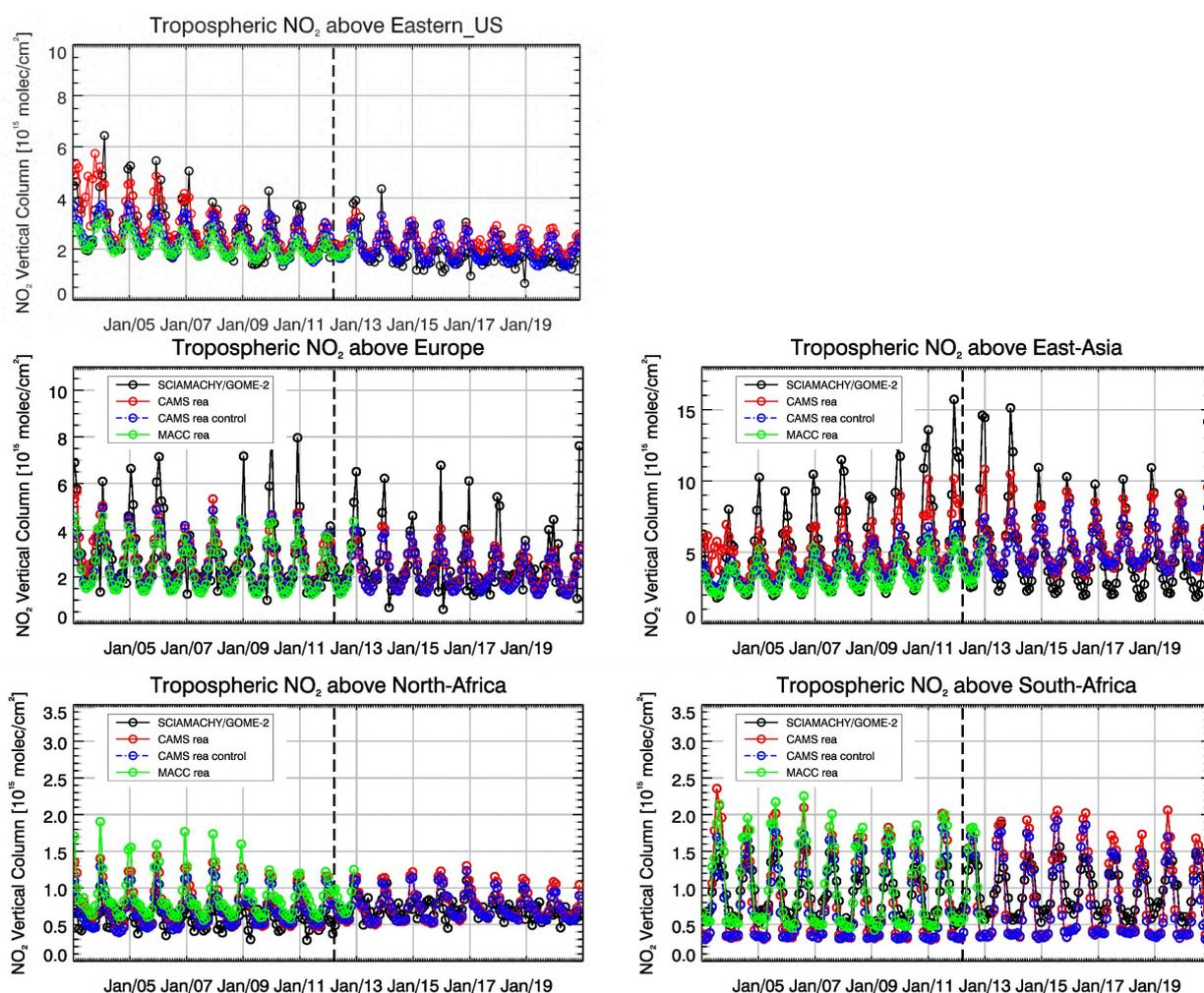


Figure S.9: Comparison of time series of tropospheric  $\text{NO}_2$  columns from SCIAMACHY (up to April 2012) and GOME-2A (from April 2012 onwards) to model results over selected regions. The switch from SCIAMACHY to GOME-2A is indicated by the vertical black dashed lines. Upper panels represent regions dominated by anthropogenic emissions; lower panels represent those dominated by biomass burning. The unusually high satellite values for December 2020 over East-Asia and Europe are possibly caused by a problem with the satellite retrievals for this month and are still under investigation.

presents a seasonal behaviour. In the UTLS, the control run overestimates systematically. At some airports the overall behaviour of the control run is to overestimates from the mid-troposphere to the UTLS (Madras, Taipei, Windhoek).

In the last years 2018-2020, the bias from the CAMS reanalysis seems to be degraded in the comparisons with IAGOS as underestimations become larger in the troposphere. This increase in the bias is not present in control run comparisons. Note that IAGOS data for 2018-2019-are Level 2 like the years before, and 2020 data is Level 1.

The CAMS reanalysis  $\text{CO}$  total column data have been compared with MOPITT version 8 and IASI on Metop-B satellite observations. The reanalysis  $\text{CO}$  total columns are in good agreement with the MOPITT data, showing the efficiency of the assimilation, and have a slight negative bias within 10%

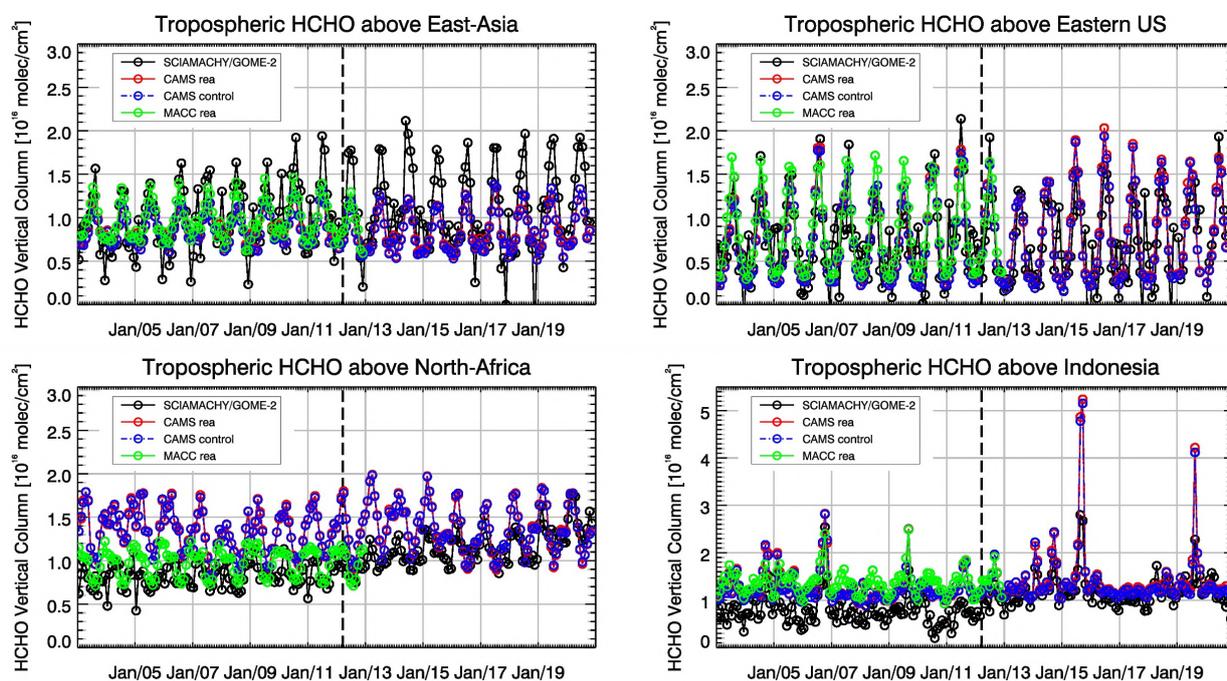


Figure S.10 Comparison of time series of tropospheric HCHO columns from SCIAMACHY (up to April 2012) and GOME-2A (from April 2012 onwards) to model results over selected regions for the years 2003-2020. The switch from SCIAMACHY to GOME-2A is indicated by the vertical black dashed lines. The regions differ from those used for NO<sub>2</sub> to better focus on HCHO hotspots: East Asia (25-40°N, 110-125°E), Eastern US (30-40°N, 75-90°W), Northern Africa (0-15°N, 15°W-25°E) and Indonesia (5°S-5°N, 100-120°E). Negative satellite retrieved values over Eastern US and East-Asia are due to the limited amount of data during Northern Hemisphere winter months for these regions.

with some exceptions. Starting from the year 2012 the negative bias increased over the low latitudes (up to 20%). The year 2015 was marked by a strong El Niño event which led to a large amount of fire emissions in Indonesia. In Autumn/winter 2015 the reanalysis shows an overestimation of satellite observations over the Southern Hemisphere (SH) low latitudes up to 20%. Summer 2012 was characterized by a large number of fire events in Siberia (Fig. S.8). Both runs reproduced the enhanced CO total columns observed by the satellites. Starting from autumn 2016, the CAMS reanalysis shows reduced negative bias (within 5%) over the globe and an improved agreement with the satellite observations. On the other hand, a slight positive bias (within 5%) appears over the SH in spring/summer season. In the first half of year 2019 the CAMS reanalysis shows a slightly higher positive bias over the SH compared to year 2018. The second half of 2019 shows improvements compared to the previous year which is reflected in a weaker bias over the SH. Year 2020 shows a change in the sign of the bias over the Southern Hemisphere from positive to negative, but within 5%. We note that MOPITT V7 data were assimilated in the CAMS reanalysis from 20170101 onwards, while before that time it was V6. In contrast, a slight positive bias (within 5%) appears over the Southern Hemisphere in spring/summer seasons.

Fire case studies in 2012, 2015, 2017 and 2020 show the skill of the CAMS model to analyse major fire events and describe the details of the transport. Often, the modelled CO values are somewhat



lower than satellite measurements, and it is observed that the further away from the fire, the more the model underestimates the CO concentrations.

### ***Tropospheric Nitrogen dioxide (NO<sub>2</sub>)***

In the comparison with SCIAMACHY and GOME-2A satellite retrievals, the CAMS reanalysis performs reasonably well regarding magnitude and seasonality, with the exception of East-Asia in 2003 due to the assimilation of NO<sub>2</sub> satellite data with degraded data quality for this year (Fig. S.9). The seasonal cycle is underestimated over East Asia and partly over Europe due to an underestimation of wintertime values. With regards to CAMS operational forecast service, similar features are observed i.e. stronger shipping signals and overestimation of boreal forest fire emissions (same for tropospheric HCHO in some cases, indicating a different performance probably related to fire emission factors depending on the trace gas, region and season), overestimation of values over the Red Sea and Persian Gulf and a tendency to underestimate values over pollution hotspots of Central Europe such as Benelux countries, while other distinct hotspots are overestimated (e.g. Moscow).

### ***Formaldehyde***

Formaldehyde (HCHO) columns over East-Asia and Eastern-US show a good agreement with SCIAMACHY and GOME-2A satellite observations. Values over North Africa and Indonesia are overestimated by up to a factor of 2, while the seasonal cycle over East-Asia is underestimated compared with observations (Fig. S.10).

### ***System performance in the Arctic***

The simulated surface O<sub>3</sub> mixing ratios in the Arctic are on average in good agreement with the observations apart from spring ozone depletion events at high-Arctic coastal stations, related to halogen chemistry reactions that are not represented in the model simulations. The levels and variability are well captured in the other seasons for the high-Arctic sites, while the overall correlations are relatively low ( $r = 0.09 - 0.41$ ) due to the pronounced ozone depletion events. The reanalysis tends to underestimate the measured levels in late winter and spring and to overestimate them in summer and autumn at the European Arctic sites. The annual average O<sub>3</sub> mixing ratios are well captured (MNMB= -1% - 11%). The model generally captures high concentration episodes in spring and summer although the peaks are underestimated. The performance of the model to predict O<sub>3</sub> concentrations over the Arctic tends to deteriorate with decreasing correlation and increasing bias over the first half of the simulated period.

Ozone sondes from the free troposphere in the Arctic show a strong negative bias for the first two years of the period (up to -24%). In the years 2005 to 2012 MNMBs are  $\pm 10\%$  for the Arctic and from 2013 onwards, MNMBs get larger, reaching up to  $\pm 25\%$  over Arctic stations. This change in bias relates to the switch from SBUV/2 13L to SBUV/2 21L data. The simulated O<sub>3</sub> concentrations pattern is the same in the UTLS.

The CO mixing ratios over the Arctic (station Alert) are reproduced accurately between 2003 and 2010 (afterwards observations are missing). The winter maxima are slightly underestimated, however. MNMBs for ALT range from -10% to 1%. Correlation is high between 0.75 and 0.97.

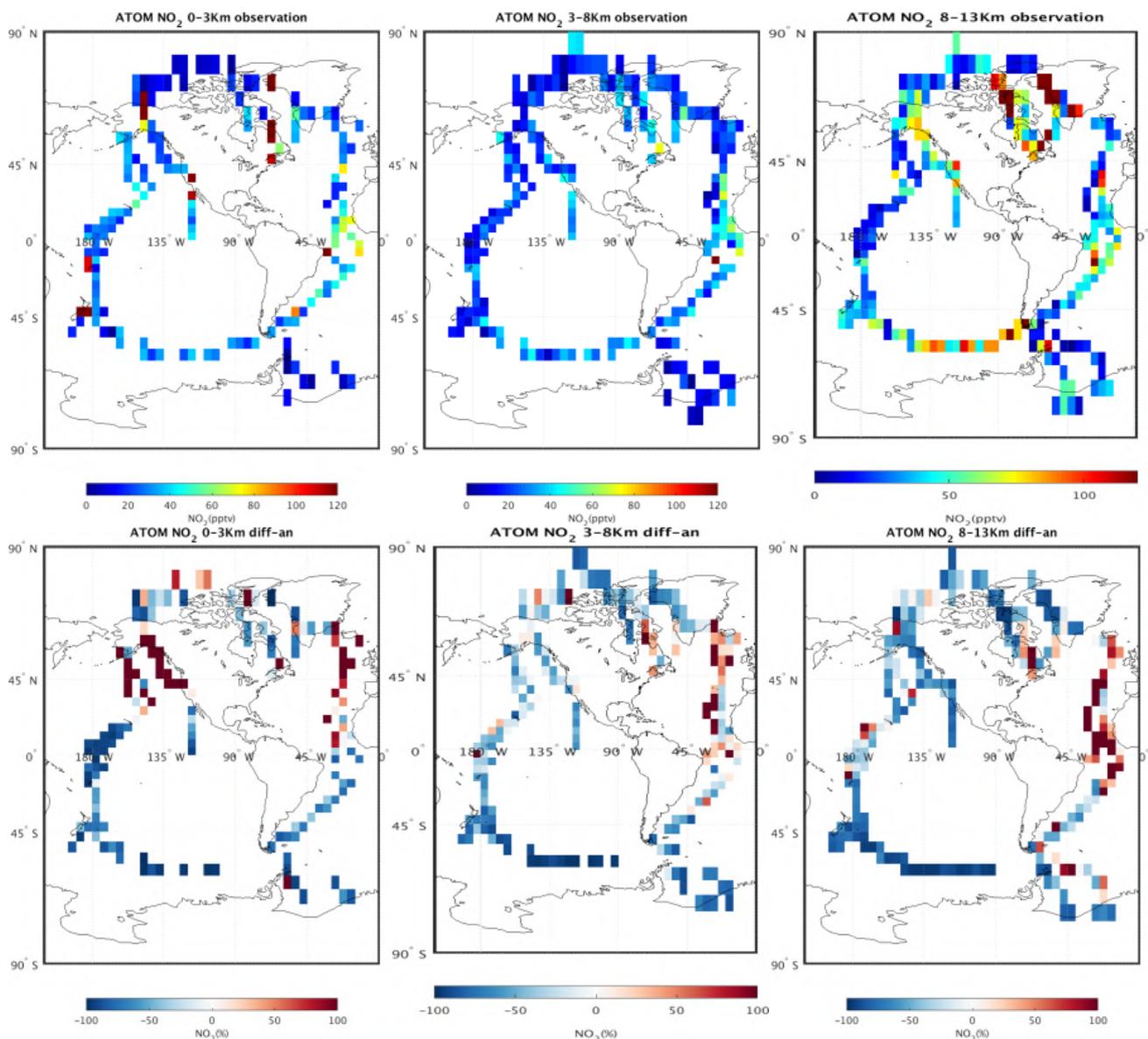


Figure S.11: ATOM campaign observations of NO<sub>2</sub> in pptv (top, see chapter 7), and the relative difference between the reanalysis and the observation (in %; difference = reanalysis – observation; bottom). Data included are from ATOM-1 to ATOM-4, between July 2016 and May 2018. The data are averaged to 5°x5° (latitude x longitude) and to three altitude bins: 0-3 km (left), 3-9 km (middle), and 9-15 km (right).

### System performance in the Mediterranean

Over the Mediterranean, the CAMS reanalysis can reproduce the Saharan long-range transport of dust with annual correlation coefficients (for 3-hourly values) between 0.60 and 0.84 for the Mediterranean in the AOD AERONET comparison. The reanalysis tends to reproduce the AOD observations in Western and Central Mediterranean sites well (with annual mean bias between -0.01 and 0.01). The reanalysis overestimates them in Eastern Mediterranean sites (with annual mean bias between 0 and 0.03).

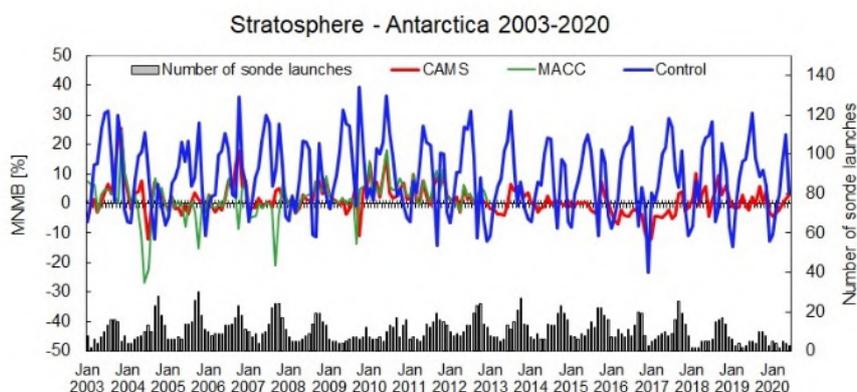


Fig. S.12: Modified Normalised Mean Bias of the CAMS reanalysis (red), its control run (blue) and the MACC reanalysis versus ozone sondes, for the Antarctic. The biases are averaged between 90 and 10hPa.

### **Comparison with aircraft campaigns**

The CAMS reanalysis and control runs are compared with aircraft measurements of O<sub>3</sub>, CO, HCHO, NO, NO<sub>2</sub>, OH, and HO<sub>2</sub>, from several campaigns between 2003 and 2018. All the campaigns are combined and averaged to grids to evaluate the spatial pattern. Additionally, the HIPPO campaigns are used to look at more details of O<sub>3</sub> and CO over the Pacific Ocean in different seasons. The evaluation of reanalysis O<sub>3</sub> is generally good with a calculated correlation coefficient squared of 0.94. The reanalysis mostly overestimates O<sub>3</sub> compared to the observations. The calculated squared correlation coefficient for the control run is 0.90. The reanalysis of CO agrees with the measurements well with a squared correlation coefficient of 0.71, and the biases of the averaged grids are mostly within 15%. For HCHO the reanalysis underestimates the observations by about 30% - 60% with a correlation of 0.53. The comparison of NO<sub>x</sub> is worse than for the other species, and the mean bias of NO<sub>x</sub> for the averaged grids is about 70%. The comparisons of OH and HO<sub>2</sub> are also reasonable with correlations of 0.44 and 0.41, respectively.

## **Ozone layer and UV**

### **Ozone partial columns and vertical profiles**

The CAMS reanalysis stratospheric ozone results have been evaluated against observations from ozonesondes, ground-based remote-sensing observations from NDACC (the Network for the Detection of Atmospheric Composition Change, <http://www.ndacc.org>) and satellite observations by several limb-profiling instruments. The performance of the reanalysis was also compared with those of the MACC reanalysis and the control run.

A good agreement, stable over time, is generally found between the CAMS reanalysis and independent observations, e.g., within  $\pm 10\%$  against ozonesondes between 10 and 90 hPa. Comparisons against limb profiles are also good in the middle stratosphere [4-100 hPa at mid-latitudes and poles, 4-70 hPa in the Tropics] where the mean bias is within  $\pm 10\%$ , its associated standard deviation is lower than 10% and with correlation often larger (better) than 0.8. In the upper troposphere lower stratosphere (between 70 to 200 hPa), comparisons degrade, especially in

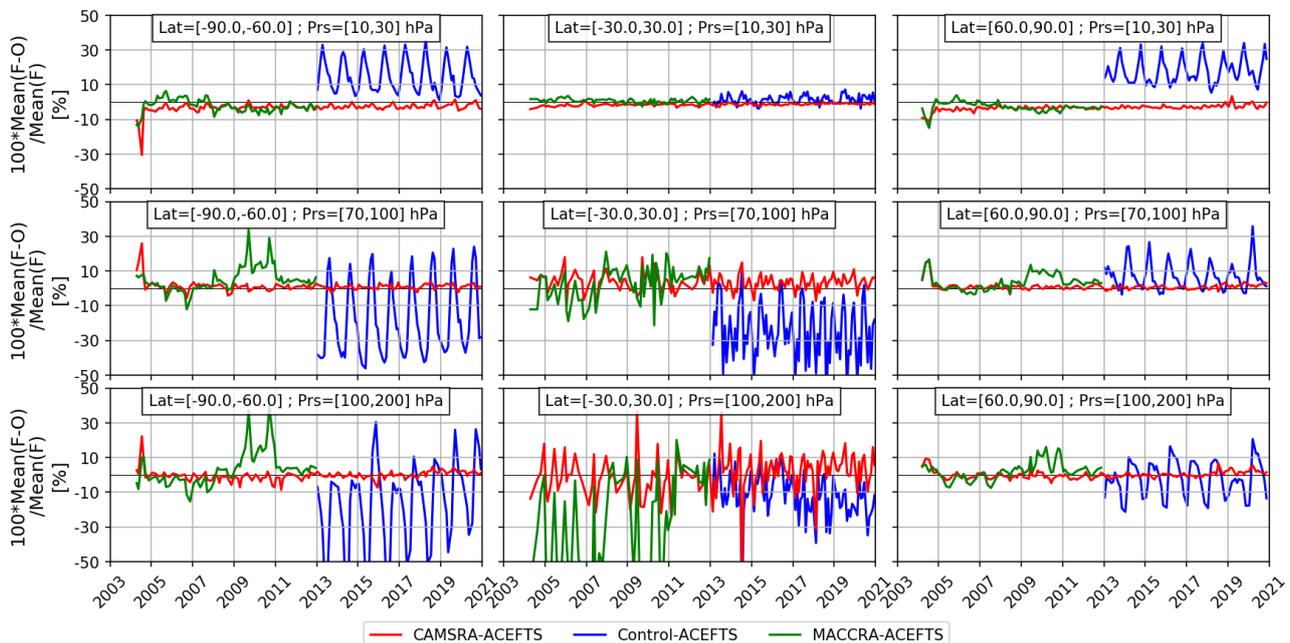


Fig. S.13: Time series of monthly normalized mean bias (NMB i.e.  $\text{Mean}(\text{Model-Obs})/\text{Mean}(\text{Model})$ ) between ACE-FTS ozone profiles and model data from CAMS reanalysis (CAMSRA, red line), its control run (blue line) and the MACC reanalysis (MACCRA, green line) in three pressure layers in the middle and lower stratosphere (from top to bottom: 10-30 hPa, 70-100 hPa and 100-200 hPa) and in three latitude bands (from left to right: 90°S-60°S, 30°S-30°N and 60°N-90°N).

the Tropics where the standard deviation of the differences can be larger than 50 % and with correlation lower than 0.7. Nevertheless, mean biases are often as good as  $\pm 10\%$ . In the upper stratosphere (pressure lower than 4 hPa), the CAMS reanalysis shows a negative bias (typically between -10 % and -20 %) and comparison with the performance of the control run reveal minor constraints from the assimilated observations. Compared to MACC, the CAMS reanalysis is much more stable over time such that users are recommended to switch to the CAMS reanalysis. Finally, time series of the performances of the reanalysis against limb profiles reveal some discontinuities associated with discontinuities in the use of the assimilated  $\text{O}_3$  profiles. Users interested to derive trend studies with CAMS-RA would need to evaluate if these discontinuities are acceptable for their work.

### **Other stratospheric trace gases**

Due to the lack of stratospheric chemistry in the IFS-CB05 scheme, the only useful product in the stratosphere is ozone. Other species, like  $\text{NO}_2$ , have also been evaluated but the results are only indicative.



## Table of Contents

<b>Summary</b>	<b>4</b>
<b>Air quality and atmospheric composition</b>	<b>4</b>
<b>Ozone layer and UV</b>	<b>18</b>
<b>1. Introduction</b>	<b>22</b>
<b>2. System summary and model background information</b>	<b>24</b>
<b>2.1 System based on the ECMWF IFS model</b>	<b>24</b>
2.1.1 CAMS reanalysis system	24
2.1.2 Control run	27
<b>2.2 Other systems</b>	<b>27</b>
2.2.1 The MACC reanalysis and CAMS forecasts	27
2.2.2 BASCOE	27
2.2.3 TM3DAM and the multi-sensor reanalysis	28
2.2.4 SDS-WAS multimodel ensemble	28
<b>2.3 CAMS reanalysis products</b>	<b>29</b>
<b>2.4 Validation results for reactive gases and aerosol</b>	<b>29</b>
<b>3. Aerosol</b>	<b>30</b>
<b>3.1 Global aerosol distribution</b>	<b>30</b>
<b>3.2 Dust evaluation over North Africa, Middle East and Europe</b>	<b>36</b>
<b>3.3 Aerosol validation over the Mediterranean</b>	<b>43</b>
<b>3.4 Exceptional dust event over the Canary Islands in late-February 2020</b>	<b>46</b>
<b>3.5 ‘Godzilla’ dust event in June 2020</b>	<b>47</b>
<b>4. Tropospheric Ozone</b>	<b>50</b>
<b>4.1 Validation with sonde data in the free troposphere</b>	<b>50</b>
<b>4.2 Validation with GAW surface ozone observations</b>	<b>57</b>
<b>4.3 Verification with European EMEP surface ozone observations</b>	<b>65</b>
4.3.1 Seasonal and Multi annual comparison	65
4.3.2 Comparison of model simulations during the extreme heatwave event in August 2003	70
<b>4.4 Verification with IAGOS ozone observations</b>	<b>74</b>
<b>4.5 Verification with observations in the Arctic</b>	<b>92</b>
<b>4.6 Comparison of the CAMS reanalysis with IASI O3</b>	<b>98</b>
<b>5. Carbon monoxide</b>	<b>100</b>
<b>5.1 Validation with Global Atmosphere Watch (GAW) Surface Observations</b>	<b>100</b>
<b>5.2 IAGOS Aircraft observations</b>	<b>106</b>
<b>5.3 Validation against FTIR CO observations from the NDACC network</b>	<b>125</b>
<b>5.4 Validation against FTIR CO observations from the TCCON network</b>	<b>130</b>
<b>5.5 Comparison of the CAMS reanalysis with MOPITT CO</b>	<b>137</b>



<b>6. Tropospheric nitrogen dioxide and formaldehyde</b>	<b>147</b>
6.1 Evaluation against SCIAMACHY and GOME-2 NO <sub>2</sub> retrievals	147
6.2 Validation against SCIAMACHY HCHO satellite data	149
<b>7. Comparisons with aircraft campaigns</b>	<b>152</b>
7.1 Campaign description	152
7.2 Comparison with several campaigns including HIPPO	154
7.3 Comparison with the ATOM campaign	167
<b>8. Stratosphere</b>	<b>175</b>
8.1 Validation against ozone sondes	175
8.2 Validation against ozone observations from the NDACC network (MWR, LIDAR)	182
8.3 Comparison with observations by limb-scanning satellites	186
8.4 Stratospheric NO <sub>2</sub>	195
<b>9. References</b>	<b>196</b>
<b>Annex 1: Acknowledgements</b>	<b>201</b>



## 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. Apart from these daily analyses, CAMS will produce a global reanalysis covering 19 years (2003-2021). The CAMS system was originally developed by a series of MACC research projects (MACC I-II-III) until it became operational in 2015, and in this report, we show comparisons between the CAMS reanalysis and the MACC reanalysis. The CAMS near-real time and reanalysis services consist of daily analysis and forecasts with the ECMWF IFS system with modelling and data assimilation of trace gas concentrations and aerosol properties. A second component of CAMS consists of the provision of air-quality forecasts and reanalyses over Europe, based on an ensemble of European air quality models.

This document presents the validation of the global CAMS reanalysis during production (up to 31 December 2020). The validation methodology and measurement datasets are discussed in Eskes et al. (2015) and Eskes et al. (2021). 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 events are documented. Table 1.1 provides an overview of the trace gas species and aerosol aspects discussed in this CAMS reanalysis validation report. The reanalysis results are compared with results for a free model run without assimilation, to document the improvements by using the (satellite) observations.

Key CAMS 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); 3D ozone fields (e.g. SPARC).

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

CAMS data are made available to users as data products (grib or netcdf files) and graphical products from the CAMS Atmosphere Data Store, <https://ads.atmosphere.copernicus.eu> and from the CAMS website, <http://atmosphere.copernicus.eu/>. The (real-time) verification results and validation reports for CAMS can be found at <https://atmosphere.copernicus.eu/global-services>.

A summary of the reanalysis system is given in section 2. Sections 3-8 gives an overview of the performance of the system for various species.



Table 1.1: Overview of the trace gas species and aerosol aspects discussed in this CAMS reanalysis validation reports. Shown are the datasets assimilated in the CAMS reanalysis (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). See Section 2 for more information on the time range of availability of the different datasets.

Species, vertical range	Assimilation	Validation
Aerosol, optical properties	MODIS Aqua/Terra AOD, AATSR	AOD, Ångström: AERONET, GAW, Skynet, MISR, OMI, lidar, ceilometer
Aerosol mass (PM10, PM2.5)	-	European AirBase stations
O <sub>3</sub> , stratosphere	MIPAS, MLS, SCIAMACHY, GOME-2, OMI, SBUV-2	O <sub>3</sub> sonde, lidar, MWR, FTIR, HALOE, GOMOS, OSIRIS, SCIAMACHY, ACE-FTS
O <sub>3</sub> , UT/LS	Indirectly constrained by limb and nadir sounders	MOZAIC, IAGOS, ozone sonde
O <sub>3</sub> , free troposphere	Indirectly constrained by limb and nadir sounders	MOZAIC, IAGOS, ozone sonde, IASI, aircraft campaigns
O <sub>3</sub> , PBL / surface	-	Surface ozone: WMO/GAW, NOAA/ESRL-GMD, AIRBASE, EMEP
CO, UT/LS	-	MOZAIC, IAGOS
CO, free troposphere	MOPITT	MOZAIC, IAGOS, MOPITT, IASI, TCCON, aircraft campaigns
CO, PBL / surface	Indirectly constrained by satellite IR sounders	Surface CO: WMO/GAW, NOAA/ESRL
NO <sub>2</sub> , troposphere	SCIAMACHY, OMI, GOME-2	SCIAMACHY, GOME-2, MAX-DOAS, aircraft campaigns
HCHO	-	SCIAMACHY, GOME-2, MAX-DOAS, aircraft campaigns
Other gases, troposphere		Aircraft campaigns (OH, NO, HO <sub>2</sub> )
Stratosphere, other than O <sub>3</sub>	-	SCIAMACHY, GOME-2 (NO <sub>2</sub> column)

The CAMS validation reports are 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>.



## 2. System summary and model background information

The specifics of the CAMS reanalysis model versions are given in section 2.1. Other systems used for the evaluation are provided in section 2.2. An overview of products derived from this system is given in section 2.3. Several external products used for validation and inter-comparison are listed in section 2.4. Timeliness and availability of the CAMS products is given in section 2.5.

### 2.1 System based on the ECMWF IFS model

Key model information is given on the CAMS reanalysis data-assimilation and its control experiment. Further details on the different model runs and their data usage can be found at <http://atmosphere.copernicus.eu/documentation-global-systems> and <http://atmosphere.copernicus.eu/user-support/operational-info>

#### 2.1.1 CAMS reanalysis system

The reanalysis system (Inness et al., 2019) 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 by Morcrette et al. (2009). The data are originally stored under the following experiment IDs (class=rd):

- gqm5 20030101 - 20030630
- gq7s 20030701 - 20040331
- gtyl 20040401 - 20040831
- gq7u 20050101 - 20081231
- gstx 20090101 - 20121231
- gstz 20130101 - 20161231
- gwoe 20170101 - 20171231
- gyy2 20180101 - 20181231
- h3x2 20181230 – 20191231
- hbsd 20200101 – 20201231

In the end, the entire reanalysis was archived as **exp=eac4, class=MC** and will be available from the Copernicus Atmosphere Data Store (ADS), and users do not have to worry about the underlying experiments. The model resolution is T255 with 60 vertical layers. Here a summary of the main specifications of the CAMS reanalysis system is given.

- The meteorological model is based on IFS version cy42r1, with interactive ozone and aerosol in radiation scheme, see also <http://www.ecmwf.int/en/forecasts/documentation-and-support/changes-ecmwf-model>; the model resolution is T255L60.

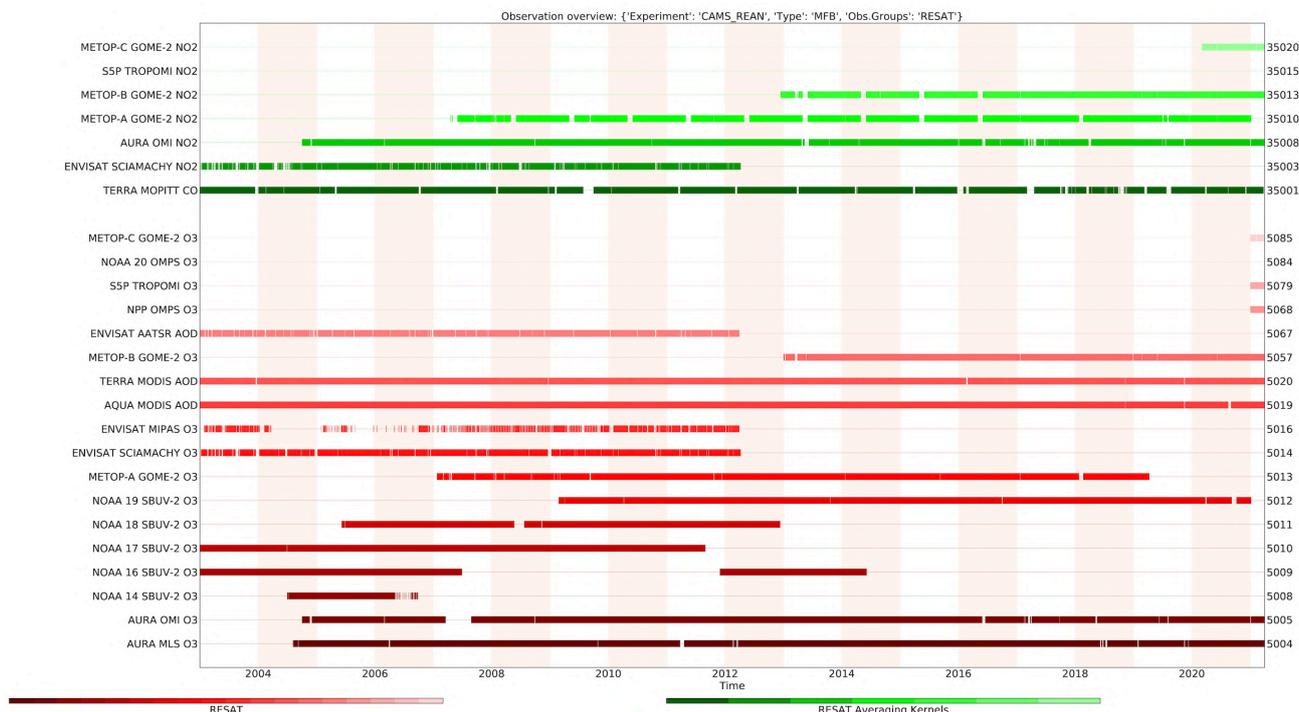


Figure 2.1: Satellite observation usage in the reanalysis for the period 2003-2020. The five green rows correspond to GOME-2, OMI and SCIAMACHY NO2, MOPITT CO, and are assimilated using the averaging kernels in the retrieval product.

- 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 Teyssère, 2007) as already available in IFS, while stratospheric NO<sub>x</sub> is constrained through a climatological ratio of HNO<sub>3</sub>/O<sub>3</sub> 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.
- Anthropogenic reactive gas emissions are based on MACCity (Granier et al., 2011), where wintertime CO emissions have been scaled up over Europe and US (Stein et al., 2014). Monthly mean biogenic emissions are derived from hourly MEGAN-MERRA (Sindelarova et al., 2014).
- CO<sub>2</sub> emission are from EDGAR v4.2 (anthropogenic), CHTESSEL (ecosystem), ACCMIP/EDGAR (aviation), Takahashi 2009 (ocean). CH<sub>4</sub> emissions are from LPJ-HYMN (wetland, natural), Bergamaschi 2013 (chemical sinks) and EDGAR v4.2 (anthropogenic).
- NRT fire emissions are taken from GFASv1.2 (Kaiser et al. 2012).



Table 2.2: Satellite retrievals of reactive gases, greenhouse gases and aerosol optical depth that are actively assimilated in the reanalysis. The table only contains datasets used for the years 2003-2020.

Variable	Instrument	Satellite	Product	Origin, period	AK
O3	SCIAMACHY	Envisat	TC	CCI; 20030101 - 20120408	no
O3	MIPAS	Envisat	PROF	ESA NRT: 20030127- 20030720 MARS ESA NRT: 20030721-20040326 CCI: 20050127-20120331	no
O3	MLS	Aura	PROF	V4: 20040803-20180312 V4 NRT: 20180313-	
O3	OMI	Aura	TC	KNMI V003; reprocessed 20040803-20150531 NRT 20150601-	no
O3	GOME-2	Metop-A	TC	CCI BIRA (fv0100): 20070123-20121231 CCI BIRA (fv0300): 201301-201612 GDP4.8 ACSAF/DLR: 20170101 -20181231	no
O3	GOME-2	Metop-B	TC	CCI BIRA (fv0300): 201301-201612 GDP4.8 ACSAF/DLR: 20170101 -	no
O3	GOME-2	Metop-C	TC	NRT GDP4.9: 202101-	
O3	SBUV/2	NOAA-14	PC 13L	NASA v8.6: 200407-200609	no
O3	SBUV/2	NOAA-16	PC 13L 21L	NASA v8.6: 200301-200706 20111201-20130708 NASA v8.6 NRT: 20130709-201406	no
O3	SBUV/2	NOAA-17	PC 13L	NASA v8.6: 200301-201108	no
O3	SBUV/2	NOAA-18	PC 13L	NASA v8.6: 200507-201211	no
O3	SBUV/2	NOAA-19	PC 13L 21L	NASA v8.6: 200903-20130708 NASA v8.6, NRT: 20130709-20201231	no
CO	MOPITT	Terra	TC	V6 (TIR): 2003-2016 V7 (TIR): 201701 – 20180924 V7 LANCE (TIR): 20180925-20190630 NRT V8 LANCE (TIR): 20190701 -	yes
NO2	SCIAMACHY	Envisat	TRC	v1p: 20030101-20101231 v2: 20110101-20120409	yes
NO2	GOME-2	Metop-A	TRC	O3MSAF GDP4.8: 20070418-20171106 NRT: 20171106-20201231	yes
NO2	GOME-2	Metop-B	TRC	O3MSAF GDP4.8: 20130101-20171106 NRT: 20171106-	yes
NO2	GOME-2	Metop-C	TRC	O3MSAF GDP4.9: 20200305	yes
NO2	OMI	Aura	TRC	COL3: 20041001-2011231 DOM: 20110101-20121231 NRT: 20130101-20140430 NRT SAPP: 20140501-	yes
AOD	AATSR	Envisat	TC	CCI, 20030101-20120331	no
AOD	MODIS	Terra	TC	COL6; 20030101-20151231 NRT COL6: 20160101-	no
AOD	MODIS	Aqua	TC	COL6; 20030101-20151231 NRT COL6: 20160101-	no



The following updates were applied to the chemistry (since the CAMS Interim Reanalysis):

- Update of heterogeneous rate coefficients for  $\text{N}_2\text{O}_5$  and  $\text{HO}_2$  based on clouds and aerosol.
- Modification of photolysis rates by aerosol.
- Dynamic tropopause definition based on T profile for coupling to stratosphere and tropospheric mass diagnostics.
- Monthly mean VOC emissions calculated by the MEGAN model for all VOCs and for whole period 2003-2015 period.
- Bug fixes, in particular for diurnal cycle of dry deposition whose correction has decreased ozone dry deposition (about 15-20%).

The aerosol model includes 12 prognostic variables, which are 3 bins for sea salt and desert dust, hydrophobic and hydrophilic organic matter and black carbon, sulphate aerosols and its precursor trace gas  $\text{SO}_2$  (Morcrette et al., 2009). Aerosol total mass is constrained by the assimilation of MODIS AOD (Benedetti et al. 2009) and AATSR AOD. A variational bias correction for the MODIS AOD is in place based on the approach used also elsewhere in the IFS (Dee and Uppala, 2009).

### 2.1.2 Control run

The control run applies the same settings as the reanalysis, based on the IFS-CB05 system with CAMS aerosol for cy42r1, except that data assimilation is not switched on. It consists of 24h cycling forecasts and uses the meteorological fields from the CAMS reanalysis.

Note that an issue was discovered in 2018 in the original control run (experiment "gqk3"), affecting in particular the aerosols. A new control run was produced for the full period 2003-2020. This new control run is used in this report (experiments h1f5, h1gb, h1in, h293, h29d, h2ol, h2om, h3dk, h7er).

## 2.2 Other systems

### 2.2.1 The MACC reanalysis and CAMS forecasts

The previous reanalysis was produced during the MACC project (Inness et al., 2013). This reanalysis is available through the CAMS website, or ECMWF archive with EXP='rean', CLASS='mc'.

In a few places the performance of the reanalysis is compared with the CAMS operational forecasts/analyses. This system is sometimes referred to as CAMS "o-suite".

### 2.2.2 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:  $\text{O}_3$ ,  $\text{H}_2\text{O}$ ,  $\text{HNO}_3$ ,  $\text{HCl}$ ,  $\text{HOCl}$ ,  $\text{N}_2\text{O}$  and  $\text{ClO}$ .
- 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  $3.75^\circ$  longitude by  $2.5^\circ$  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](http://www.copernicus-stratosphere.eu/4_NRT_products/3_Models_changelogs/BASCOE.php).

### 2.2.3 TM3DAM and the multi-sensor reanalysis

One of the MACC products was a 30-year reanalysis, near-real time analysis and 10-day forecast of ozone column amounts performed with the KNMI TM3DAM data assimilation system, the Multi-Sensor Reanalysis (MSR) system (van der A et al., 2010, 2015), [http://www.temis.nl/macc/index.php?link=o3\\_msr\\_intro.html](http://www.temis.nl/macc/index.php?link=o3_msr_intro.html).

The corresponding validation report can be found at

[http://www.copernicus-atmosphere.eu/services/gac/global\\_verification/validation\\_reports/](http://www.copernicus-atmosphere.eu/services/gac/global_verification/validation_reports/).

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

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

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

### 2.2.4 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/>) started its activities in 2012. During this period, the Regional Center has established a protocol to routinely exchange products from dust forecast models and observations (i.e. ground-based and satellite aerosol products) as the basis for both near-real-time and delayed common model evaluation.

Global and regional dust models for international operational and research institutions are currently providing daily operational dust forecasts (i.e. dust optical depth, DOD, and dust surface concentration). Different multi-model products are generated from the different prediction models. Two products describing centrality (multi-model median and mean) and two products describing spread (standard deviation and range of variation) are daily computed. In order to generate them, the model outputs are bi-linearly interpolated to a common grid mesh of 0.5° x 0.5°. The SDS-WAS



multi-model DOD (at 550 nm) Median from available dust prediction models participating in the SDS-WAS Regional Center is used for the validation of the CAMS NRT streams.

The current routine evaluation of dust predictions is focused on total-column dust optical depth (DOD) and uses remote-sensing retrievals from sun-photometric (AERONET) and satellite (MODIS) measurements.

The updated list of dust models participating in the model inter-comparison can be found at <https://sds-was.aemet.es/forecast-products/forecast-evaluation/model-inter-comparison-and-forecast-evaluation/>

### 2.3 CAMS reanalysis products

The entire reanalysis was archived as exp=**eac4**, class=**MC**. Data is made available through the Copernicus Atmosphere Data Store (ADS), <https://ads.atmosphere.copernicus.eu>. The CAMS 3D reanalysis products are stored as 3-hourly fields and will update the MACC global reanalysis which is available through the CAMS catalogue, <https://atmosphere.copernicus.eu/catalogue#/>. The following fields are archived:

- Forecast fields: From 0z, 3-hourly, step=0,3,..., 48
- Analysis fields: Every 3 hours, e.g. 0z, 3z,...21z
- Surface forecast fields: From 0z, 1-hourly, step=0,1,2,...,48

Several parameters are also available as synoptic monthly means, for each particular time and forecast step (stream=mnth) and as monthly means of daily means, for the month as a whole (stream=moda), see

<https://confluence.ecmwf.int/display/CKB/CAMS+Reanalysis+data+documentation>.

### 2.4 Validation results for reactive gases and aerosol

The following sections present the validation results of the CAMS global reanalysis for aerosols and reactive gases. Naming and color-coding conventions predominantly follow the scheme as given in the table below:

Name in figs	Experiments	Colour
{obs name}	{obs}	black
<b>CAMS reanalysis</b>	<b>eac4</b>	<b>red</b>
<b>CAMS control</b>	<b>h1f5 etc.</b>	<b>blue</b>
<b>MACC reanalysis</b>	<b>rean, class=mc</b>	<b>green</b>



## 3. Aerosol

### 3.1 Global aerosol distribution

The global aerosol fields are analysed for the CAMS reanalysis. For some comparisons also the previous MACC reanalysis is used. More detailed results can be found on the [AeroCom/CAMS website](#) in the CAMS reanalysis section.

The evaluation in table 3.1.1 shows the average global AOD and speciated AOD for the two reanalysis and illustrates a shift from MACC to CAMS in aerosol composition. Changes from MACC to CAMS in total AOD are relatively smaller than the changes in species AOD. The decrease in total AOD is due to reductions in dust, sulphate and sea salt AOD. Table 3.1.2 shows global means of AOD in two periods where different satellite data were assimilated, because AATSR was ceasing to operate in early 2012. AATSR seems to have amplified the annual cycle of AOD by circa 0.01, while the global annual mean is nearly indistinguishable for the two periods. Figures 3.1.1 shows where the changes from MACC to CAMS reanalysis occur exemplary for sulphate AOD.

Table 3.1.1: Comparison of the CAMS reanalysis and MACC reanalysis. Mean annual global total and speciated aerosol optical depth (AOD) in for the years 2003-2012, available in common from both reanalyses.

	CAMS reanalysis	MACC reanalysis
AOD@550	0.146	0.188
BC-OD@550	0.006	0.007
Dust-OD@550	0.019	0.045
OA-OD@550	0.053	0.023
SO4-OD@550	0.033	0.051
SS-OD@550	0.035	0.061

Table 3.1.2: Comparison of two periods, one including AATSR and MODIS assimilation (2008-2011) and one when only MODIS was available (2013-2016). Total aerosol optical depth (AOD) in the global area and at Aeronet stations.

<i>AOD global mean</i>	2008-2011	2013-2016
Assimilation	MODIS+AATSR	MODIS
Global area	0.146	0.147
Global area winter (DJF)	0.123	0.129
Global area summer (JJA)	0.165	0.164
@Aeronet sites	0.213	0.202
@Aeronet sites winter (DJF)	0.157	0.157
@Aeronet sites summer (JJA)	0.250	0.233

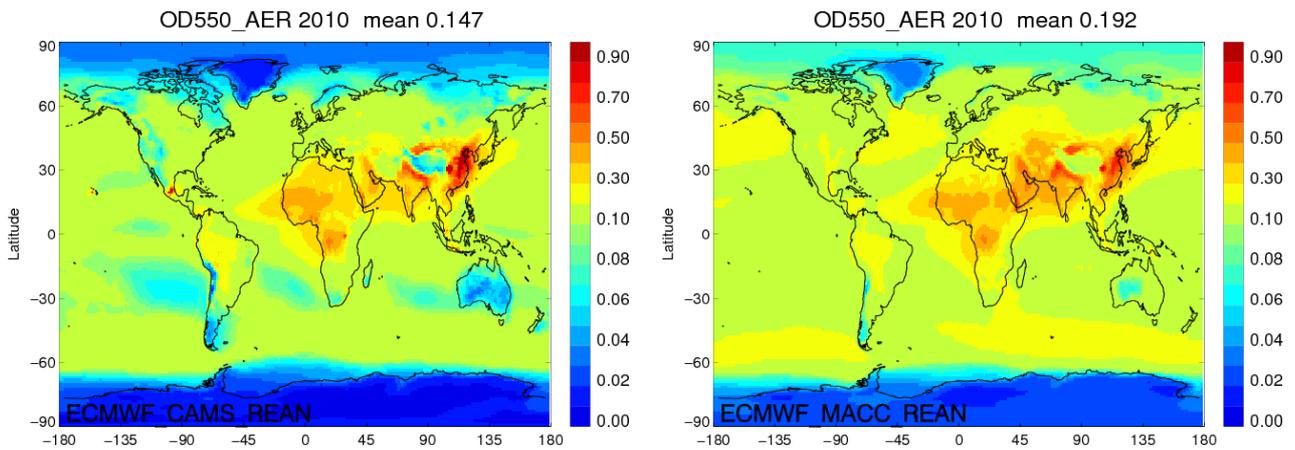


Figure 3.1.1: Comparison of aerosol optical depth (AOD) from IFS experiments CAMS reanalysis (left) and MACC reanalysis (right) for the year 2010.

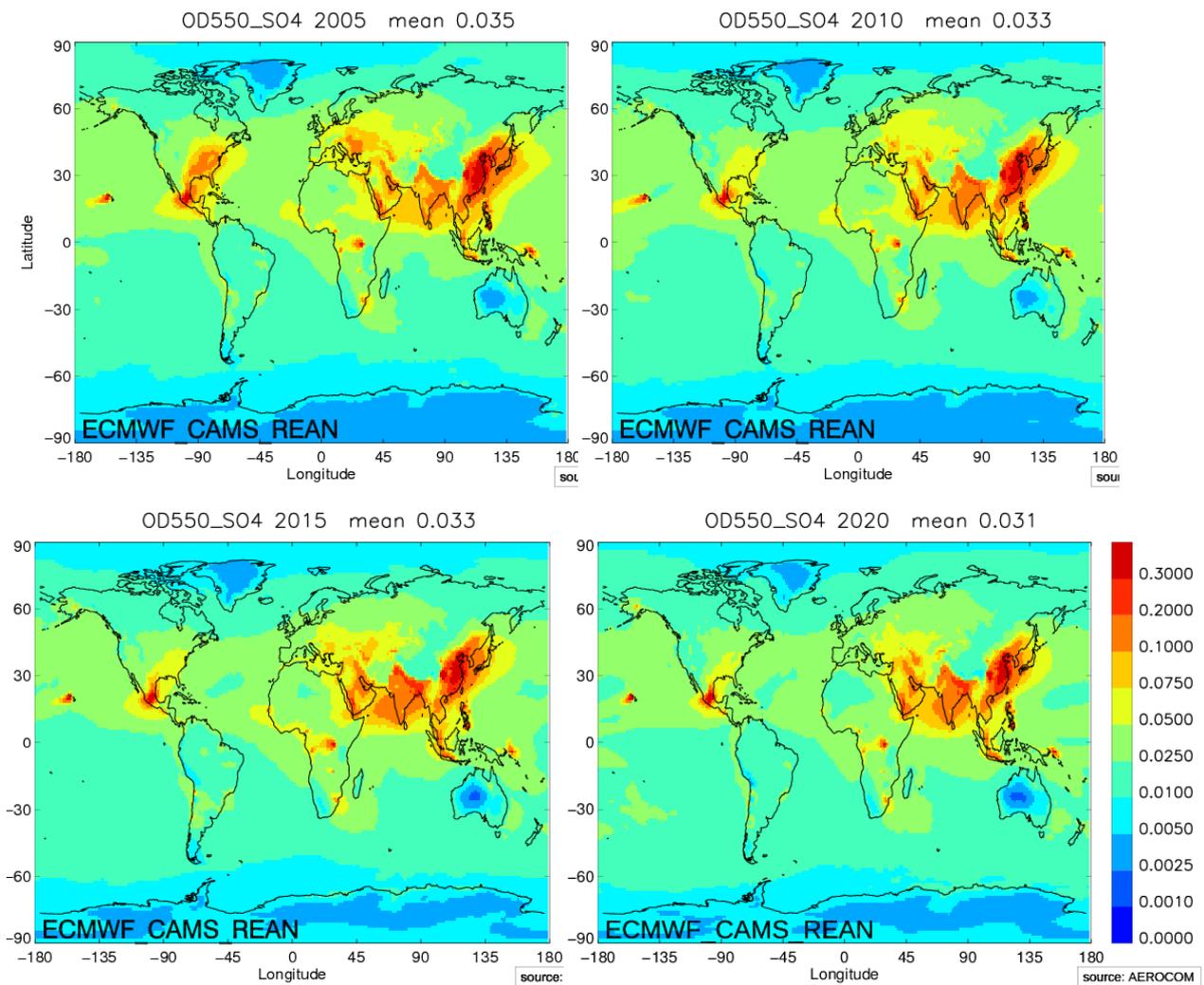


Figure 3.1.2: Evolution of sulphate optical depth in CAMS reanalysis from 2005, 2010, 2015 to 2020.

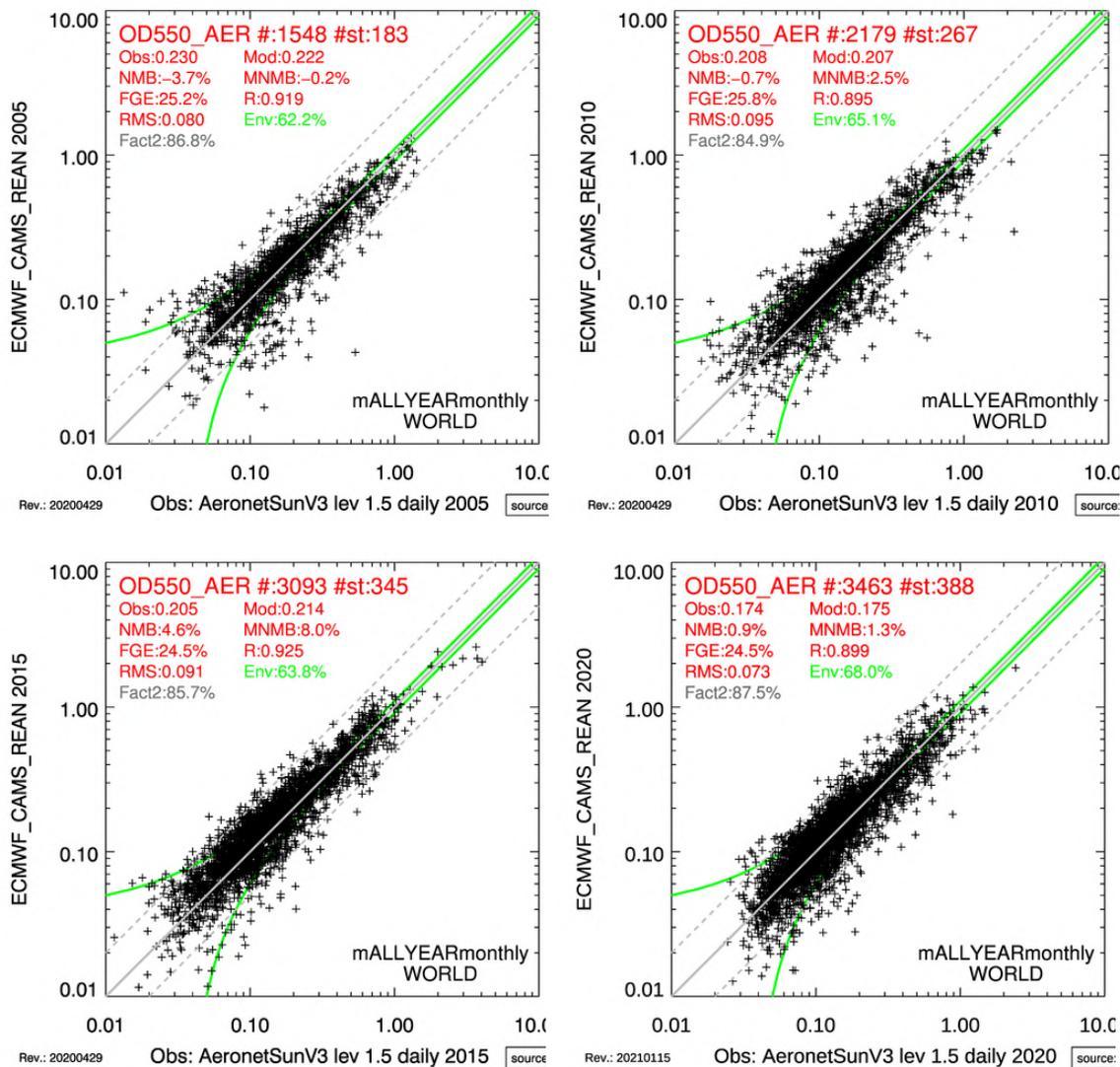


Figure 3.1.3: Evaluation of simulated monthly AOD against Aeronet version 3, level 1.5 sun photometer measurements in CAMS reanalysis for the years 2005, 2010, 2015, 2020. (# number of measurements; #st: number of stations, Obs: mean observations, Mod: mean model; NMB: Normalised mean bias; MNMB: Modified normalised mean bias, FGE: Fractional gross error, R: correlation, RMS: root mean square error, Env: GCOS error envelope, Fact2: Model data fraction within factor two of measurement.)

Volcanic eruptions add sulphate near Hawaii and Middle America, seen as little hot spots on the map. Figure 3.1.2 illustrates the decreasing trend in sulphate from 2005 to 2020. High values in CAMS reanalysis near Hawaii lead to some local outliers in the scatterplot in figure 3.1.3, where Aeronet measurements show much smaller values, indicating an overestimation of the volcanic derived sulphate. Aerosol dispersion is less pronounced in CAMS than in MACC, which is visible in smaller sulphate AOD at polar locations (eg sulphate AOD in figure 3.1.2).

The overall performance of the CAMS reanalysis is very good and stable compared exemplary for years 2010 and 2020 in figures 3.1.4, 3.1.5 and 3.1.6.

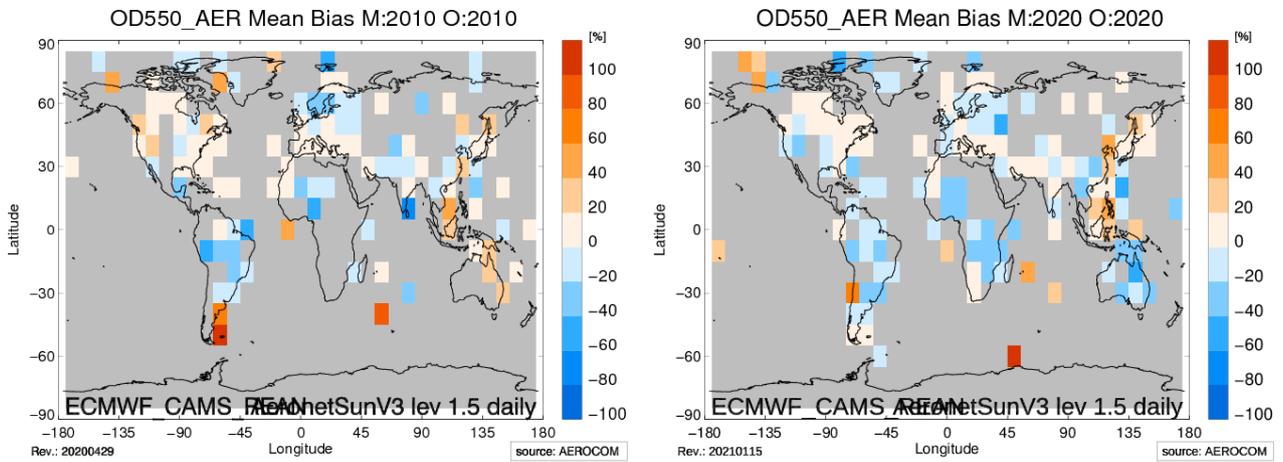


Figure 3.1.4: Regional relative mean bias of simulated daily AOD against Aeronet sun V3 level 1.5 photometer measurements in CAMS reanalysis for the year 2010 and 2020.

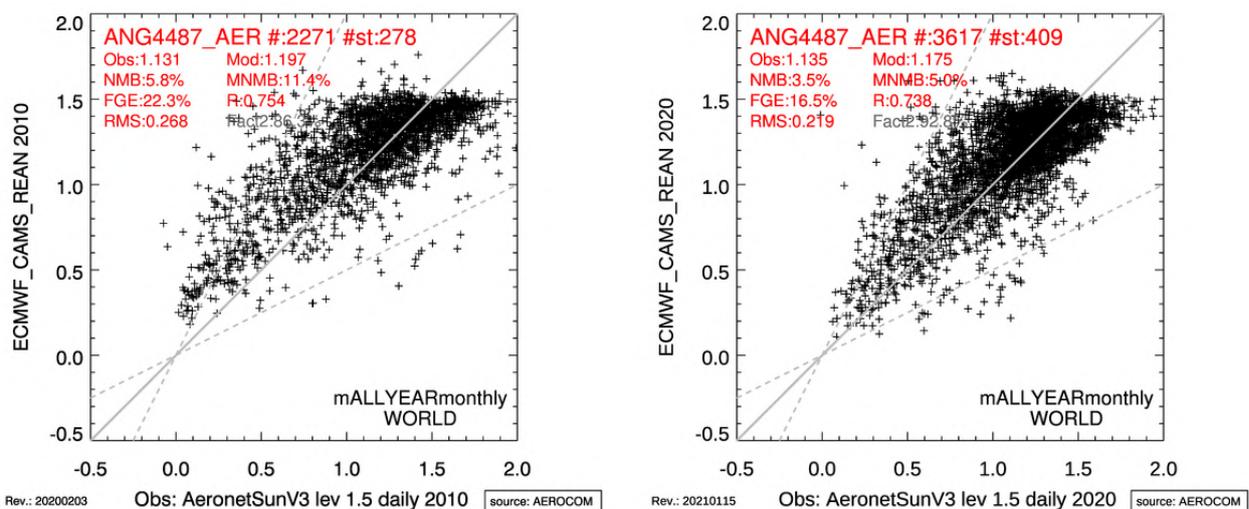


Figure 3.1.5: Simulated monthly mean Angstrom Coefficient in CAMS reanalysis (2010 and 2020) against Aeronet Sun version 3 level 1.5 photometer measurements. Statistics shown as in figure 3.1.3.

For the full reanalysis period of 2003-2020, correlation (Fig. S.1-b) based on daily optical depth and Aeronet Direct-sun Version 3 Level 1.5 observations shows that the CAMS reanalysis improves slightly with time, in particular in the beginning. MACC reanalysis shows quite similar correlation throughout the period but is a little lower than the CAMS reanalysis. The modified normalized mean bias of AOD (MNMB, see Figure S.1-a), shows a seasonal variation and the CAMS reanalysis has bias values between -10% and +20 %. CAMS reanalysis tends to overestimate in summer. MACC reanalysis consistently overestimates more (biases between 0% and +40 %) along with a smaller seasonal cycle.

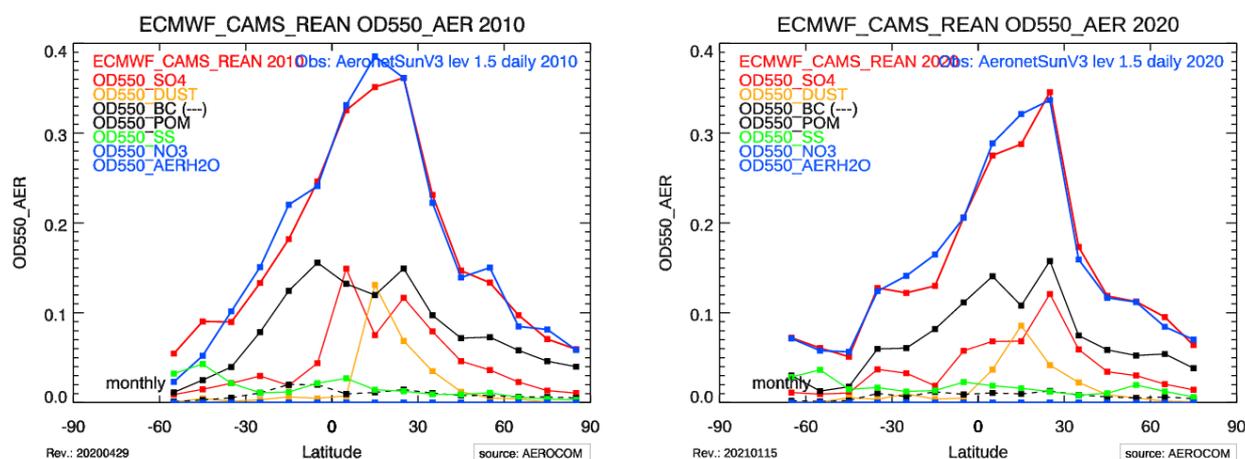


Figure 3.1.6: Evaluation of latitudinal distribution of aerosol in CAMS reanalysis (2010 and 2020). Total AOD from model (light red curve) at Aeronet sites against total AOD from Aeronet (dark blue curve), aggregated as mean in 10 degree latitude bands. AOD speciation at Aeronet sites split in sulphate (dark red), dust (brown), black carbon (black dashed), organic (black) and sea salt (green).

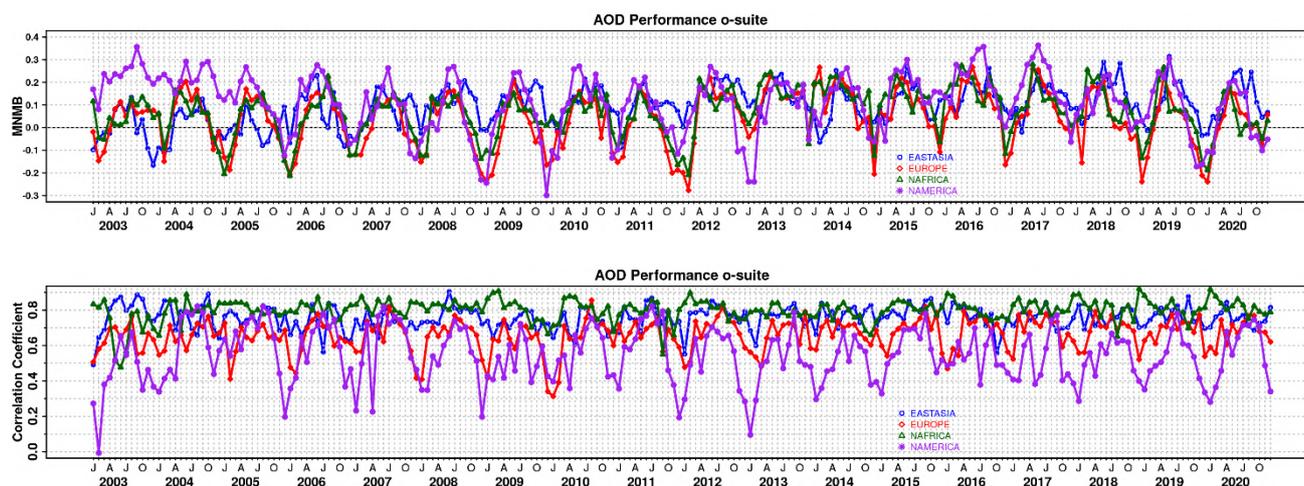


Figure 3.1.7: a) Modified normalized mean bias (MNMB) and b) correlation coefficient in AOD for 2003-2020 based on daily AOD comparison (Aeronet V3 level 1.5) in four world regions [East Asia (blue); Europe (red); North Africa (green); North America (purple)] for the CAMS reanalysis.

A seasonal cycle in AOD bias is also visible in continental regions, especially for the modified normalized mean bias (MNMB) (Fig. 3.1.7-a). A persistent overestimation of AOD in North America in the first two years (2003, 2004) disappears later and the bias becomes more like for the other regions. Also, for correlation (see Figure 3.1.7-b) North America is lowest and with the strongest seasonal variability among the regions. CAMS values in Europe show higher correlation and a less pronounced seasonal variability. Dust influenced data in Northern Africa/Southern Europe are best reproduced by the CAMS reanalysis, with a higher bias after 2012 due to a change in data assimilated.

The composition variations over time in the model are shown in figure 3.1.8, with no particular trend. Organic aerosol and dust show seasonal cycles due to variation in source strength over the year.

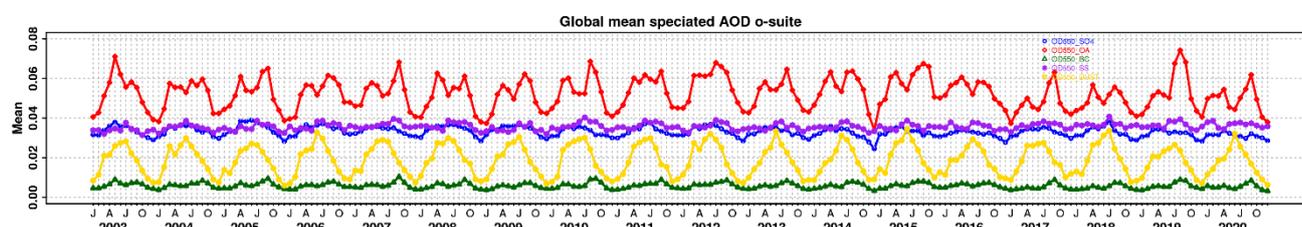


Figure 3.1.8: Evolution of the aerosol components of total AOD@550nm [OD550\_SO4 = sulphate(blue); OD550\_OA = organics(red); OD550\_BC = black carbon(green); OD550\_SS = sea salt(purple); OD550\_DUST = dust(yellow)] for the CAMS reanalysis.

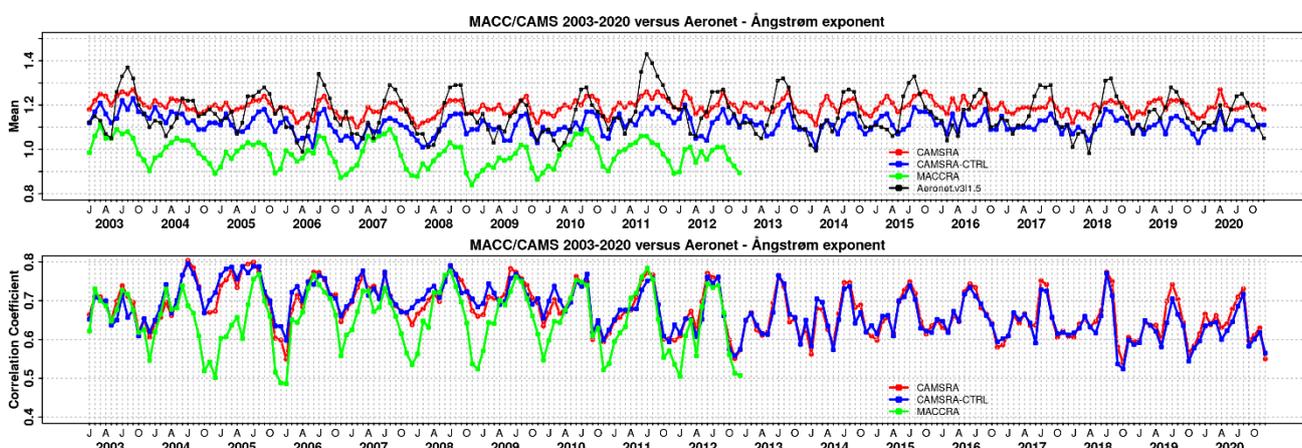


Figure 3.1.9: a) Evolution of mean Ångström exponent for 2003–2020 at Aeronet sites (Aeronet V3 level 1.5 based on matching monthly mean values [CAMS reanalysis (red) and control (blue); MACC reanalysis (green); observations (black)]. b) Correlation using daily matching Ångström exponent.

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. Figure 3.1.9 a) shows the temporal evolution of simulated and observed mean Ångström exponent, while the correlation is found in figure 3.1.9 b). The CAMS reanalysis shows less seasonal variability than the observations, on average slightly overestimating the Ångström exponent (5-20%).

This would be consistent with dust size more correct without assimilation, as suggested above. Total AOD is dominated by OA and with less dust for CAMS reanalysis compared to MACC reanalysis. This may explain the higher overall Ångström values in the CAMS reanalysis. Temporal-spatial correlation figure 3.1.9 b) being higher in winter in both CAMS simulations than in MACC indicates partially better tracing of aerosol size and implicitly composition variability.

PM 10 and PM25 concentrations are both biased high against a climatology. Figure 3.1.10 shows that the surface PM10 concentrations are within a factor of 2 for 40-70% of the monthly station data. Summer concentrations are better reproduced. A considerable (more) decreasing trend of simulated PM25 concentrations in North America and Europe reduces the bias with time and leads to a larger fraction of simulated PM25 concentrations being within factor 2 of the climatology. Lower aerosol loads in the control simulation in general, thus a smaller average bias, lead to a better PM25 score for the control simulation.

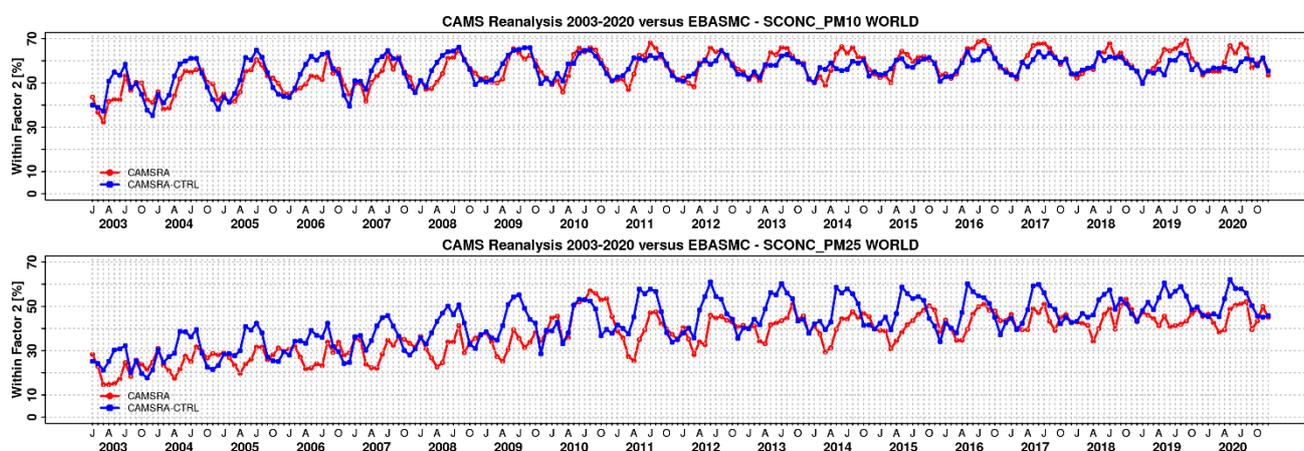


Figure 3.1.10: a) Timeseries (2003-2020) of PM 10 concentrations within a factor 2 of observational climatology at 166 rural/remote sites in Europe and North America [CAMS reanalysis (red) and control (blue)]. b) As a) but for PM25 concentrations at 143 sites.

### 3.2 Dust evaluation over North Africa, Middle East and Europe

Over North Africa, Middle East and Europe, the seasonal DOD fields from CAMS reanalysis show a distinct seasonal pattern linked to the spatial distribution of dust emissions and transport throughout the period 2003-2020, in good qualitative agreement with both the MODIS Terra/Aqua and MISR (Figure 3.2.1) and AERONET (Table 3.2.1, Figures 3.2.2 and Figure 3.2.3). The Bodéle, as well as the desert dust sources in Mauritania, Maghreb, Saudi Arabia and Oman, are systematically underestimated throughout the year.

The CAMS reanalysis tends to underestimate AOD in comparison with observations, particularly over the subtropical North Atlantic and the Sahel belt, as well as over the Red Sea, the Gulf of Oman and the Arabian Sea (see Figure 3.2.1). The most striking result is the big difference between AOD and DOD from CAMS in the desert areas (i.e., the Sahara and the Middle East) where mineral dust aerosol unambiguously dominates (see Figure 3.2.1).

The comparison with AERONET direct-sun cloud-screened dust-filtered AOD version 3 observations (i.e., DOD, on 3-hourly basis) shows that the CAMS reanalysis reproduces rather well the annual variability, showing annual correlation coefficients between 0.80 and 0.90, on average for all the AERONET sites. However, the CAMS reanalysis tends to underestimate DOD with an annual mean bias between -0.05 and -0.09, RMSE between 0.13 and 0.22 and MNMB between -0.38 and -0.51 on average for all the AERONET sites. These results are consistent along the 18-years study period (see Table 3.2.1).

Additionally, the big differences between AOD and DOD observed over Northern Africa and Middle East (causing strong DOD underestimations) when compared with AERONET (see Table 3.2.1) are explained by an overestimation of organic matter (OM) from biomass burning during wintertime and also an overestimation of secondary organics over heavily populated areas during summertime. This makes OM a bit too preponderant through the assimilation step and hence, leads to a DOD underestimation.

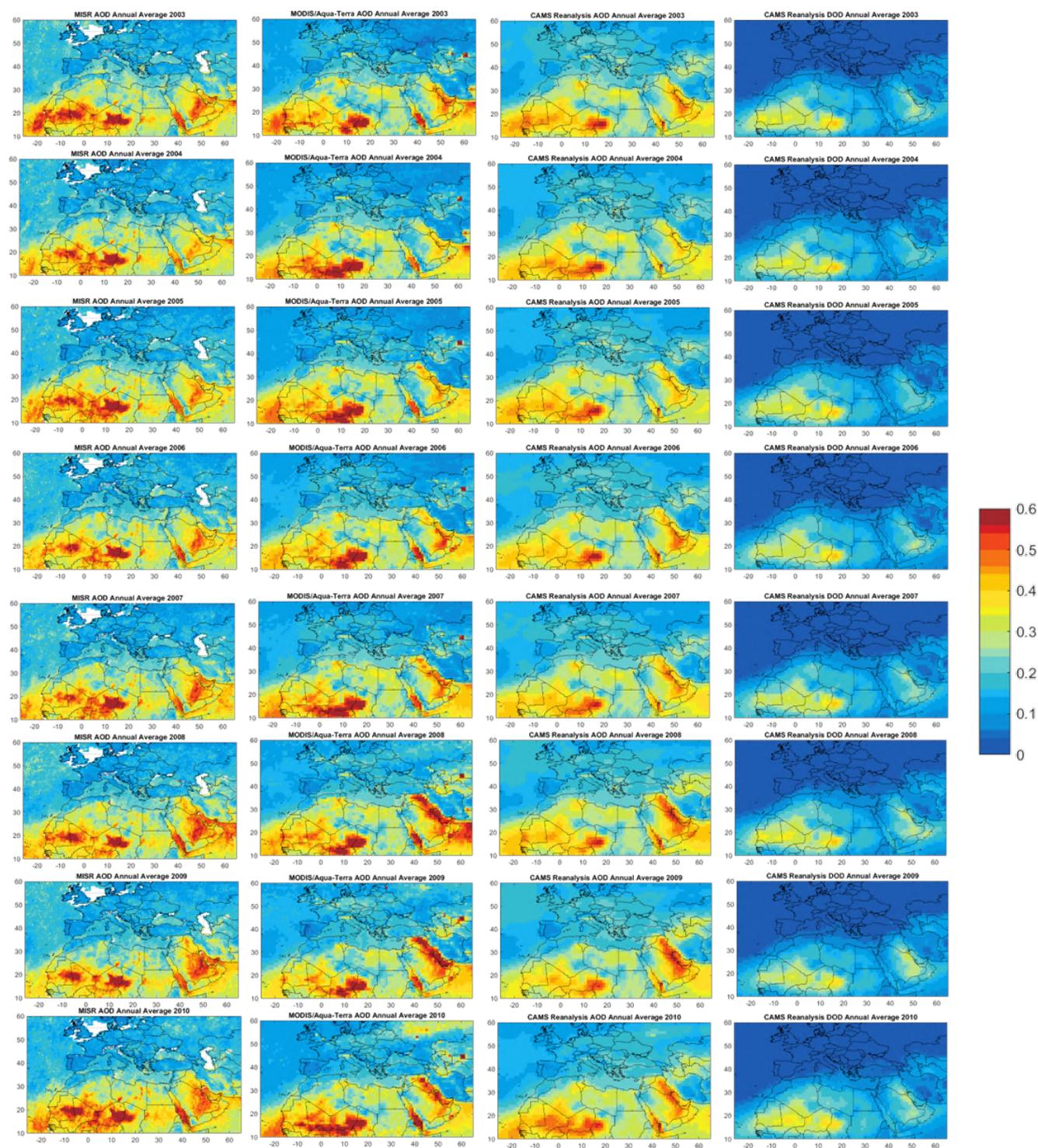


Figure 3.2.1a: Annual averaged AOD from MISR monthly Level 3 monthly  $0.5^\circ \times 0.5^\circ$  global product (first column), MODIS Collection 6 Terra/Aqua merged Dark target and Deep Blue Level 3 daily  $1^\circ \times 1^\circ$  global product (second column), as well as AOD (third column) and DOD from CAMS reanalysis (fourth column) (at 12UTC) for the years 2003-2010.

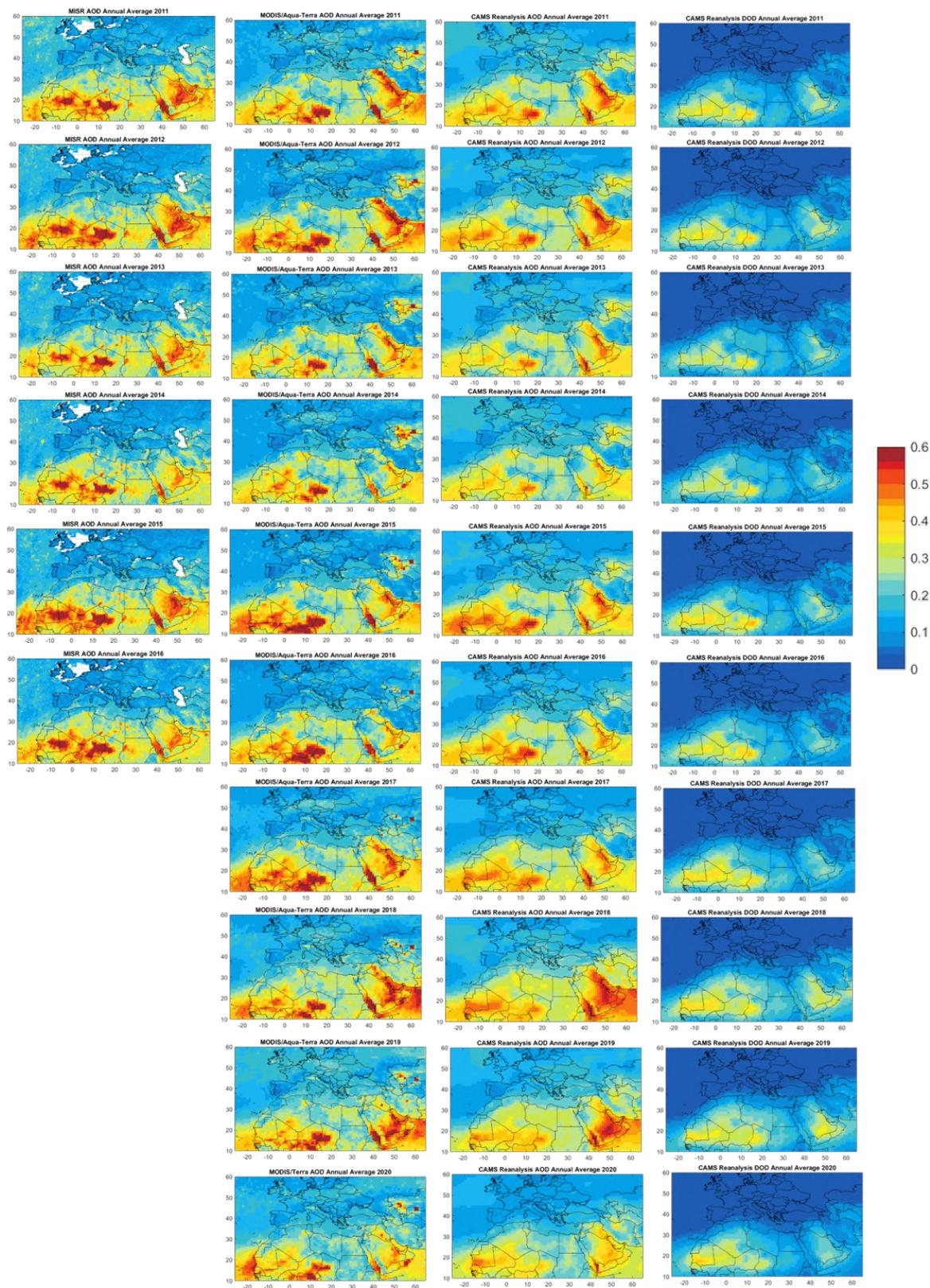


Figure 3.2.1b: Same than Figure 3.2.1b but for the years 2011-2020.

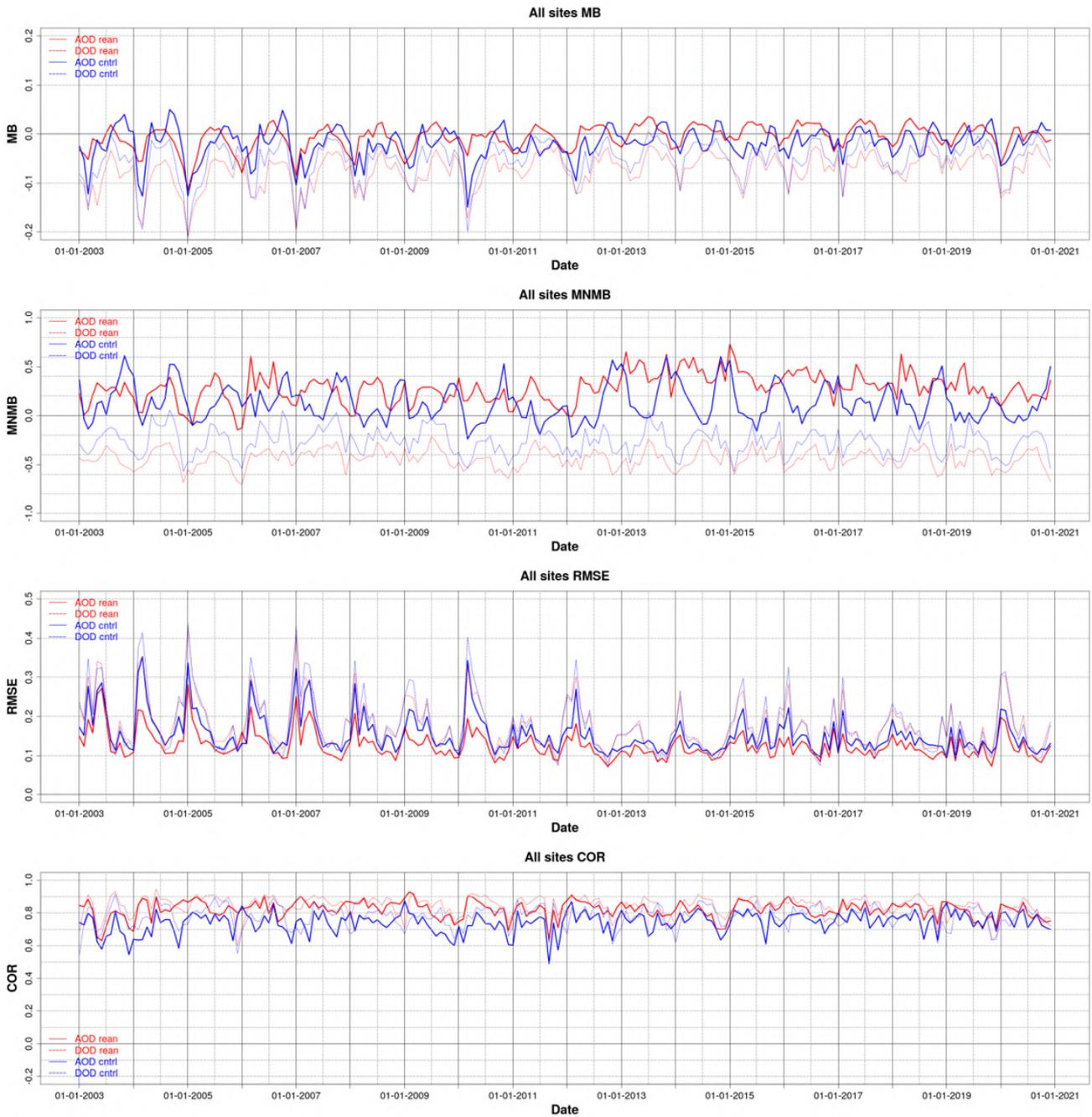


Figure 3.2.2: Monthly MB, RMSE, MNMB and correlation coefficient results for the years 2003-2020 for all the available AERONET observations over Northern Africa, the Middle East and Southern Europe. AOD and dust-filtered AOD (DOD; dashed lines) from AERONET Version 3 cloud-screened is the reference and the colours identify the CAMS reanalysis (red colour) and control (blue colour) experiments.



Table 3.2.1: Annual skill scores (MEAN, SD, MB, MNMB, RMSE and r) of CAMS reanalysis and control run for the years 2003-2020 for Northern Africa, the Middle East and Europe. 3-hourly AOD and dust-filtered AOD (DOD) from AERONET Version 3 cloud-screened is the reference.

		DOD											
		Reanalysis						Control					
	NDATA	MEAN	SD	MB	MNMB	RMSE	r	MEAN	SD	MB	MNMB	RMSE	r
2003	16867	0.11	0.15	-0.08	-0.43	0.22	0.81	0.14	0.17	-0.06	-0.25	0.22	0.75
2004	18726	0.12	0.16	-0.09	-0.42	0.21	0.87	0.15	0.19	-0.06	-0.20	0.22	0.75
2005	23428	0.10	0.14	-0.09	-0.48	0.21	0.86	0.13	0.17	-0.06	-0.29	0.20	0.78
2006	22001	0.11	0.16	-0.09	-0.42	0.20	0.87	0.14	0.18	-0.06	-0.20	0.21	0.76
2007	25482	0.11	0.15	-0.08	-0.38	0.21	0.86	0.13	0.17	-0.06	-0.19	0.22	0.75
2008	25254	0.11	0.15	-0.08	-0.40	0.19	0.86	0.13	0.17	-0.06	-0.25	0.20	0.77
2009	24993	0.10	0.14	-0.07	-0.39	0.18	0.86	0.13	0.17	-0.04	-0.17	0.19	0.75
2010	27332	0.10	0.16	-0.08	-0.51	0.19	0.90	0.12	0.17	-0.06	-0.36	0.21	0.78
2011	34166	0.08	0.13	-0.05	-0.40	0.15	0.86	0.11	0.17	-0.03	-0.21	0.15	0.80
2012	34009	0.09	0.14	-0.07	-0.44	0.18	0.86	0.11	0.16	-0.05	-0.28	0.19	0.78
2013	37286	0.08	0.12	-0.05	-0.44	0.13	0.85	0.10	0.15	-0.03	-0.23	0.13	0.79
2014	34926	0.08	0.12	-0.05	-0.43	0.14	0.85	0.10	0.15	-0.03	-0.25	0.14	0.79
2015	38507	0.10	0.15	-0.07	-0.42	0.18	0.87	0.11	0.16	-0.05	-0.29	0.19	0.79
2016	40394	0.09	0.14	-0.06	-0.44	0.17	0.85	0.11	0.16	-0.04	-0.28	0.17	0.77
2017	38541	0.08	0.13	-0.06	-0.42	0.15	0.86	0.10	0.16	-0.04	-0.27	0.15	0.79
2018	34430	0.09	0.14	-0.06	-0.43	0.16	0.86	0.11	0.16	-0.04	-0.30	0.16	0.78
2019	35303	0.08	0.12	-0.05	-0.43	0.15	0.80	0.10	0.16	-0.02	-0.25	0.15	0.75
2020	27569	0.08	0.12	-0.07	-0.47	0.19	0.81	0.10	0.14	-0.05	-0.32	0.18	0.74
		AOD											
		Reanalysis						Control					
	NDATA	MEAN	SD	MB	MNMB	RMSE	r	MEAN	SD	MB	MNMB	RMSE	r
2003	20522	0.25	0.19	-0.01	0.23	0.17	0.76	0.25	0.19	-0.02	0.18	0.19	0.67
2004	23376	0.25	0.20	-0.01	0.19	0.14	0.84	0.25	0.20	-0.01	0.20	0.20	0.66
2005	28678	0.23	0.19	-0.02	0.19	0.14	0.83	0.23	0.19	-0.03	0.11	0.17	0.74
2006	27531	0.24	0.20	-0.01	0.29	0.14	0.84	0.24	0.20	-0.02	0.18	0.18	0.72
2007	32400	0.23	0.19	-0.01	0.26	0.14	0.84	0.22	0.19	-0.02	0.16	0.18	0.71
2008	32129	0.23	0.19	-0.01	0.22	0.13	0.85	0.21	0.18	-0.03	0.05	0.16	0.74
2009	33026	0.22	0.18	0.00	0.23	0.12	0.84	0.21	0.19	-0.01	0.10	0.16	0.72
2010	35883	0.22	0.20	-0.01	0.20	0.12	0.87	0.19	0.18	-0.03	0.04	0.18	0.72
2011	44176	0.21	0.16	0.00	0.19	0.12	0.80	0.18	0.18	-0.03	-0.01	0.14	0.72
2012	46319	0.21	0.19	0.00	0.25	0.12	0.86	0.18	0.18	-0.02	0.05	0.15	0.77
2013	50050	0.19	0.15	0.01	0.45	0.10	0.81	0.17	0.17	-0.01	0.19	0.12	0.74
2014	47046	0.19	0.15	0.01	0.45	0.11	0.81	0.17	0.17	-0.01	0.16	0.12	0.75
2015	51550	0.21	0.19	0.01	0.38	0.12	0.85	0.18	0.18	-0.02	0.10	0.15	0.77
2016	55283	0.20	0.17	0.01	0.39	0.12	0.82	0.18	0.18	-0.01	0.14	0.14	0.74
2017	53440	0.20	0.17	0.01	0.33	0.11	0.83	0.18	0.18	-0.01	0.13	0.13	0.76
2018	45742	0.21	0.18	0.01	0.34	0.12	0.83	0.19	0.18	-0.02	0.12	0.14	0.74
2019	49402	0.19	0.15	0.00	0.30	0.11	0.80	0.17	0.18	-0.01	0.04	0.13	0.75
2020	39348	0.18	0.21	-0.01	0.22	0.13	0.80	0.17	0.21	-0.02	0.06	0.14	0.73

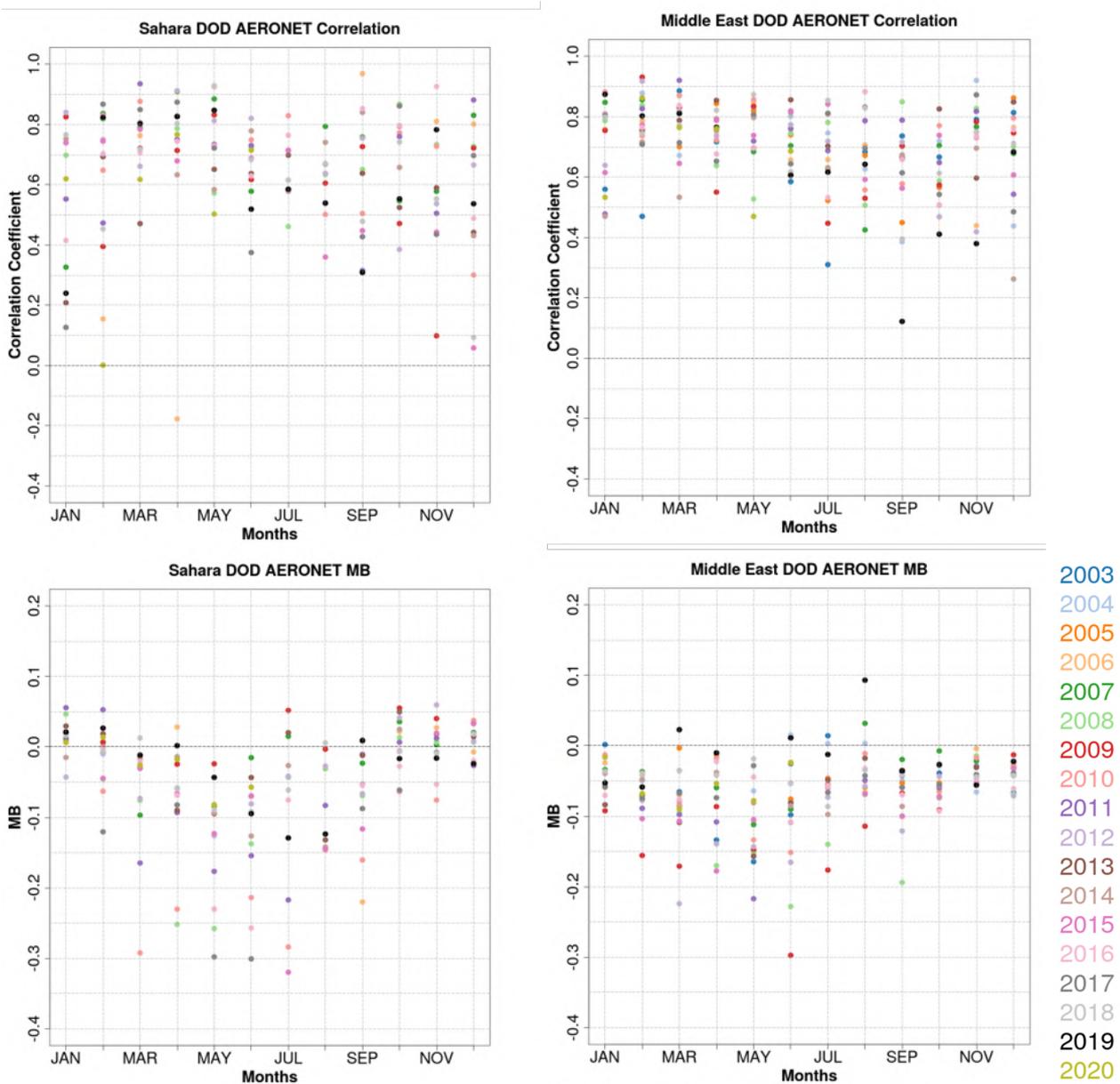


Figure 3.2.3a: Monthly correlation coefficient ( $r$ , upper panels) and Mean Bias (MB, bottom panels) of CAMS reanalysis (on 3-hourly basis) for the years 2003-2020 in the Middle East and Sahara regions. Dust-filtered DOD from AERONET Version 3 cloud-screened is the reference.

In terms of variability, the CAMS reanalysis is able to reproduce the dust seasonal cycle in the AERONET observations with monthly correlation coefficients  $> 0.5$  (see Figure 3.2.2). Higher DOD underestimations are observed during wintertime (see MB in Figure 3.2.2). These strong underestimations are associated to the strong underestimations in the Sahel (see Figure 3.2.3). The CAMS control run that has associated weaker underestimations than the CAMS reanalysis in the comparison with AERONET observations (see MNMB time series in Figure 3.2.2). As pointed out before, the assimilation is reducing the contribution of desert dust (DOD) to the total AOD in comparison in those regions where the model present strongest underestimations as the Sahara.

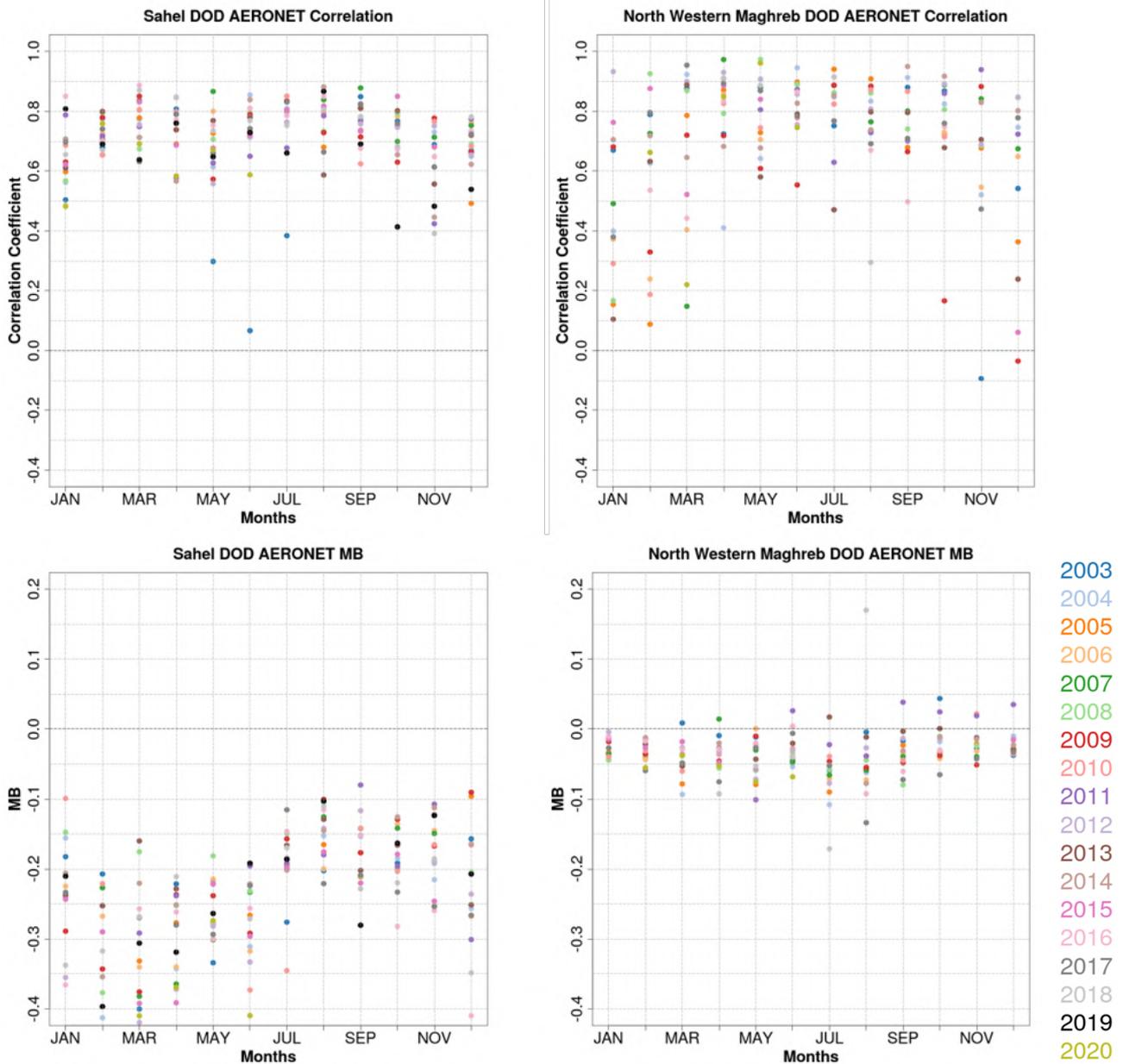


Figure 3.2.3b: Same than Figure 3.2.3a but for the Sahel and North Western Maghreb.

Over desert dust source regions (see Figure 3.2.3), the skill scores show a strong seasonal cycle associated with the dust seasonal activity. Over the Sahara, the largest DOD underestimations ( $< -0.05$ ) and the lower correlation coefficient values ( $< 0.5$ ) are observed during summertime. These underestimations are related to strong and fast dust outbreaks associated with mesoscale convective systems that the model is not able to reproduce. The same summer behaviour is also observed in Northern Western Maghreb. In the Sahel, the CAMS reanalysis presents the largest DOD underestimations (monthly MB up to  $-0.4$ ) particularly in wintertime when there are low-level dust outbreaks. In the Middle East, the model results show a systematic DOD underestimation throughout the year but particularly in summertime the lowest MB and correlation coefficients are

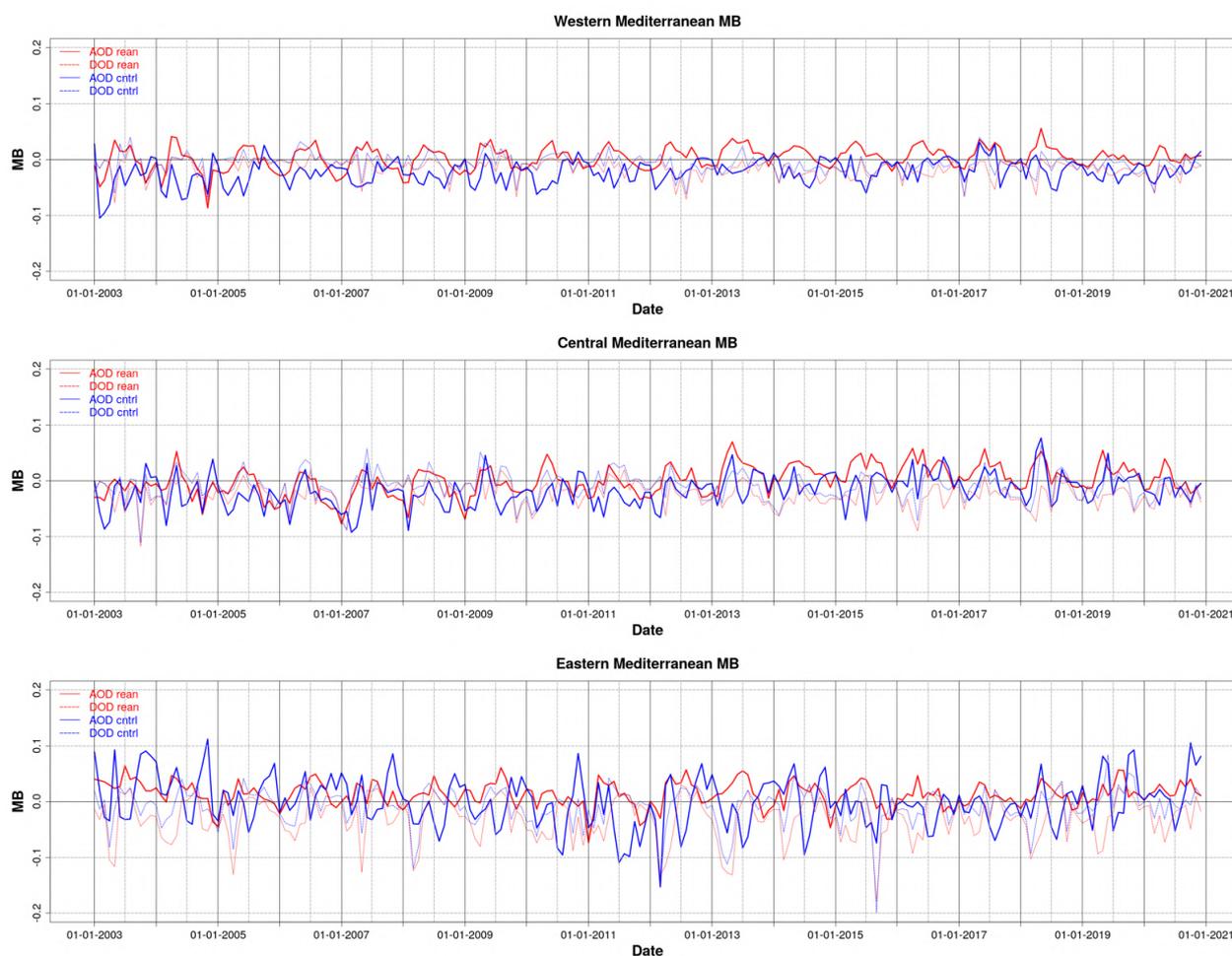


Figure 3.3.1: Monthly MB results for the years 2003-2020 for all the available AERONET observations over Western, Central and Eastern Mediterranean. AOD and dust-filtered AOD (DOD; dashed lines) from AERONET Version 3 cloud-screened is the reference and the colours identified the CAMS reanalysis (red) and control run (blue). The observed DOD underestimations in comparison with AERONET in the Middle East can be partly attributed to a poor representation of small-scale emission processes such as the wind peak associated with the breakdown of the nocturnal low-level jet, the meteorological effects in the vicinities over complex topography, sea breezes and cold pools.

### 3.3 Aerosol validation over the Mediterranean

Over the Mediterranean, the CAMS reanalysis matches well the 3-hourly AOD variability of AERONET cloud-screened observations (Table 3.3.1, Fig. 3.3.1) with annual correlation coefficients between 0.62 and 0.84, with lower skill in the Eastern Mediterranean. In general, the model reproduces the highest AOD peaks, which are associated with desert dust intrusions (see Barcelona, Lecce University and Sede Boker in Fig. 3.3.2). The assimilation of aerosol observations is increasing the background levels in the Mediterranean reducing the higher underestimations observed in the CAMS control experiment (annual MB up to -0.04, see Table 3.3.1 and Fig. 3.3.1) in comparison with the CAMS reanalysis. The CAMS reanalysis tends to keep the annual mean bias between -0.01 and 0.03 in the AOD comparison in the Mediterranean. In the Western and Central Mediterranean sites, a positive mean bias for 2013-2020 is observed (see annual MB in Table 3.3.1).



Table 3.3.1: Annual skill scores (MEAN, SD, MB, MNMB, RMSE and r) of CAMS reanalysis and control run for the years 2003-2020 for Western, Central and Eastern Mediterranean. 3-hourly AOD and dust-filtered AOD (DOD) from AERONET Version 3 cloud-screened is the reference.

Western Mediterranean													
	Reanalysis							Control					
	NDATA	MEAN	SD	MB	MNMB	RMSE	r	MEAN	SD	MB	MNMB	RMSE	r
2003	1162	0.17	0.13	0.00	0.06	0.08	0.79	0.14	0.09	-0.04	-0.09	0.11	0.66
2004	1934	0.17	0.11	0.00	0.17	0.09	0.78	0.13	0.09	-0.04	-0.05	0.10	0.72
2005	4464	0.16	0.11	0.00	0.19	0.08	0.75	0.12	0.09	-0.03	-0.04	0.09	0.67
2006	4418	0.16	0.11	0.00	0.21	0.08	0.77	0.13	0.11	-0.03	-0.05	0.09	0.69
2007	6000	0.16	0.10	0.01	0.22	0.08	0.76	0.13	0.10	-0.03	-0.03	0.09	0.65
2008	5331	0.14	0.10	0.00	0.13	0.07	0.75	0.11	0.09	-0.03	-0.13	0.08	0.69
2009	5218	0.15	0.10	0.01	0.13	0.07	0.74	0.12	0.09	-0.03	-0.13	0.08	0.67
2010	6067	0.13	0.09	0.01	0.22	0.06	0.77	0.10	0.08	-0.03	-0.03	0.08	0.63
2011	8440	0.14	0.10	0.00	0.15	0.07	0.77	0.11	0.09	-0.03	-0.09	0.08	0.65
2012	11311	0.15	0.12	0.01	0.22	0.07	0.83	0.12	0.12	-0.02	-0.05	0.09	0.75
2013	11012	0.13	0.10	0.02	0.43	0.07	0.78	0.10	0.09	-0.01	0.07	0.07	0.68
2014	11602	0.13	0.09	0.01	0.34	0.07	0.75	0.10	0.08	-0.02	0.03	0.07	0.70
2015	12059	0.13	0.10	0.01	0.33	0.07	0.79	0.10	0.09	-0.03	-0.01	0.08	0.72
2016	13504	0.13	0.11	0.01	0.35	0.07	0.80	0.11	0.11	-0.01	0.07	0.07	0.76
2017	13344	0.13	0.11	0.01	0.30	0.07	0.80	0.11	0.13	-0.01	0.07	0.10	0.67
2018	10215	0.14	0.10	0.02	0.40	0.07	0.74	0.10	0.09	-0.02	0.05	0.08	0.66
2019	10497	0.12	0.09	0.00	0.20	0.06	0.80	0.09	0.10	-0.02	-0.16	0.07	0.76
2020	8920	0.12	0.08	0.00	0.21	0.06	0.71	0.09	0.08	-0.02	-0.05	0.07	0.63
Central Mediterranean													
	Reanalysis							Control					
	NDATA	MEAN	SD	MB	MNMB	RMSE	r	MEAN	SD	MB	MNMB	RMSE	r
2003	3802	0.21	0.14	-0.01	0.02	0.10	0.77	0.20	0.12	-0.03	0.00	0.13	0.64
2004	1911	0.17	0.11	0.00	0.11	0.09	0.71	0.15	0.12	-0.02	0.04	0.12	0.52
2005	3945	0.18	0.13	-0.01	0.06	0.10	0.71	0.16	0.11	-0.03	-0.03	0.12	0.60
2006	3230	0.16	0.12	-0.02	-0.04	0.10	0.70	0.15	0.13	-0.03	-0.07	0.12	0.59
2007	3806	0.19	0.13	-0.01	0.05	0.09	0.74	0.18	0.14	-0.03	-0.08	0.12	0.65
2008	4320	0.17	0.12	0.00	0.10	0.09	0.73	0.15	0.12	-0.03	-0.06	0.11	0.63
2009	3663	0.17	0.10	0.00	0.03	0.08	0.71	0.15	0.12	-0.03	-0.10	0.11	0.55
2010	4391	0.16	0.13	0.01	0.10	0.08	0.78	0.14	0.13	-0.02	-0.03	0.10	0.69
2011	7264	0.16	0.11	-0.01	0.01	0.07	0.77	0.13	0.13	-0.03	-0.17	0.11	0.63
2012	5935	0.16	0.11	0.00	0.06	0.07	0.79	0.13	0.11	-0.02	-0.08	0.09	0.68
2013	7400	0.18	0.13	0.02	0.26	0.09	0.75	0.15	0.13	-0.01	0.03	0.10	0.65
2014	7744	0.17	0.14	0.02	0.28	0.09	0.79	0.15	0.14	-0.01	0.08	0.10	0.74
2015	8690	0.18	0.13	0.03	0.29	0.09	0.76	0.14	0.13	-0.01	0.00	0.10	0.65
2016	10074	0.18	0.15	0.03	0.27	0.10	0.81	0.16	0.16	0.01	0.11	0.10	0.77
2017	10203	0.17	0.13	0.02	0.24	0.08	0.80	0.15	0.14	0.00	0.02	0.10	0.72
2018	8769	0.20	0.15	0.01	0.17	0.09	0.80	0.18	0.16	0.00	0.06	0.12	0.71
2019	8422	0.18	0.14	0.02	0.18	0.09	0.81	0.16	0.17	0.00	-0.05	0.10	0.79
2020	6678	0.13	0.09	0.00	0.06	0.07	0.71	0.12	0.09	-0.02	-0.10	0.13	0.65
Eastern Mediterranean													
	Reanalysis							Control					
	NDATA	MEAN	SD	MB	MNMB	RMSE	r	MEAN	SD	MB	MNMB	RMSE	r
2003	3945	0.24	0.13	0.04	0.38	0.11	0.74	0.23	0.16	0.02	0.31	0.13	0.63
2004	4715	0.22	0.12	0.02	0.25	0.10	0.69	0.23	0.16	0.02	0.25	0.13	0.59
2005	3856	0.22	0.11	0.01	0.23	0.10	0.64	0.21	0.13	0.00	0.18	0.13	0.49
2006	4178	0.23	0.13	0.02	0.22	0.10	0.72	0.22	0.14	0.01	0.16	0.12	0.64
2007	3253	0.20	0.13	0.01	0.20	0.09	0.78	0.20	0.15	0.01	0.23	0.13	0.60
2008	4529	0.22	0.13	0.01	0.24	0.10	0.76	0.19	0.13	-0.02	0.08	0.12	0.63
2009	5439	0.21	0.12	0.02	0.29	0.10	0.64	0.18	0.12	-0.01	0.06	0.11	0.53
2010	7867	0.23	0.13	0.01	0.17	0.10	0.75	0.20	0.15	-0.02	0.02	0.14	0.59
2011	8388	0.23	0.12	0.00	0.15	0.11	0.62	0.17	0.13	-0.06	-0.16	0.16	0.42
2012	7225	0.24	0.13	0.02	0.29	0.12	0.68	0.20	0.15	-0.01	0.10	0.15	0.49
2013	7122	0.24	0.13	0.03	0.36	0.13	0.70	0.20	0.15	-0.02	0.07	0.15	0.60
2014	6434	0.22	0.12	0.02	0.34	0.11	0.72	0.21	0.16	0.00	0.25	0.15	0.57
2015	7375	0.23	0.15	0.01	0.22	0.11	0.79	0.20	0.16	-0.01	0.07	0.15	0.62
2016	8903	0.20	0.13	0.01	0.23	0.10	0.72	0.17	0.14	-0.02	0.05	0.13	0.60
2017	7877	0.21	0.12	0.01	0.21	0.10	0.73	0.18	0.14	-0.02	0.00	0.12	0.62
2018	6954	0.24	0.15	0.01	0.21	0.11	0.73	0.22	0.17	-0.01	0.10	0.14	0.63
2019	4598	0.23	0.13	0.02	0.28	0.11	0.70	0.23	0.20	0.02	0.17	0.16	0.60
2020	5229	0.20	0.11	0.02	0.25	0.09	0.69	0.19	0.11	0.01	0.14	0.13	0.57

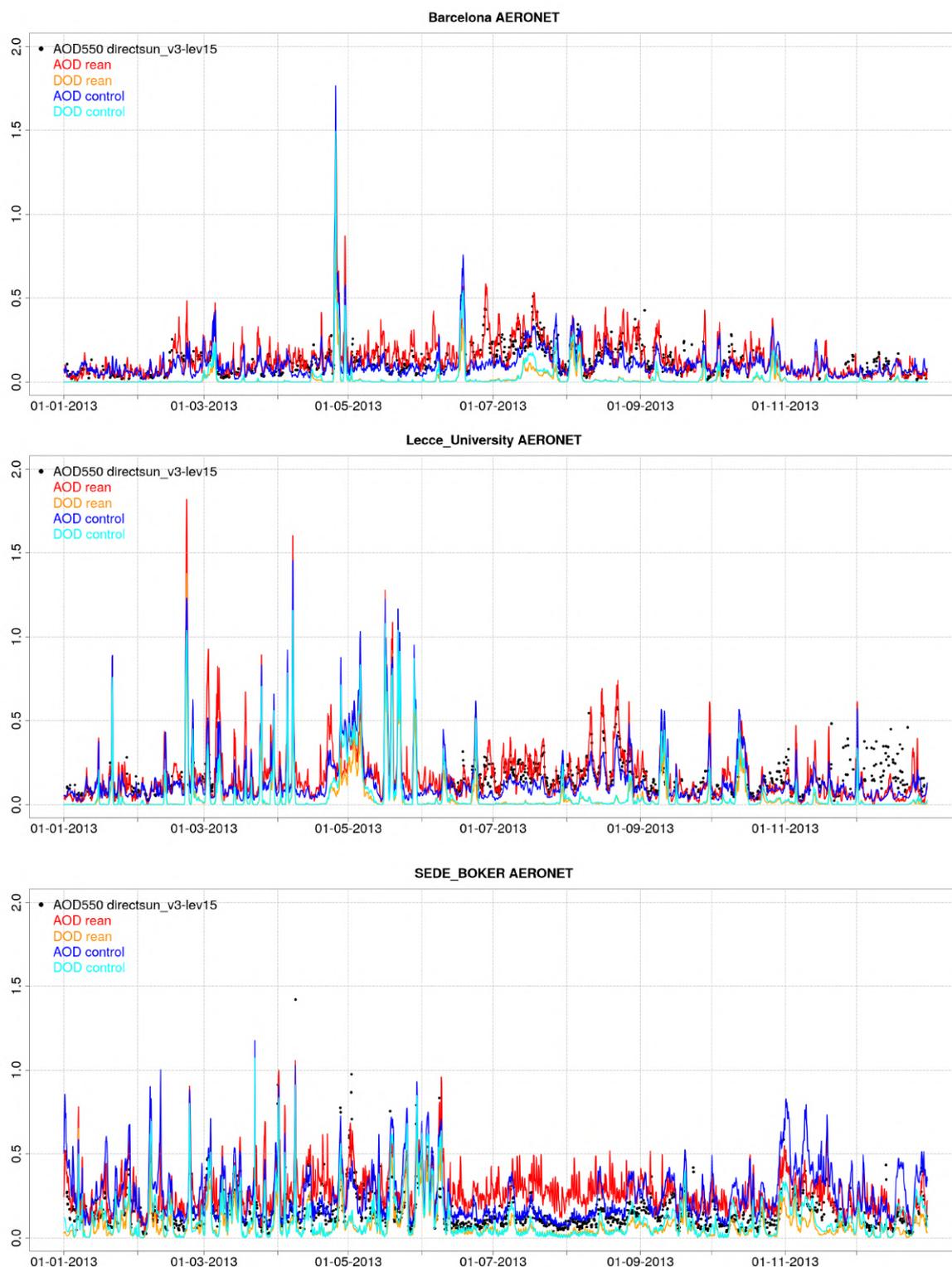


Figure 3.3.2: AOD and AE Level 1.5 Direct-Sun AERONET observations (dots), as well as CAMS AOD and DOD reanalysis and control run over Barcelona (Spain, Western Mediterranean), Lecce University (Italy, Central Mediterranean) and Sede Boker (Israel, Eastern Mediterranean) for the year 2013.

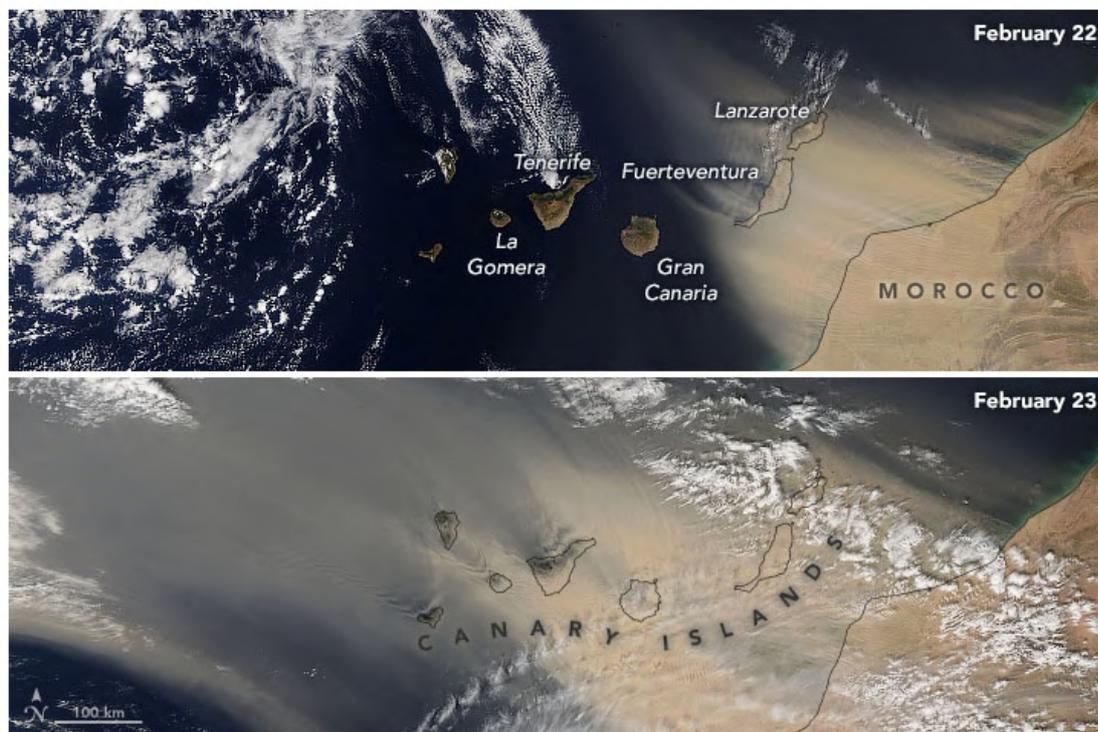


Figure 3.4.1: MODIS images on 22-23 February 2020. Source NASA Earth Observatory.

### 3.4 Exceptional dust event over the Canary Islands in late-February 2020

In late February 2020, strong Saharan winds picked up dust from Africa and carried it over the Canary Islands, achieving surface concentrations over  $3000 \mu\text{g}/\text{m}^3$  and reducing visibility to less than 10m. Due to strong winds and poor visibility, all airports across the Canary Islands were closed on February 22, and most stayed closed until February 24 (Cuevas et al., 2021). Close to 800 flights were cancelled or re-routed. Some roads were also shut down due to limited visibility. Schools and universities were closed on February 24, and people were advised to keep their windows shut and to stay indoors due to poor air quality. Some Carnival events were postponed or curtailed due to the dusty conditions. The strong winds spread several wildfires in Tenerife and Gran Canaria (the largest island in the archipelago).

MODIS satellite detected an intense outbreak of dust over Canary Islands (see Figure 3.4.1 and Figure 3.4.2) leading to maximum AOD values up to 5. The event is originating in Mauritania on 22<sup>nd</sup> February in the early morning. Dust was transported towards the West achieving Canary Islands a few hours later (see Santa Cruz de Tenerife in Fig. 3.4.3).

The event was associated with the presence of some clouds as shown in Figure 3.4.1. The comparison with the MODIS AOD product show that the CAMS reanalysis AOD reproduces the spatial distribution and timing of the dust plume as shown by the comparison with MODIS/Aqua (see Figure 3.4.2) and AERONET (see Santa Cruz de Tenerife in Figure 3.4.3). However, the comparison shows that the CAMS reanalysis underestimates the aerosol concentrations during this event, reaching AOD values up to 1.5 and PM<sub>10</sub> up to  $300 \mu\text{g}/\text{m}^3$  in over the Canary Islands.

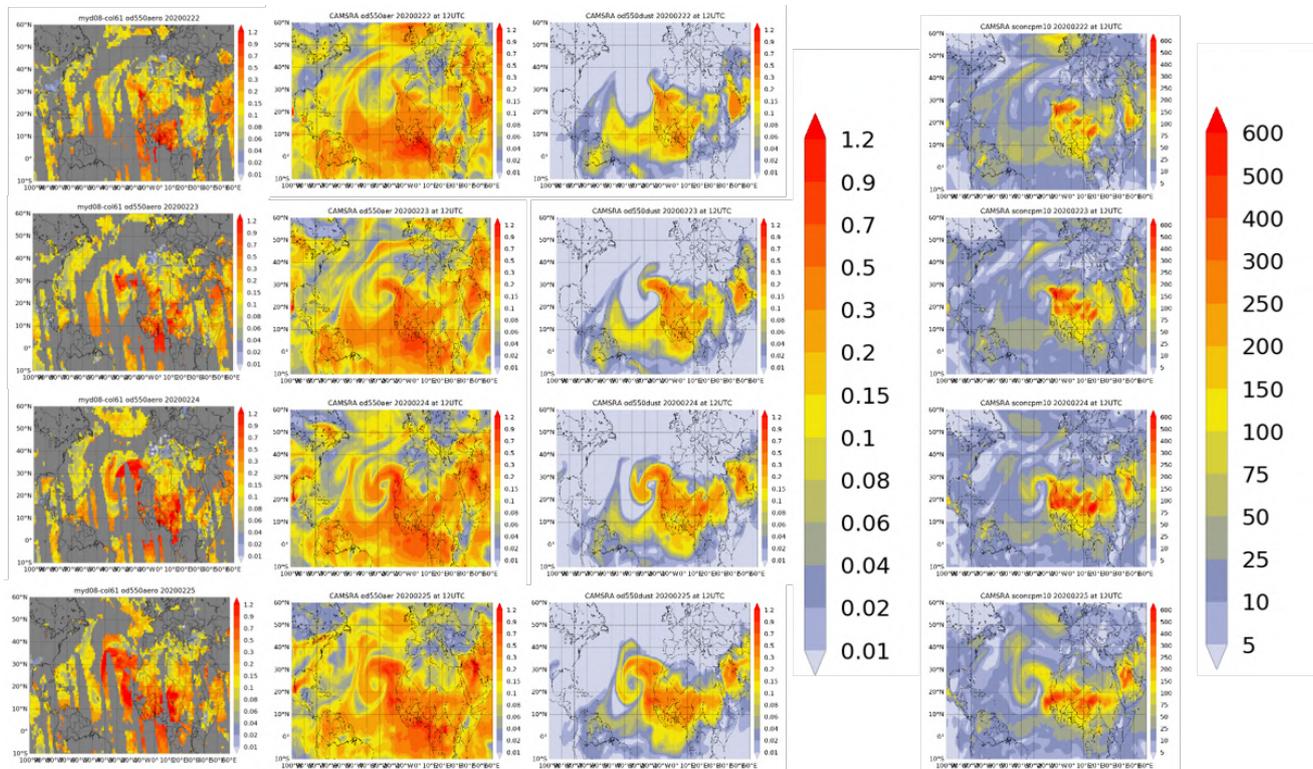


Figure 3.4.2: Daily composite of NASA MODIS/Aqua well as AOD, DOD and PM10 at 12UTC from CAMSRA for 22-25 February 2020.

### 3.5 ‘Godzilla’ dust event in June 2020

Each year, dust from the Sahara Desert blows off Africa and across the Atlantic. There was much interest in the media for the spectacular African dust event during late June that affected a huge area of the tropical Atlantic and, in the next days, the Caribbean and Gulf of Mexico.

During June 2020, the unusually large quantity of dust remaining after this plume's trans-Atlantic voyage earned it the nickname “Godzilla” as well as the title of the dustiest such event in the 20 years that scientists have kept records of these storms. The records of mineral dust concentrations in Barbados since the late 1960s shows the exceptionality of this event, as it is shown in the comments of Prof. Joseph Prospero published in the WMO SDS-WAS Regional Center (<https://sds-was.aemet.es/news/godzilla-dust-event>).

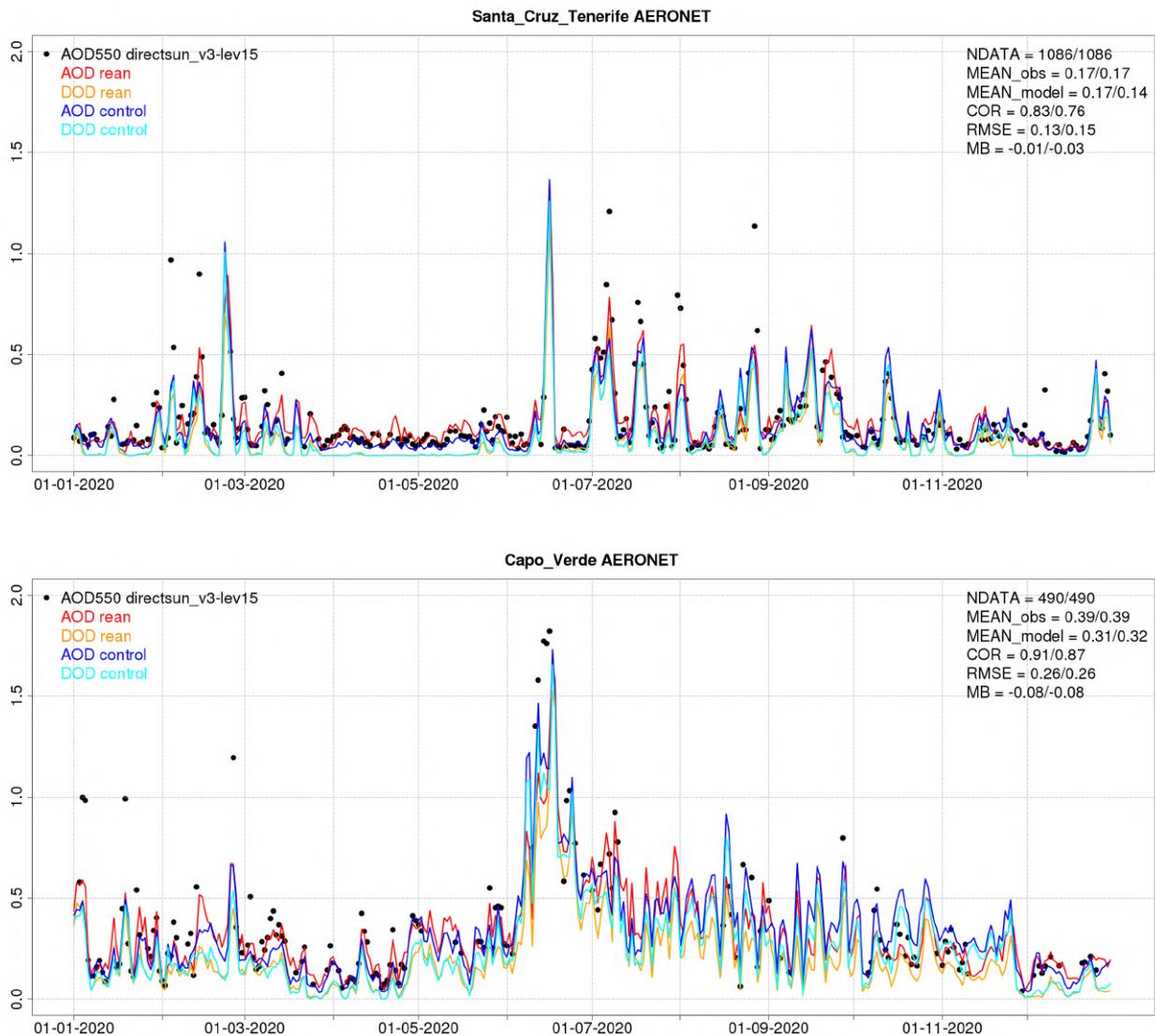


Figure 3.4.3: Daily averages of AOD from AERONET (black dot), AOD CAMSRA (red line), AOD control (blue line), AOD-Nat CAMSRA (orange line), AOD-Nat control (cyan line), for the study period over Santa Cruz de Tenerife (Canary Islands, Spain) and Capo Verde. AOD-Nat corresponds to the natural aerosol optical depth that includes dust and sea-salt. Skill scores per each site and model (CAMSRA/control) are shown in the upper right corner (NDATA: available 3-hourly values used for the calculations, MEAN observations, MEAN model, COR, RMSE, MB).

Satellites tracked how the plume swept out from the west coast of Africa over the Atlantic Ocean on June 17, 2020 (see Figure 3.5.1) with AOD values above 1.5. Within days it extended across the Atlantic, degrading air quality and lowering visibility in the Caribbean with hourly PM<sub>10</sub> peaks above 150 µg/m<sup>3</sup> in the EEA Guadeloupe site (not shown here). It reached the south-east of the United States on June 24 (see Figure 3.5.1), resulting in an air pollution event lasting several days that led to the U.S. Environmental Protection Agency’s (EPA) air quality index (AQI) reaching code orange (unhealthy for sensitive groups) for Florida.

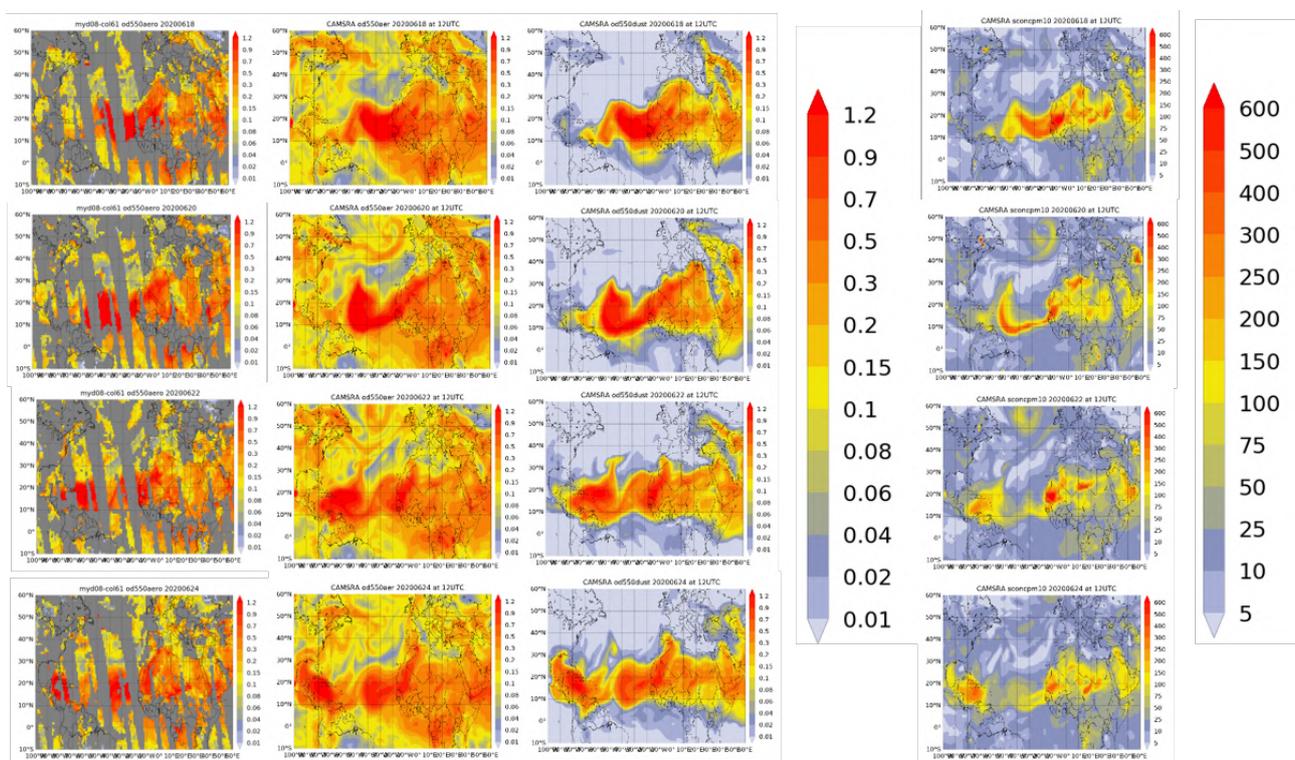


Figure 3.5.1: Observed AOD from MODIS/Aqua Level 3 and modelled AOD, DOD and PM10 (in  $\mu\text{g}/\text{m}^3$ ) at 12UTC from CAMSRA for 18, 20, 22 and 24 June 2020.

The AOD values above 1.2 across the Atlantic and PM10 values above  $400 \mu\text{g}/\text{m}^3$  are well predicted by the CAMS reanalysis in the Caribbean, with maximum values up to  $700 \mu\text{g}/\text{m}^3$  in the outflow of the Senegal coast. The comparison with MODIS shows how CAMSRA reproduces the spatial and temporal distribution of the observed AOD plume. The main contributor to the predicted AOD by the reanalysis is the presence of mineral dust as confirmed by Figure 3.5.1. The comparison in Capo Verde with AERONET AOD (see Figure 3.4.3) shows how the reanalysis did timely reproduce the distribution of the main dust plume despite the model tendency to underestimate the observed maximum values.



## 4. Tropospheric Ozone

### 4.1 Validation with sonde data in the free troposphere

Model profiles of the CAMS reanalysis runs were compared to free tropospheric balloon sonde measurement data of 38 stations taken from the NDACC, WOUDC, NILU and SHADOZ databases for January 2003 to December 2020. The methodology for model comparison against the observations is described in Eskes et al., 2018. The free troposphere is defined as the altitude range between 750 and 200 hPa in the tropics and between 750 and 350 hPa elsewhere, and the UTLS between 100 and 60 hPa in the tropics, and between 100 and 300 hPa elsewhere.

In the free troposphere, ozone mixing ratios are underestimated over all regions until September 2004, except over the Tropics. In the years 2005 to 2012, MNMBs are  $\pm 10\%$  for the Arctic and for the Northern midlatitudes and up to  $\pm 20\%$  for Antarctica. From 2013 onwards, MNMBs get larger, reaching up to  $\pm 38\%$  over Antarctica and up to  $\pm 25\%$  over Arctic stations. This change in bias relates to the switch from SBUV/2 13l to SBUV/2 21l data. Before July 2013 a 13-layer product was assimilated, after 9 July 2013 a 21-layer product (see Table 2.2). Unfortunately, this had an impact on the ozone bias in the troposphere (Figs 4.1.1. and 4.1.2). The same behaviour in tropospheric ozone is seen in the CAMS near-real-time analysis, which also assimilates the 21L data from May 2013 onwards. Larger biases during DJF 2120/2013 over Antarctica also appear due to the loss of ENVISAT data in the assimilation after 2012 (Fig. 4.1.2).

Over the Tropics, ozone mixing ratios are mostly slightly overestimated throughout the whole period 2003 to 2020 with maximum MNMBs of 30% in early 2003, and late 2018, see Fig. 4.1.1. The improvement observed since the second half of 2004 may be linked to the availability of Aura MLS ozone profile data. Towards the end of the period, MNMBs increase slightly (Fig. 4.1.1). It is unclear if this also relates to the changes in ECC sondes from 2014 onwards (Stauffer et al., 2020), see also the stratospheric section.

#### Impact of data analysis from the comparison with the control run

Except over the Tropics, the control run has generally lower O<sub>3</sub> mixing ratios as the CAMS reanalysis over all regions and thus mostly negative MNMBs. Over Antarctica, for DJF the loss of ENVISAT and the switch in SBUV data lead to large positive MNMBs whereas the control run remains stable. Over the Arctic, changes in MNMBs show up during the summer season in JJA (Fig. 4.1.3).

Apart from these inconsistencies, the seasonal-time series remain mostly stable. However, over the Arctic during JJA, a positive trend in bias is observed in both CAMS runs. The reason for this trend needs to be further investigated.

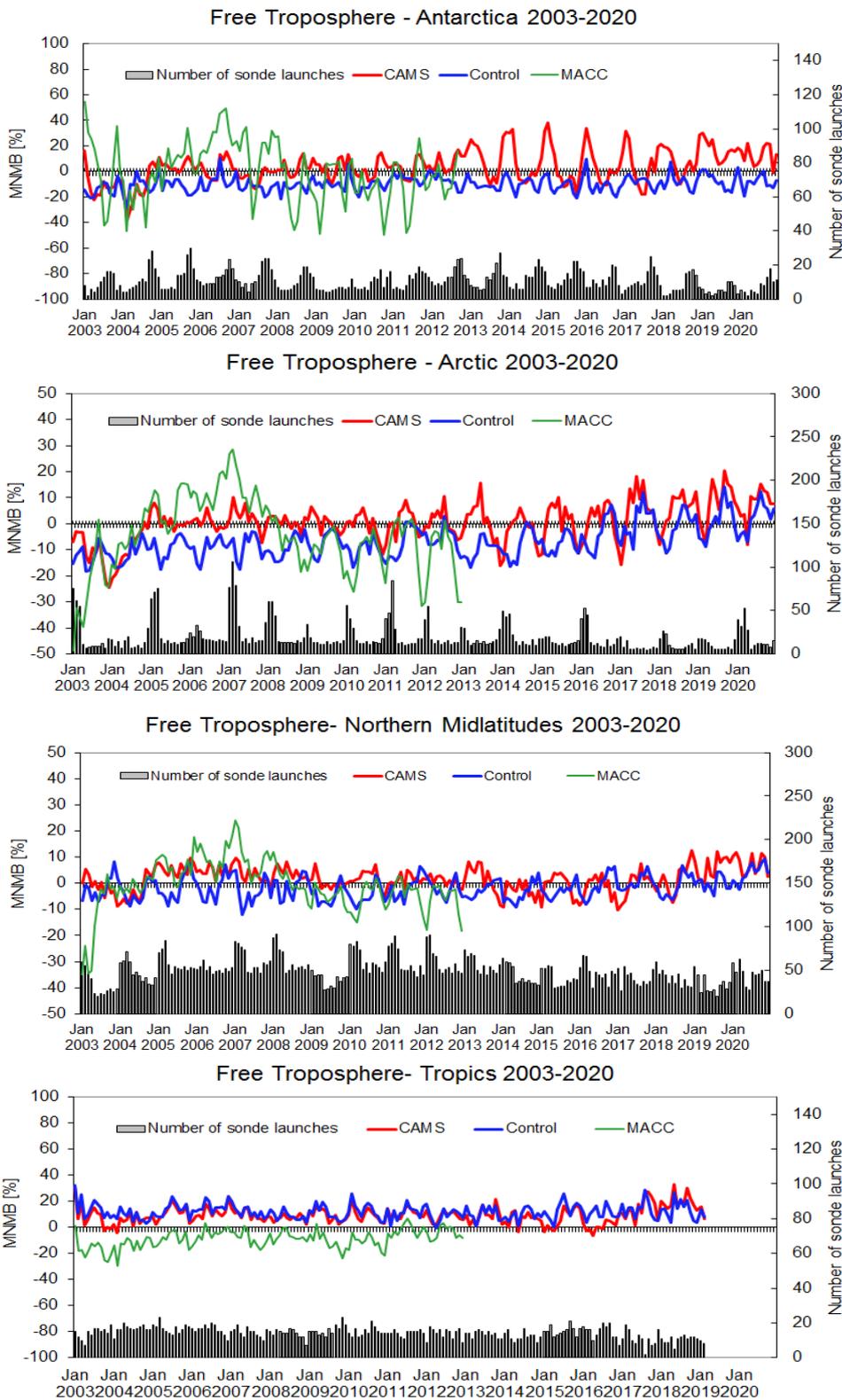


Figure 4.1.1: Comparison with ozone sondes in the free troposphere, period 2003-2020: MNMBs for 4 regions (first row left: Antarctica, first row right: Arctic, second row left: Northern Midlatitudes, second row right: Tropics) in the free troposphere. Red: CAMS reanalysis, green: MACC reanalysis, blue: Control run.

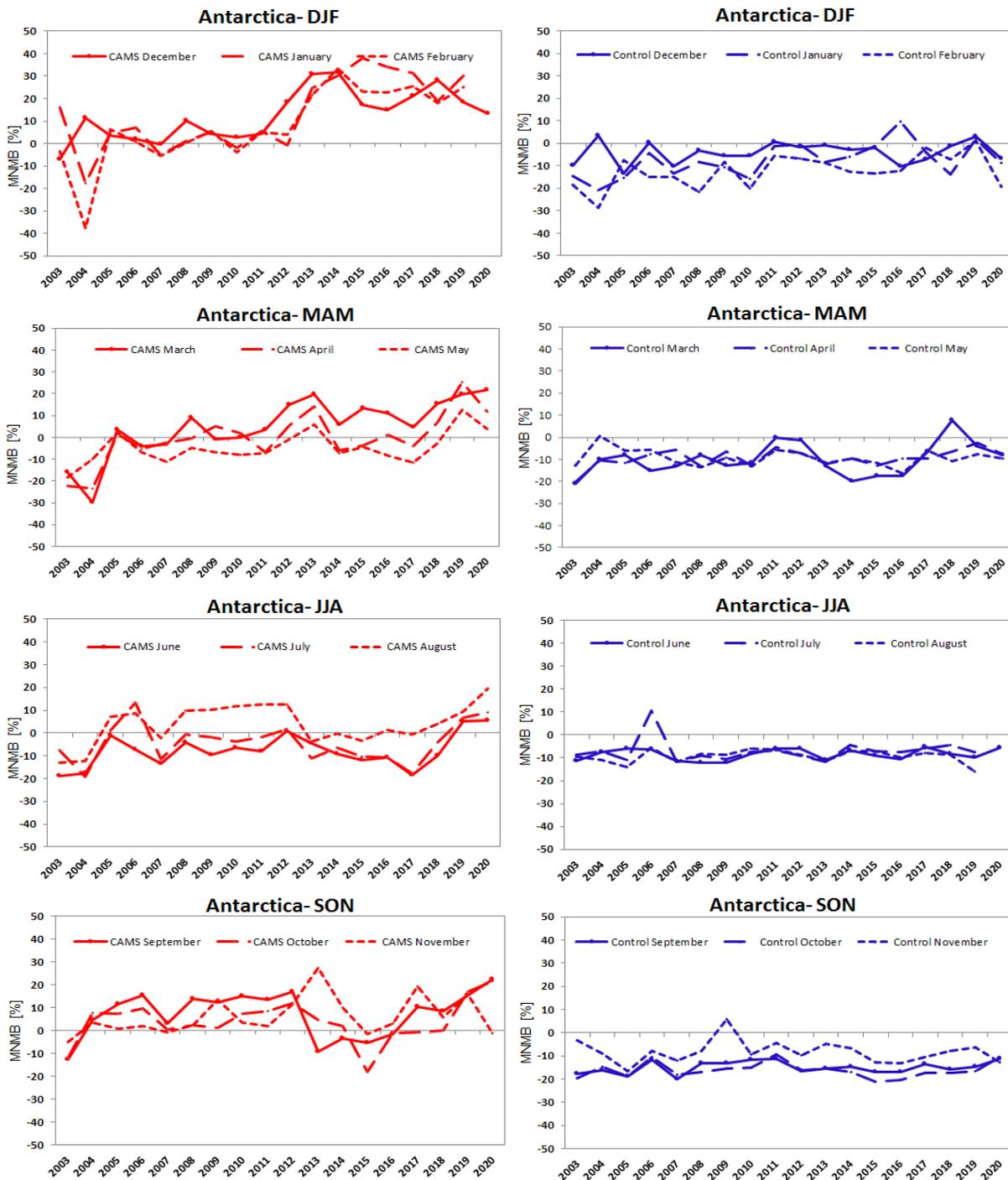


Figure 4.1.2: seasonal time-series (first row: DJF, second row: MAM, third row: JJA, fourth row: SON) of MNMBs from the comparison with ozone sondes in the free troposphere over Antarctica, period 2003-2020 (left: CAMS reanalysis in red, right: control in blue)

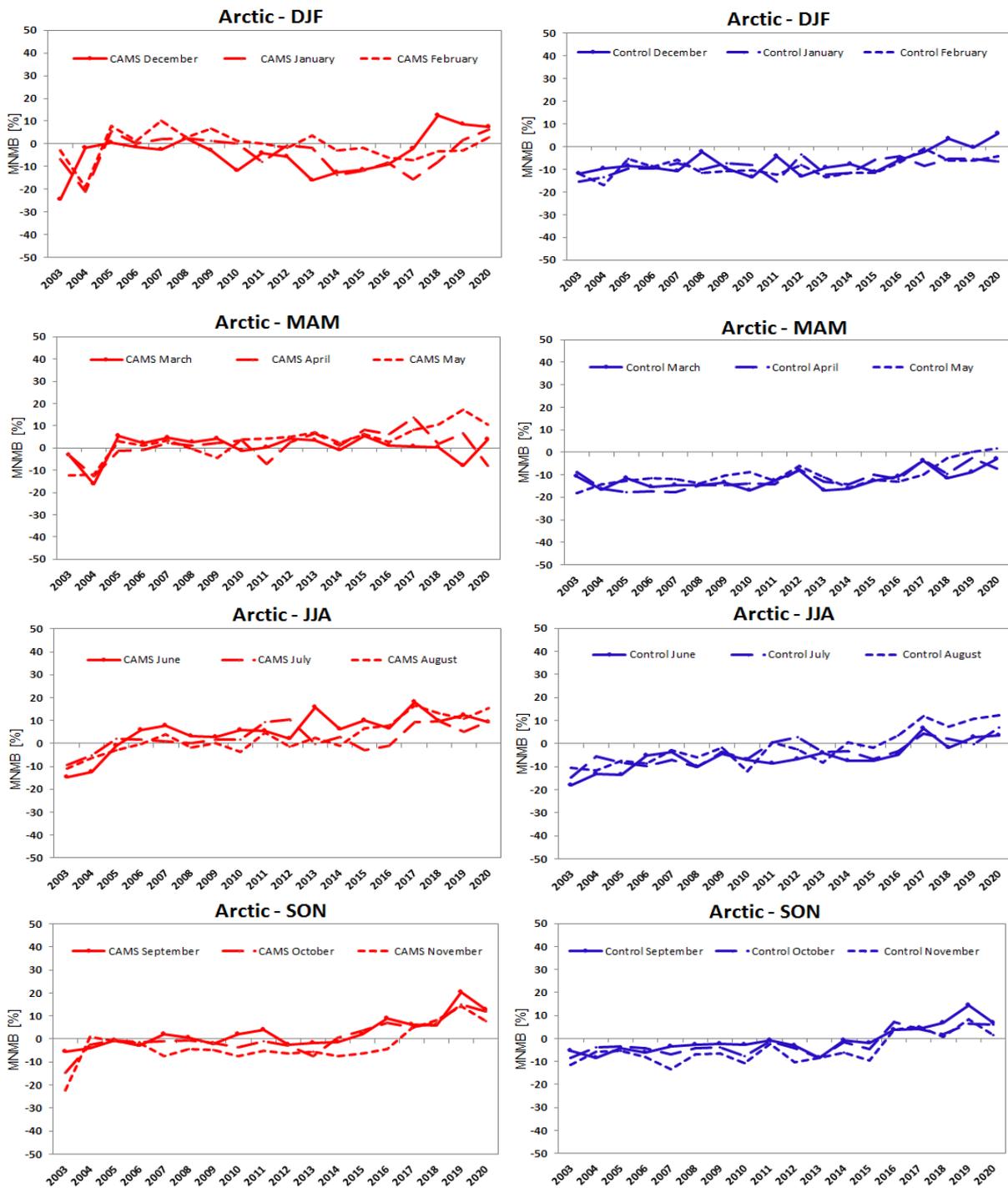


Figure 4.1.3: Same as 4.1.2 over the Arctic

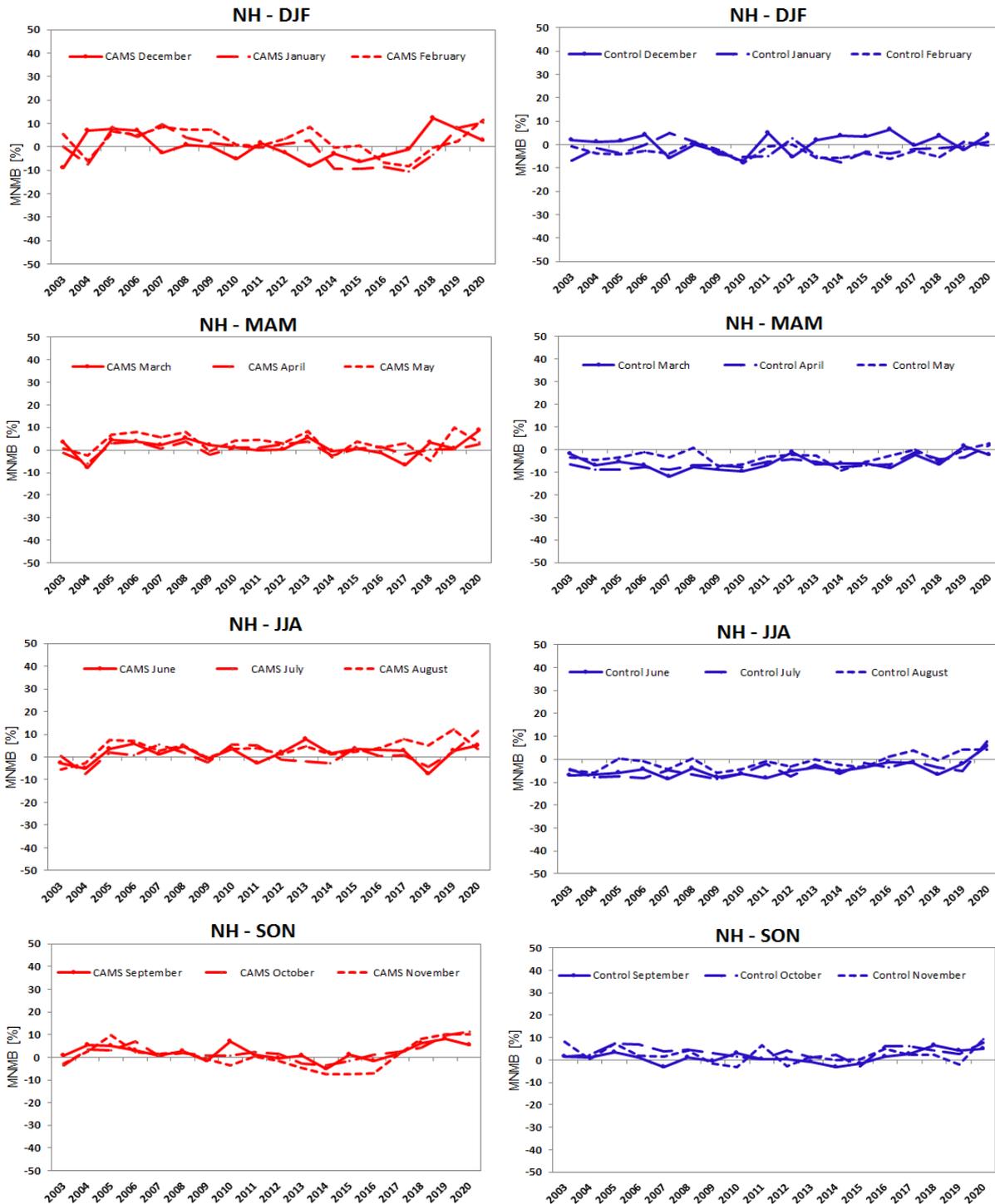


Figure 4.1.4: Same as 4.1.2 over the Northern Midlatitudes

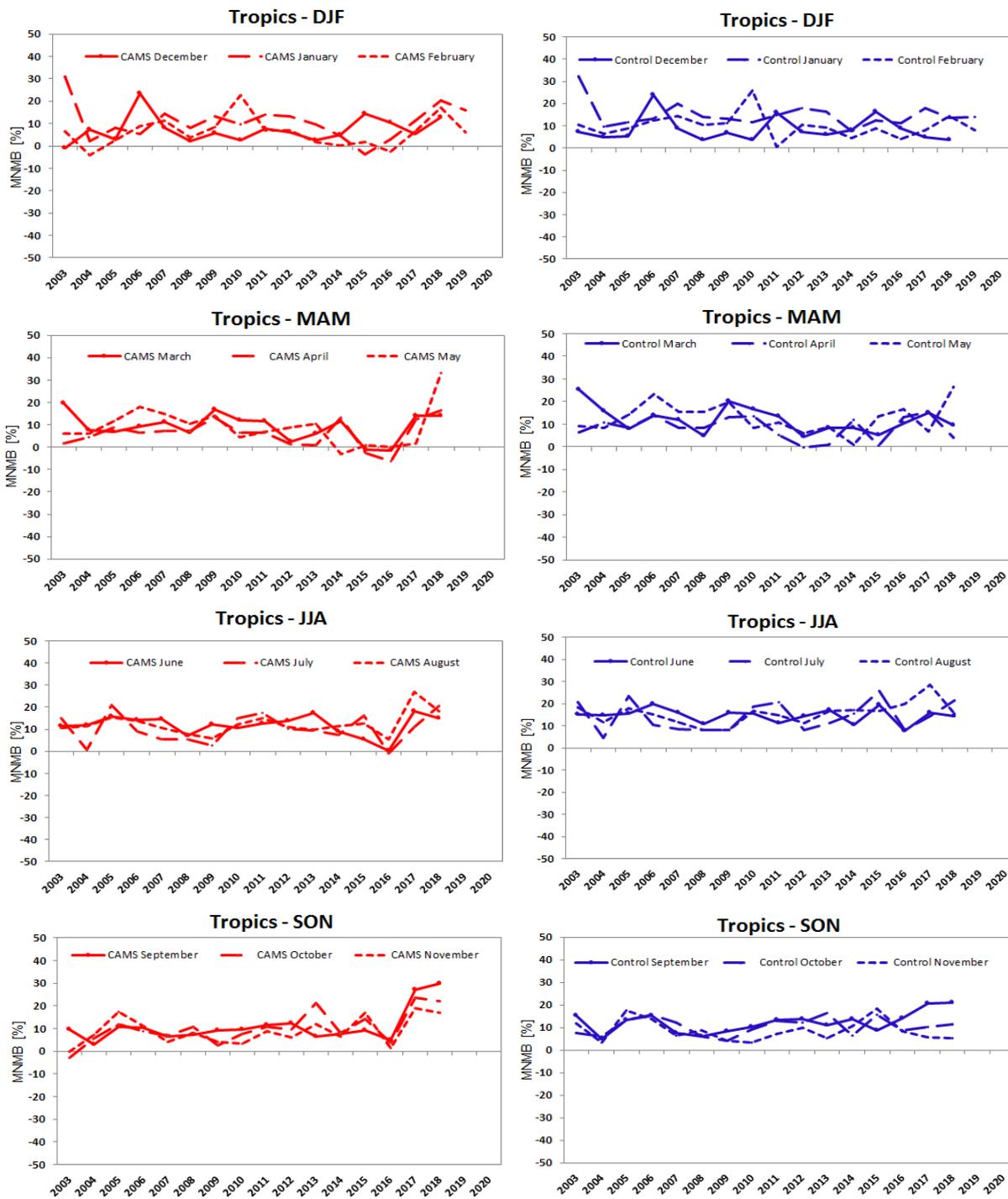


Figure 4.1.5: Same as 4.1.2 over the Tropics

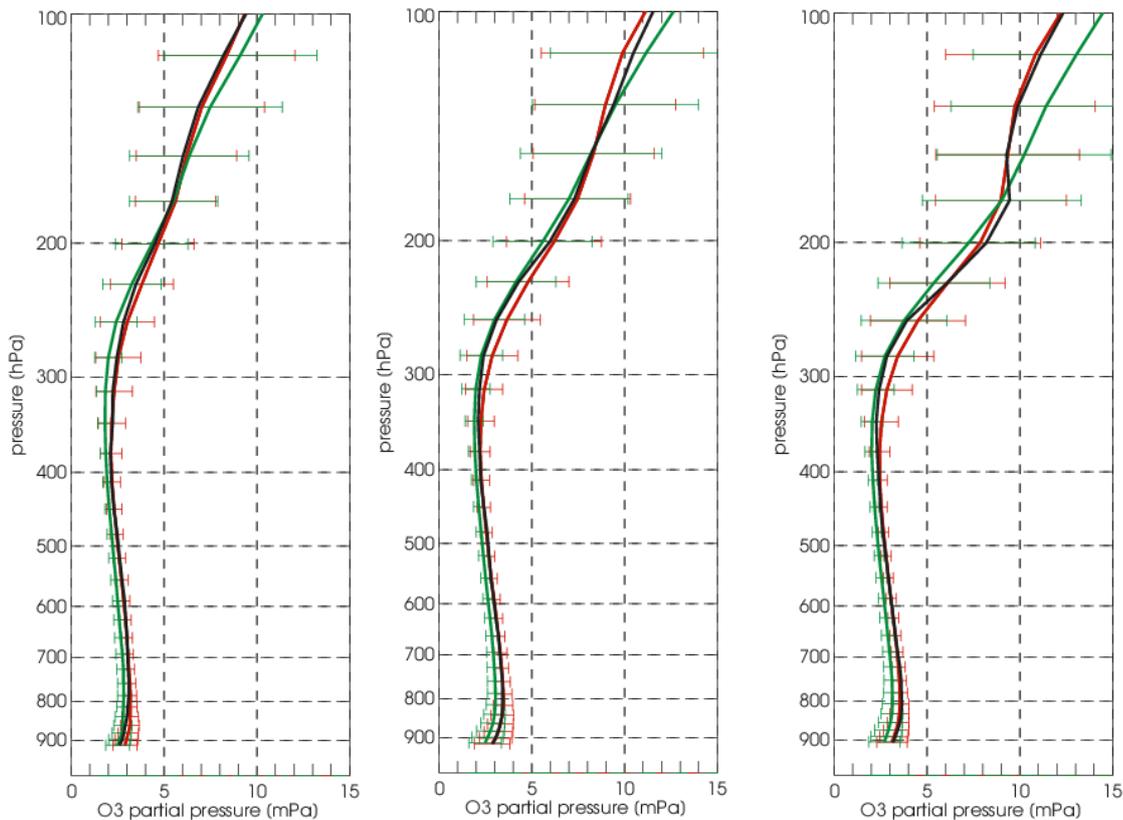


Figure 4.1.4: Ozone profiles for December 2009, January 2010 and February 2010 for the Northern Midlatitudes (free troposphere and UTLS), CAMS reanalysis (red), MACC reanalysis (green), ozone sondes (black).

#### Comparison with the MACC reanalysis.

The MACC reanalysis shows larger biases for all regions and a larger variability of the bias with time. Monthly MNMBs in the free troposphere are  $\pm 35\%$  over the Northern midlatitudes,  $\pm 50\%$  over Arctic and Antarctica and  $\pm 25\%$  over the Tropics.

#### Profiles.

Over the northern mid-latitudes, the CAMS reanalysis provides an accurate description of the average monthly-mean profiles, an improvement compared to the MACC reanalysis, see Fig. 4.1.4 and 4.1.5.

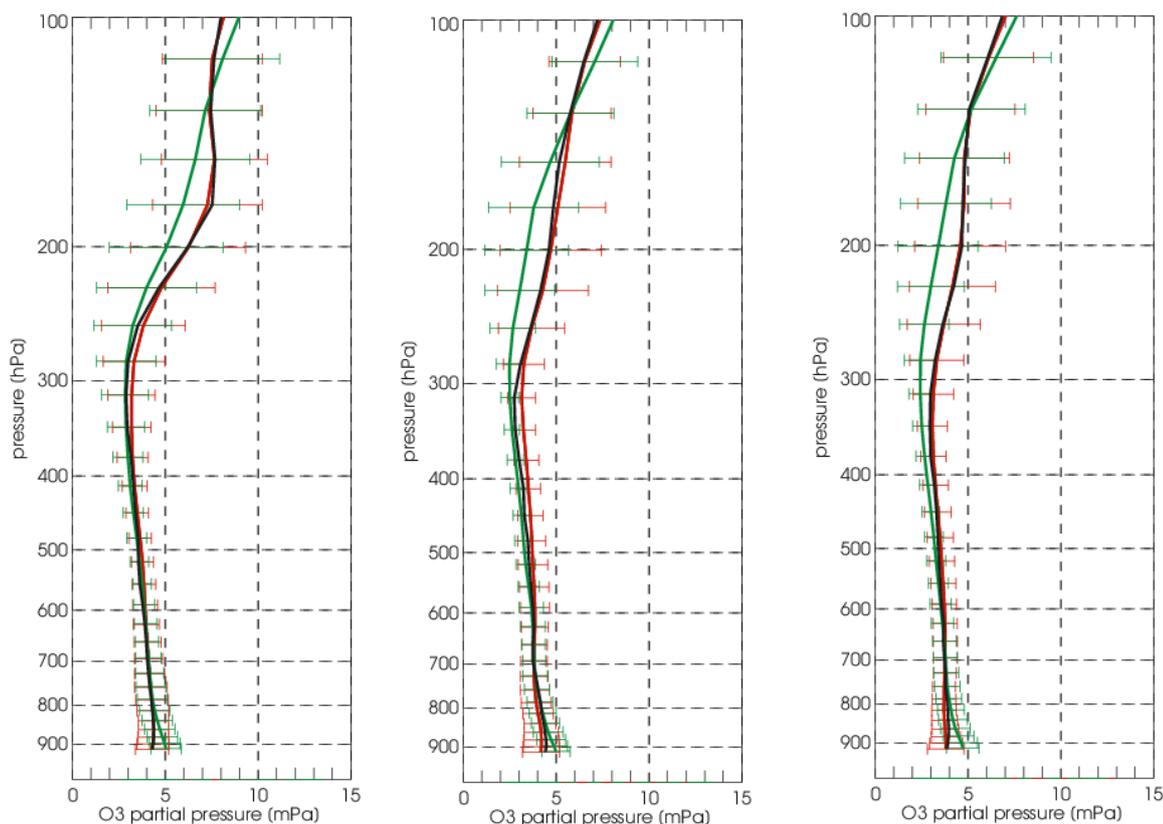


Figure 4.1.5: Ozone profiles for June 2010, July 2010 and August 2010 for the Northern Midlatitudes (free troposphere and UTLS), CAMS reanalysis (red), MACC reanalysis (green), ozone sondes (black).

## 4.2 Validation with GAW surface ozone observations

In the following, an evaluation of model performance for the CAMS reanalysis has been carried out for the period January 2003 to December 2020. GAW hourly data from 42 stations from the World Data Centre for Greenhouse Gases (WDCGG) has been used for model verification for the years 2003-2015, see Table 4.2.1. For the years 2018-2020 near-real-time station data has been used additionally (also included in Table 4.2.1). MNMBs have been calculated on a yearly and monthly basis between 2003 and 2020.

The validation shows that the reanalysis can reproduce well the observed  $O_3$  mixing ratios, see Fig. 4.2.1. Monthly MNMBs for the Northern midlatitudes range between -38% and 26%. For Antarctica MNMBs are within  $\pm 44\%$ , respectively (except for 2004). Larger MNMBs (up to 75%) appear for Arctic stations during spring (ozone depletion events). For the Tropics, monthly MNMBs are within -6 and 51%.

Fig. 4.2.1 show the monthly time series of MNMBs. Surface ozone of Antarctic stations is mostly slightly underestimated during the Antarctic spring season (August-October), whereas the model overestimates the maximum of ozone during the Antarctic winter season (July-September), Fig. 4.2.2. From 2012 onwards inconsistencies in the time-series for the CAMS reanalysis show up due to the switch in SBUV/2 data and also potentially influenced by the loss of ENVISAT data and (see

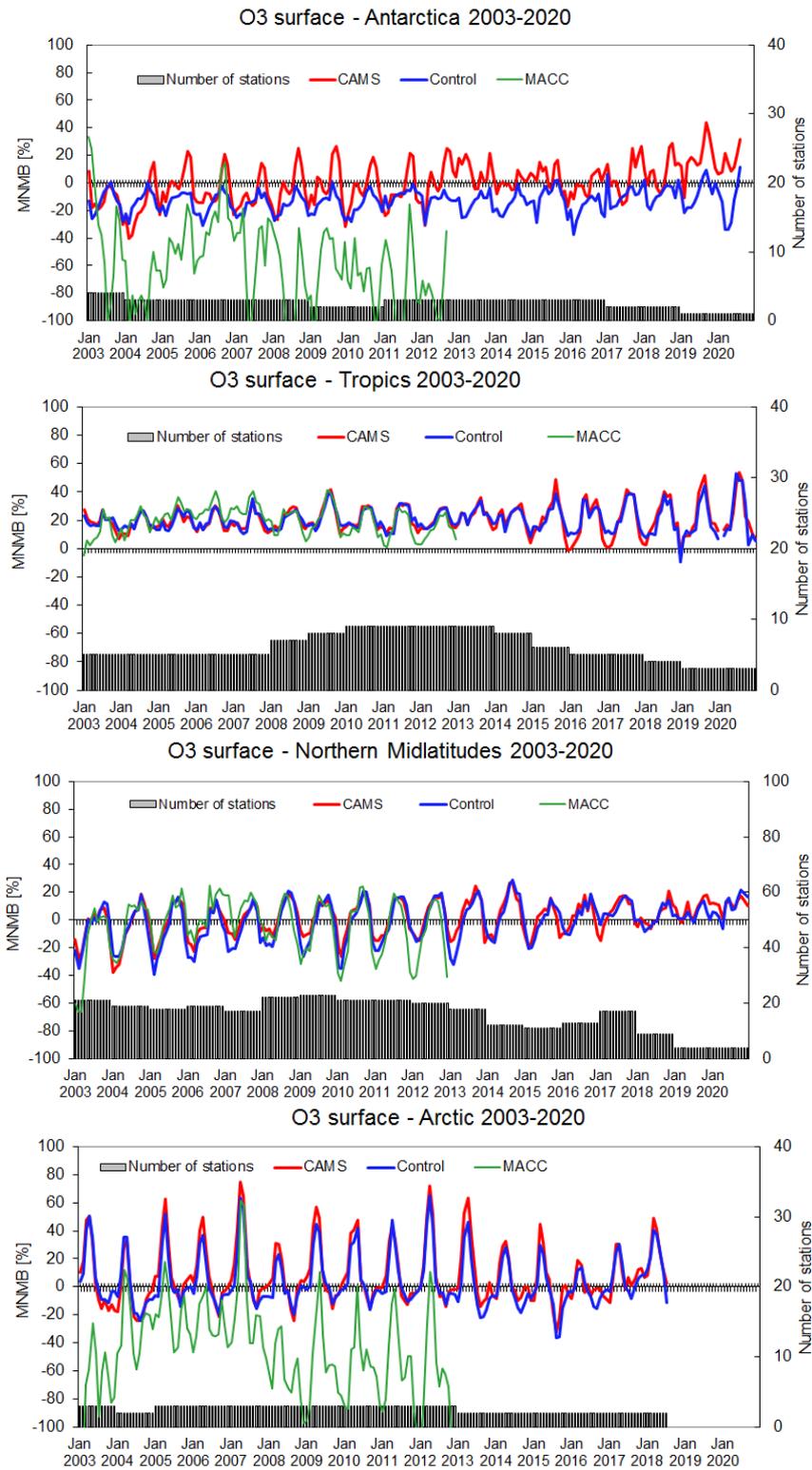


Figure 4.2.1: Mean Normalised bias (MNMB) for GAW stations calculated for the years 2003-2020 from 3-hourly data, averaged on a monthly basis, averaged over the zonal belts (first row left: Antarctica, first row right: Arctic, second row left: Northern Midlatitudes, second row right: Tropics). Red: CAMS reanalysis, green: MACC reanalysis, blue: Control run. The availability of observations reduces towards the end of the period.



stratospheric section). The control runs remains stable, except during DJF towards the end of the period due to a limited number of stations in recent years. After 2013, with the change of SBUV data, the O<sub>3</sub> maxima are reduced. From the year 2019, the discrepancy between the reanalysis and control run increased. In the year 2020, the control run shows an increase of the negative bias in the spring season.

For Arctic stations, MNMBs are positive during winter and negative during summer. The changes in the data assimilation after 2012 show up in the months February and June.

For stations of the Northern midlatitudes, positive biases can be noted during the summer season and negative biases during the winter season. The seasonal time-series is very stable for SON, during other quarters; a positive trend appears in both, the CAMS reanalysis and control run. Starting from the year 2015, both model runs show a smaller bias compared to the observations.

For tropical stations, both model runs show positive biases all year round. For DJF, data assimilation leads to a drop in MNMBs after 2012. For JJA and SON, there is a positive trend in MNMBs for both, the CAMS reanalysis and control run.

For Asian stations, the model shows an overestimation of the minimum O<sub>3</sub> mixing ratios especially during the summer monsoon season when observed surface ozone is low. This effect increases with time.

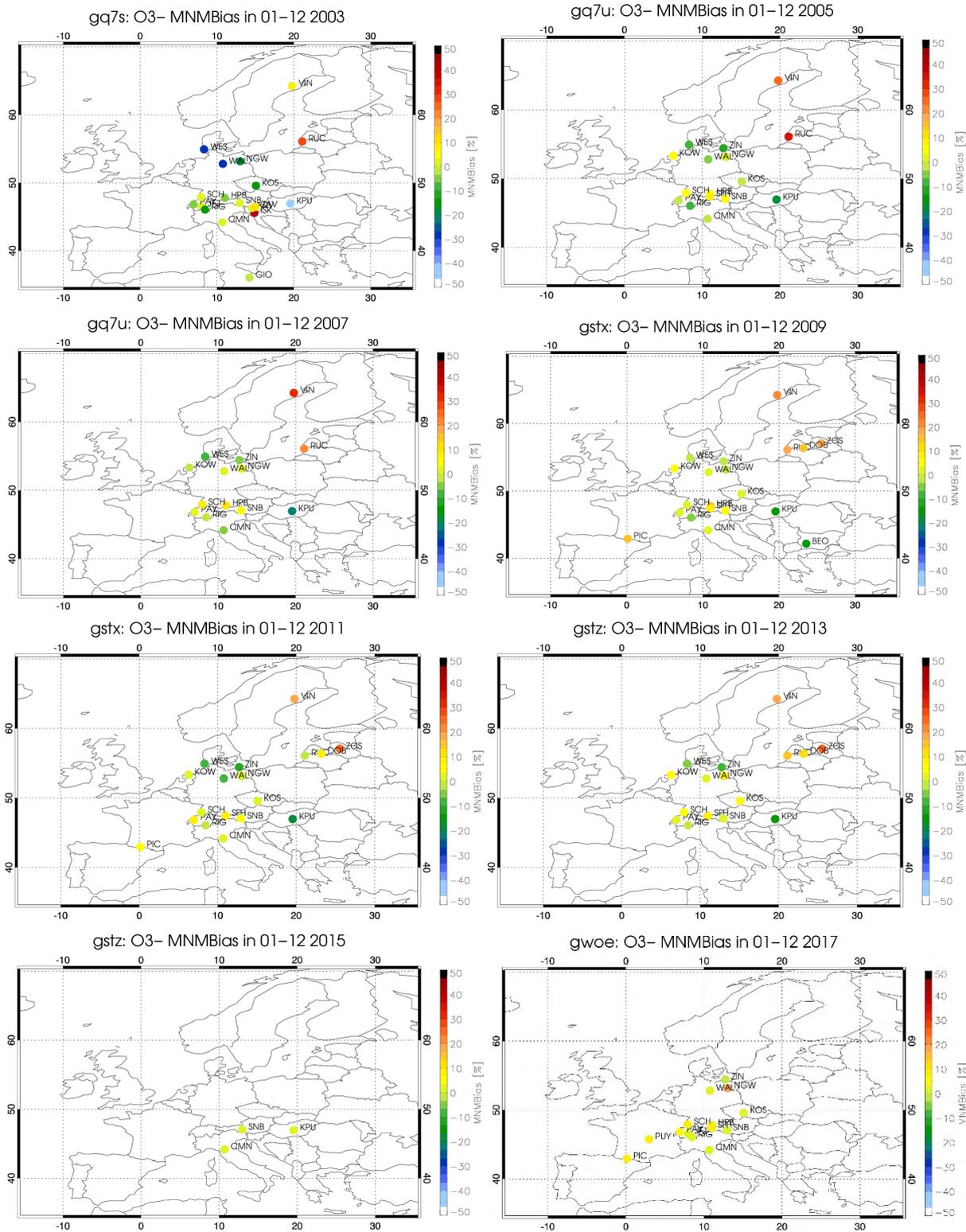
Figure 4.2.2 displays maps with annual MNMBs for Europe. Figure 4.2.3 shows Taylor plots with annual standard deviation and correlation.

The location of the stations plays an important role for the quality of the results in Europe. The more remote mountain stations (HPB, SNB, SCH, CMN, PAY and JFJ) show very accurate results (see also map in Fig. 4.2.2) with low MNMBs (between -8 and 13%) and high correlation (>0.7). This shows that long-range transport processes and atmospheric background in the troposphere are well represented in the reanalysis. Stations in the north (WES, WAL, NGW) and east of Europe (KPU and KOS), show larger negative biases, especially during the year 2003, see maps in Fig 4.2.6. These negative MNMBs are reduced significantly for later years. Two stations in the far north-east of Europe (RUC and VIN), show surface ozone that is strongly overestimated, with MNMBs of up to 35%.

Annual correlation coefficients are mostly over 0.7 showing the skill of the reanalysis to reproduce the variability of the observations. Lower correlation for the Arctic stations is associated with the ozone depletion events during spring. The standard deviation in the Taylor plots is relatively close to the reference point, which shows that the modelled variability of ozone is similar to the observed variability for most stations (Fig. 4.2.3).

#### Comparison with the MACC reanalysis

The MACC reanalysis in comparison shows larger and more variable MNMBs (Fig. 4.2.1), especially for the high latitude regions (Antarctica, Arctic). Relevant improvements are the change of CTM (MOZART in MACC and IFS-TM5 in CAMS), but also to changes in the data assimilation (bias correction applied to MLS in MACC).



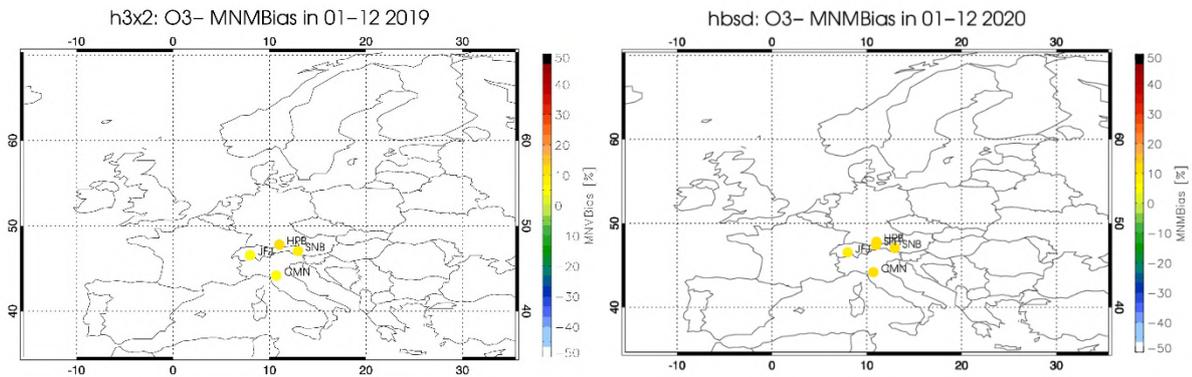
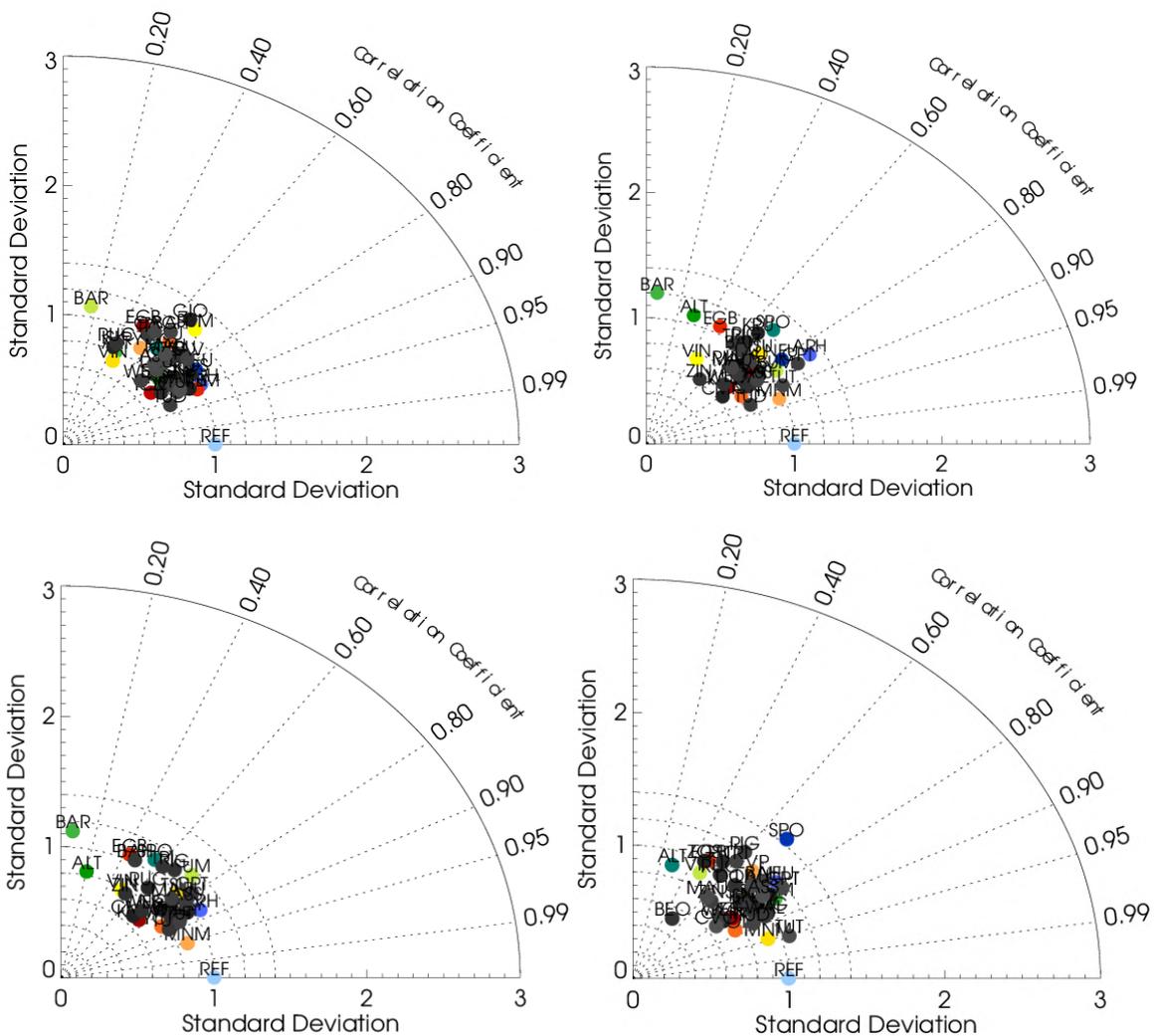


Figure 4.2.2: Map of MNMBs for the CAMS reanalysis against European GAW stations for the years 2003 (first row left), 2005 (first row right), 2007 (second row left), 2009 (second row right), 2011 (third row left), 2013 (third row right), 2015 (fourth row left), 2017 (fourth row right), and 2019 (last row left) and 2020 (last row right).



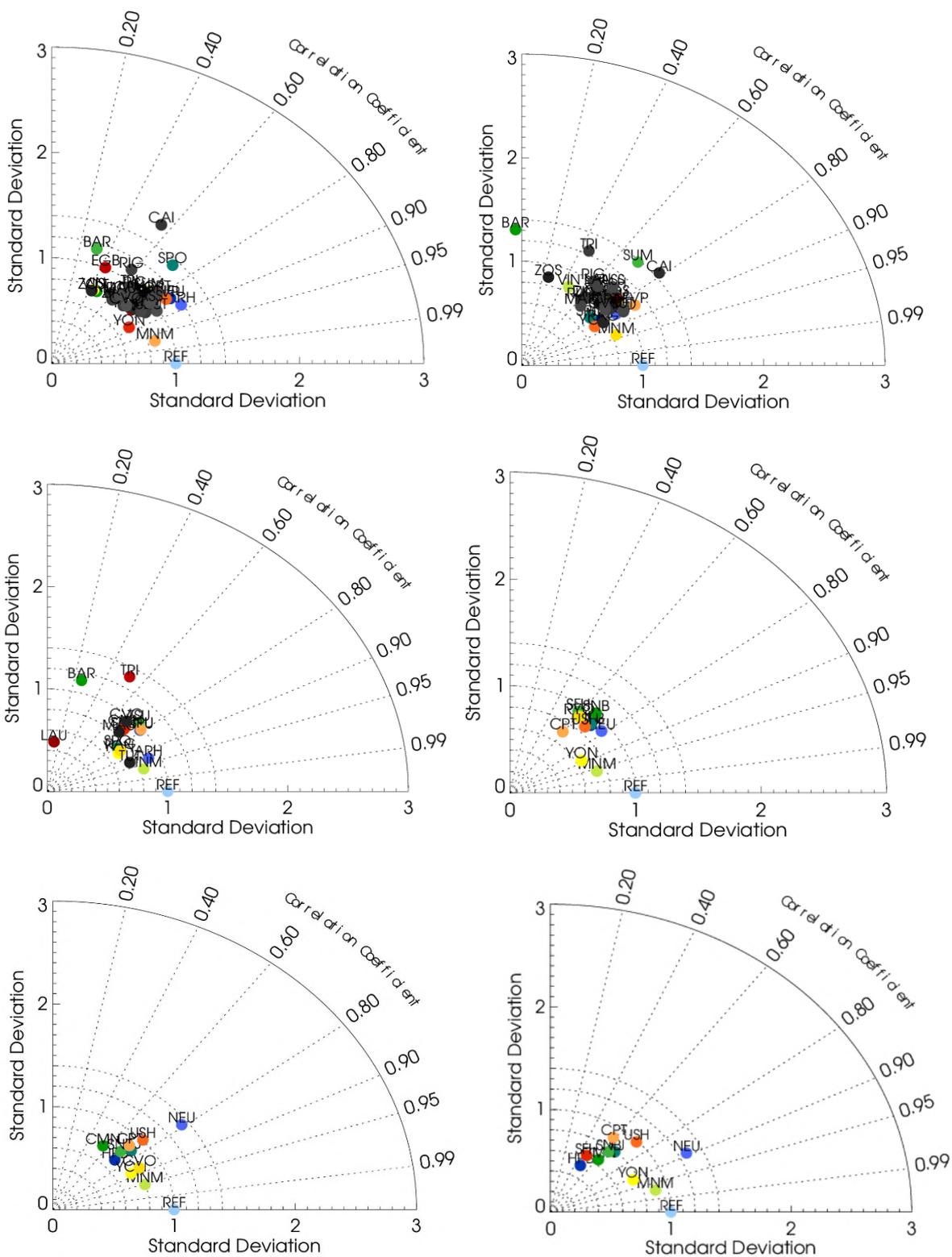


Figure 4.2.3: Taylor diagram with standard deviation and correlation for O<sub>3</sub> in 2003 (first row left), 2005 (first row right), 2007 (second row left), 2009 (second row left), 2011 (third row left), 2013 (third row right), and 2015 (fourth row left), 2017 (fourth row right), and 2019 (last row left) and 2020 (last row right).

Table 4.2.1: GAW station data (O<sub>3</sub>) used in the validation for 2003-2020.

Station	Name	Latitude	Longitude	Altitude	Year
Alert	ALT	82.45	-62.52	210	2003,2005-2012
Arrival Heights	ARH	-77.8	166.67	184	2003-2008, 2011- 2015
Assekrem	ASS	23.27	5.63	2710	2003-2013
Baring Head	BAH	-41.41	174.87	85	2003-2014
Barrow	BAR	71.32	-156.6	11	2003-2015
BEO Moussala	BEO	42.18	23.586	2925	2006-2009, 2011, 2012
Cairo	CAI	30.08	31.28	35	2010-2014
Cape Grim	CAG	-40.68	144.68	94	2003
Monte Cimone	CMN	44.18	10.7	2165	2003-2012, 2014,2015, 2020
Cape Point	CPT	-34.35	18.48	230	2003-2020
Cape Verde	CVO	16.8	-24.9	10	2006-2020
Dobele	DOB	56.37	23.19	42	2009-2013
Egbert	EGB	44.23	-79.78	253	2003-2012
Everest Pyramid	EVP	27.96	86.82	5079	2009-2013
Giordan Lighthouse	GIO	36.07	14.22	167	2003
Hohenpeissenberg	HPB	47.8	11.02	985	2003-2010, 2016-2020
Iskrba	ISK	45.57	14.87	520	2003
Izana	IZO	28.3	-16.5	2367	2003
Jungfraujoch	JFJ	46.55	7.98	3578	2003,2016-2020
Kollumerwaard	KOW	53.33	6.28	0	2003-2014
Kosetice	KOS	49.58	15.08	534	2003-2006,2008-2014
K-Puszt	KPU	46.97	19.55	125	2003-2015
Krvavec	KRV	46.3	14.53	1720	2003
Lauder	LAU	-45.03	169.67	370	2003-2015
Mauna Loa	MAU	19.539	-155.578	3397	2003-2015
Minamitorishima	MNM	24.28	153.98	8	2003-2020
Neumayer	NEU	-70.65	-8.25	42	2003-2020
Neuglobsow	NGW	53.17	13.03	65	2003-2013
Payerne	PAY	46.82	6.95	490	2003-2014
Pic du Midi	PIC	42.94	0.14	2877	2008-2012
Ragged Point	RAG	13.17	-59.43	45	2006-2015
Rigi	RIG	46.07	8.45	1031	2003-2014
Rucava	RUC	56.17	21.18	18	2003-2014
Ryori	RYO	39.03	141.82	260	2003,2018,2020
Schauinsland	SCH	47.92	7.92	1205	2003-2013
Schneefernerhaus	SFH	47.42	10.98	2656	2004-2006,2008-2014,2016-2020
Sonnblick	SNB	47.05	12.95	3106	2003-2020
South Pole	SPO	-89.98	-24.8	2810	2003-2016
Summit	SUM	72.58	-38.48	3238	2003-2015
Syowa Station	SYO	-69	39.58	2	2003
Trinidad Head	TRI	41.05	-124.15	120	2003-2009,2011-2015
Tsukuba	TSU	36.05	140.13	25	2003-2015
Tudor Hill	TUD	32.27	-64.87	30	2003-2014
Tutuila	TUT	-14.24	-170.57	42	2003-2015
Vindeln	VIN	64.25	19.77	271	2003-2014
Waldhof	WAL	52.8	10.77	74	2003-2013
Westerland	WES	54.93	8.32	12	2003-2013
Yonagunijima	YON	24.47	123.02	30	2003-2020
Zavodnje	ZAV	46.43	15	770	2003
Zingst	ZIN	54.43	12.73	1	2003-2013
Zoseni	ZOS	57.08	25.54	183	2008-2014

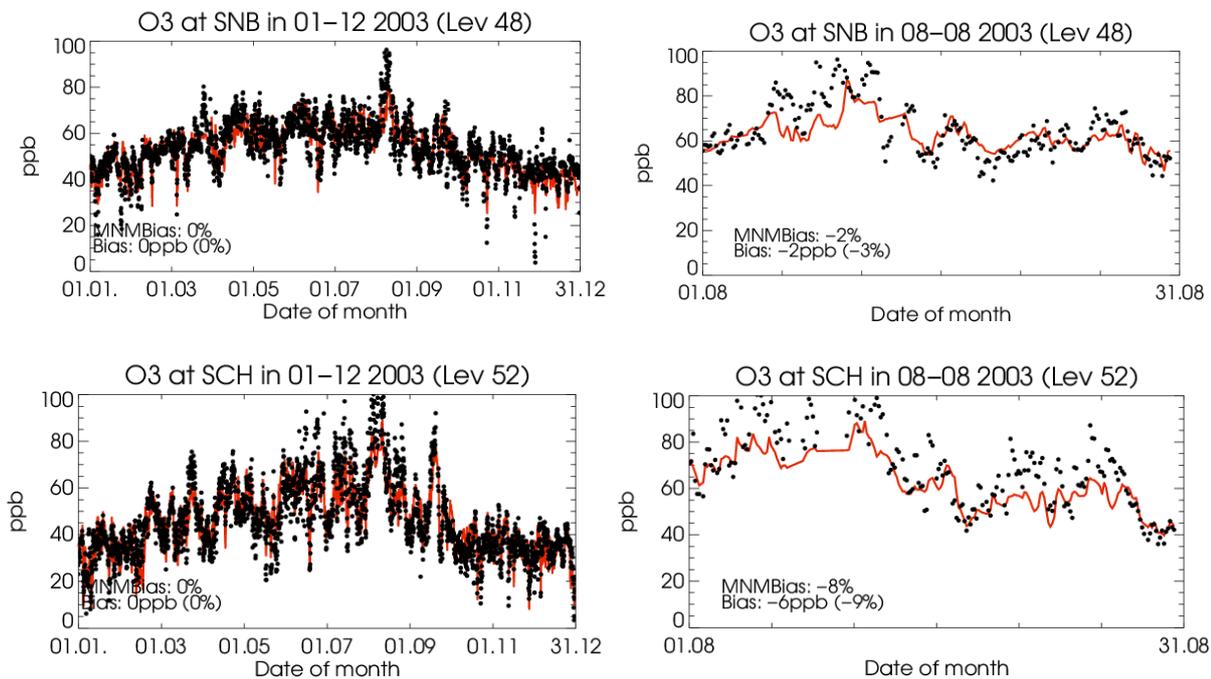


Figure 4.2.4: Time series for O<sub>3</sub> for Sonnblick Station (upper panel) and Schauinsland station (lower panel) during the whole year of 2003 (left) and for August 2003 (right).

### Heat wave in Europe 2003

During August 2003 there was an episode of exceptionally warm and dry weather with increased surface ozone in Europe. This peak in surface ozone can even be seen in the observations of high-altitude stations like Sonnblick (SNB), see Fig. 4.2.4. The model could correctly reproduce this episode of increased O<sub>3</sub> with only little underestimation of the maximum concentrations, see Fig. 4.2.4.

A comprehensive inter-annual seasonal evaluation of surface O<sub>3</sub> data can be found in Wagner et al. (2021).



## 4.3 Verification with European EMEP surface ozone observations

### 4.3.1 Seasonal and Multi annual comparison

The CAMS reanalysis, the CAMS control run and the MACC reanalysis were all compared to the surface EMEP and AirBase European background rural observations at the latitudinal zones: 30°N-40°N, 40°N-50°N and 50°N-70°N as well as using only the high altitude stations at elevations exceeding 1000 m above mean sea level. The results grouped together are shown in Fig. 4.3.1. Note that the MACC reanalysis shows the largest departures relative to the observations. However, both CAMS and MACC reanalysis biases show a high seasonal component in the time series studied. The amplitude of the seasonal variation is slightly higher for MACC as is evident in all comparisons shown in Fig. 4.3.1(a, c, e, g). The panels on the right show that the biases between models and observations are increasing as we move to higher latitudes and to higher altitudes in central and southern Europe.

Over Northern Europe, the CAMS reanalysis underestimates O<sub>3</sub> levels during the winter and spring seasons (MNMBs down to -20% during February), overestimates O<sub>3</sub> during autumn (MNMBs up to 15%) and reproduces well mean concentrations of ozone during the summer season. It should be noted that the CAMS reanalysis performs significantly better compared to the MACC reanalysis in terms of biases during June to September. Also, over Northern Europe the data assimilation reduces significantly the negative biases during January to April.

Over Central Europe CAMS reanalysis, the CAMS control and the MACC reanalysis show a large negative bias during January to April (down to -15% during February and March for the CAMS reanalysis and -25% for the MACC reanalysis). The data assimilation reduces significantly the negative biases during January to April (CAMS control biases during this period are almost equal with MACC reanalysis biases). During summer months the CAMS reanalysis reproduces well the mean ozone concentrations while the MACC reanalysis overestimates it up to 15%. During autumn the two reanalysis experiments show a positive offset which is almost equal to the MACC reanalysis bias (CAMS reanalysis MNMBs≈10%, MACC reanalysis MNMBs≈15%).

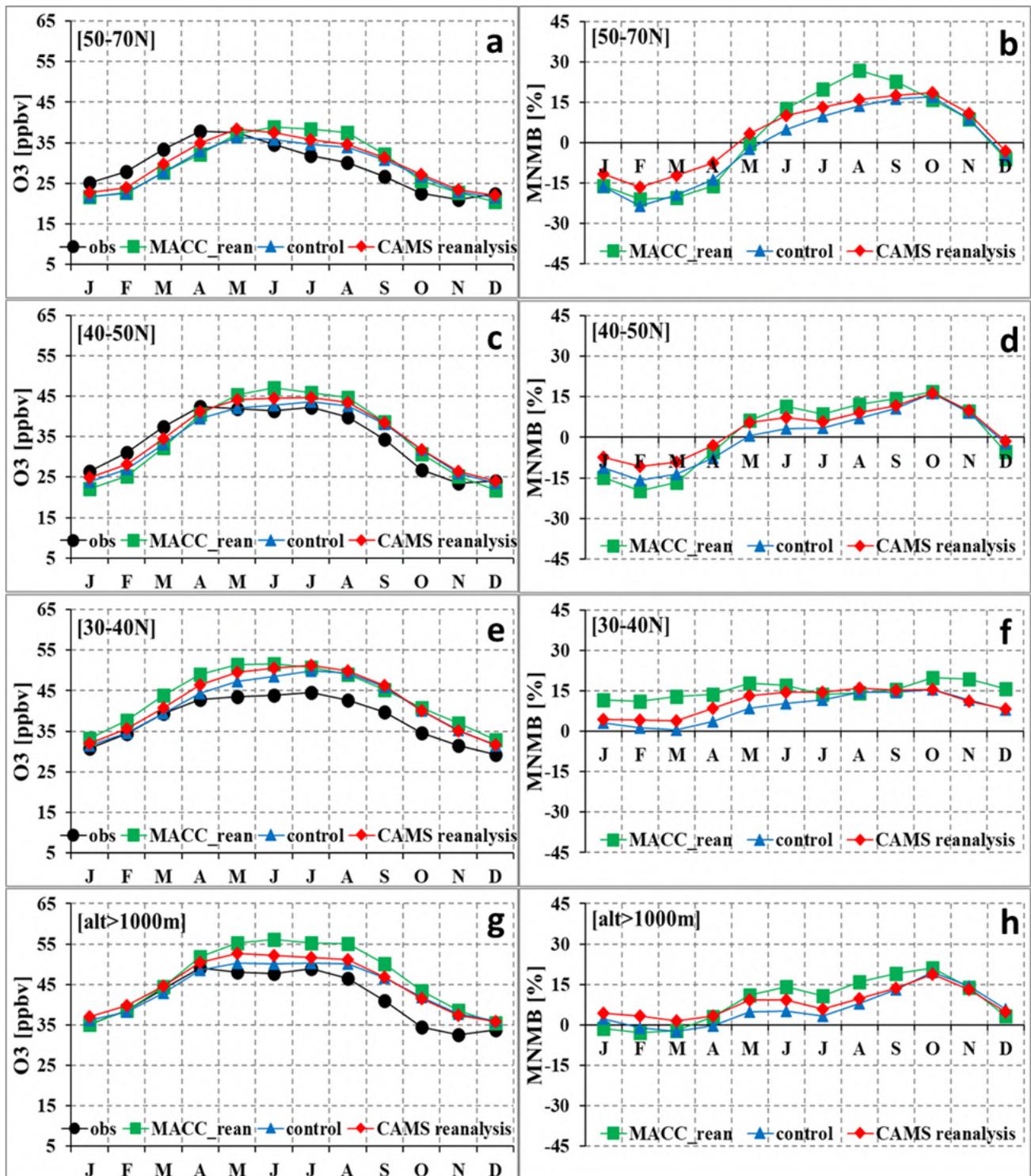


Figure 4.3.1: Mean monthly ozone variability for the period 2003-2020 (left) and the MNMBs (right) of the CAMS reanalysis (red robs), the associated CAMS control run (blue triangles), the MACC reanalysis (green squares), and the EMEP observations (black circles) over Northern Europe (1st row, a and b), Central Europe (2nd row, c and d), Southern Europe (3rd row, e and f) as well as for stations with altitude greater than 1000m above sea level (4rd row, g and h).

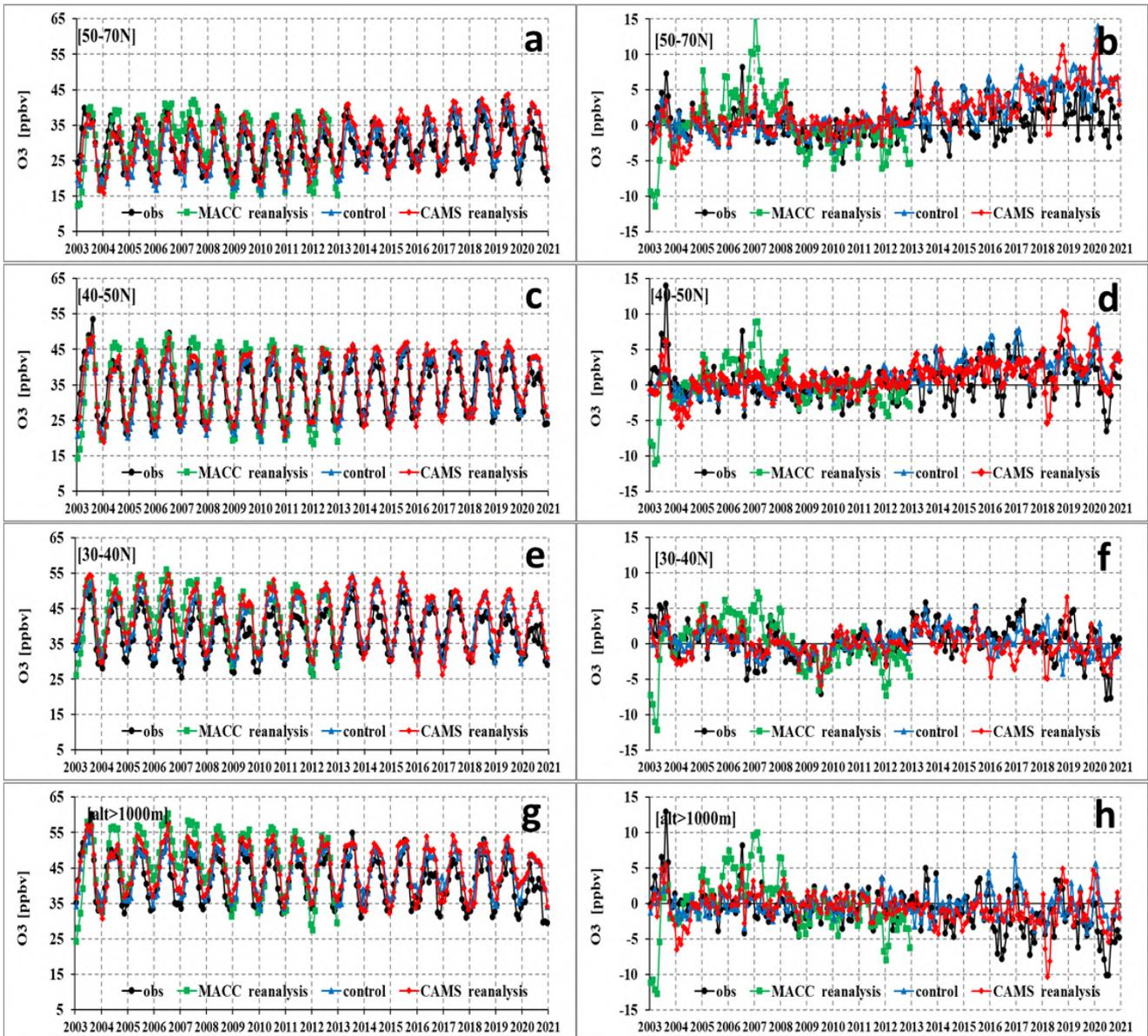


Figure 4.3.2: Monthly mean ozone concentrations during the period 2003-2020 before (left) and after removing the seasonal cycle (right), of CAMS reanalysis (red), the associated CAMS control run (blue), the MACC reanalysis (green) and the EMEP observations (black) averaged over European latitudinal zones of 50N-70N (1st row), 40N-50N (2nd row) and 30N-40N (3rd row) as well as for the group containing stations above 1000m (4th row).

Over Southern Europe all reanalysis experiments overestimate surface ozone values during all months and a higher positive offset is observed during autumn (up to 15%). It should be noted that the CAMS reanalysis is behaving better than the MACC reanalysis. Over Southern Europe the positive bias of the CAMS control run over the period January to June is almost half the bias observed for the CAMS reanalysis.



season	latitudinal belt	Correlation Coefficient			Correlation Coefficient (deseasonalised)		
		A	B	C	D	E	F
		CAMS_RE	CAMS_CONTROL	MACC_RE	CAMS_RE	CAMS_CONTROL	MACC_RE
Cold (Oct-Mar)	50°N-	0.57*	0.50*	0.35*	0.60*	0.67*	0.34*
	40°N-	0.70*	0.67*	0.44*	0.53*	0.71*	-0.22
	30°N-	0.66*	0.71*	0.52*	0.10	0.37*	-0.05
	alt>1000	0.64*	0.55*	0.32*	0.28*	0.42*	-0.01
Warm (Apr-Sep)	50°N-	0.62*	0.56*	0.15	0.52*	0.60*	0.18
	40°N-	0.72*	0.67*	0.48*	0.65*	0.67*	0.21
	30°N-	0.66*	0.56*	0.61*	0.61*	0.57*	0.45*
	alt>1000	0.78*	0.82*	0.47*	0.70*	0.80*	0.32*
Year	50°N-	0.77*	0.74*	0.63*	0.55*	0.63*	0.23*
	40°N-	0.90*	0.89*	0.83*	0.56*	0.65*	0.03
	30°N-	0.89*	0.89*	0.84*	0.37*	0.49*	0.17
	alt>1000	0.88*	0.87*	0.77*	0.50*	0.61*	0.13

Table 4.3.1: Correlation Coefficients (from monthly zonal mean values) between CAMS reanalysis (CAMs\_REA), the associated CAMS control run (CAMs\_CONTROL), the MACC reanalysis (MACC\_REA) and EMEP/AIRBASE observations for the period 2003-2020 over European latitudinal zones of 50N-70N, 40N-50N and 30N-40N as well as for the group containing stations above 1000m. Statistically significant correlations (at 95% confidence level) are denoted with asterisks.

The same analysis was repeated for the high-altitude stations (stations with altitude higher than 1000 m a.s.l.) and it was found that the CAMS reanalysis overestimates slightly ozone mean concentrations during December to July (MNMBs up to 5%) and overestimate it more during autumn (MNMBs up to 15%). Again, the CAMS reanalysis experiment performs better than the MACC reanalysis in terms of biases especially for the period May to November.

Figure 4.3.2 presents on the left the monthly time series and to the right shows the same series after removing the seasonal cycle. This was done to examine the residuals in both the reanalyses and the observations without the dominant seasonal cycle. The correlation coefficients between the reanalysis experiments and the observations during the period 2003-2020 (2003-2012 for the MACC reanalysis) are highly significant over all grouped zones, as seen in Table 3.2.3.1, columns A, B, C. In that table, the correlation coefficients have been repeated after removing the seasonal cycle (columns D, E, F). It is evident from Table 1 that the CAMS runs are representing the variability of ozone in all cases even when the seasonal variability has been removed. It does show also the much poorer performance of the MACC reanalysis in reproducing the temporal variability at all latitudes when the seasonal cycle is removed.

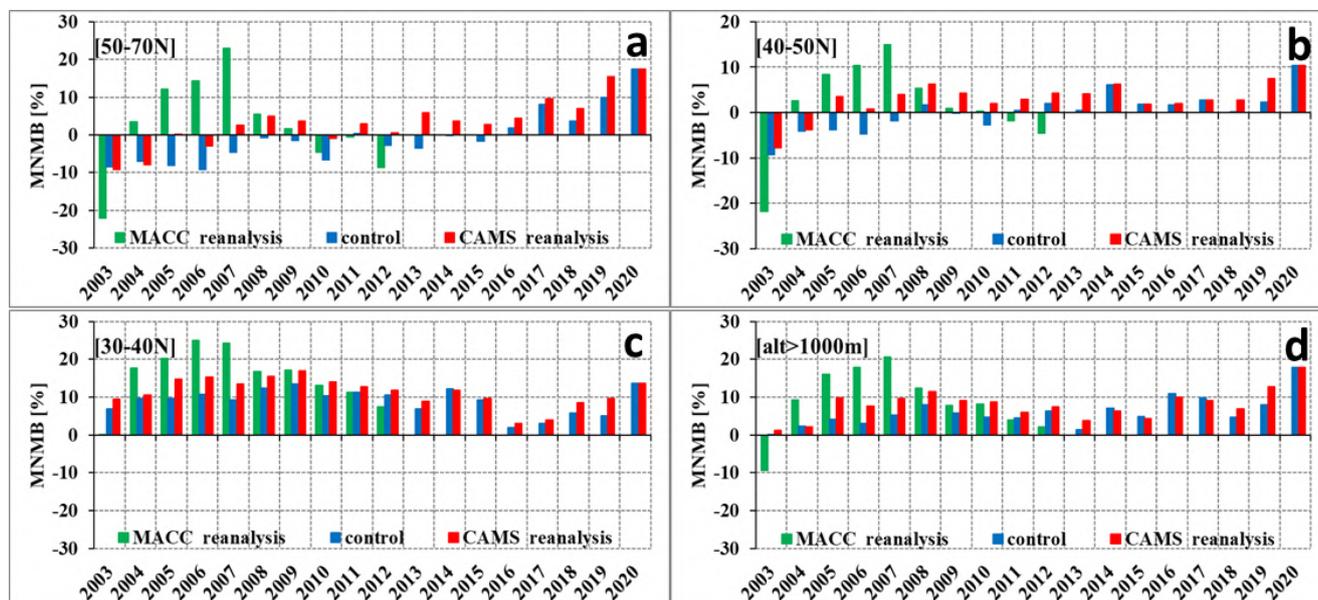


Figure 4.3.3: Annual MNMBs (in %) for the CAMS reanalysis (red), the associated CAMS control run (blue) and MACC reanalysis (green) for the years 2003-2020 of the model simulations and the EMEP observations (black) averaged over European latitudinal zones of 50N-70N (a), 40N-50N (b) and 30N-40N (c) as well as for the group containing stations above 1000m (d).

Additionally, the multiannual variability of the MNMBs for the three latitudinal zones has been examined for the period 2003-2020 (Fig. 4.3.3). It is noteworthy that in 2003 and especially during the 1st half of the year, large negative biases were recorded for the MACC reanalysis over all selected regions, possibly due to the non-stabilized conditions following the spin-up month (December 2002). These biases are significantly reduced in the new CAMS reanalysis. For the years 2004-2018 both reanalyses overestimate ozone mixing ratios over Southern Europe and high-altitude stations. The MNMBs in the new CAMS reanalysis experiments are reduced by about 50% compared to the MACC reanalysis for the years 2004-2007 and have almost the same magnitude for the years 2008-2012. Over Central and Northern Europe the new CAMS reanalysis and CAMS control with the exception of 2003 reproduce well the annual surface ozone levels (MNMBs between -5% and +5% almost equals for the individual years 2004-2020; it is notable that there is a positive tendency in the biases especially over Northern Europe after 2015). This is a significant improvement against MACC reanalysis during the period 2004-2007 that shows a strong positive bias over Central and Northern Europe which is increased up to 2007. For years after 2008 the performance of CAMS and MACC reanalysis are almost the same in terms of annual MNMBs. To summarize we conclude that although the bias is improving especially during years 2003-2007, there is still a seasonal dependence of the bias and MNMB clearly seen in Fig. 4.3.4. The amplitude of the seasonality of MNMB is significantly reduced in the latitudinal zones of 30°N-40°N as well as the high-altitude stations (stations with altitude higher than 1000 m a.s.l.).

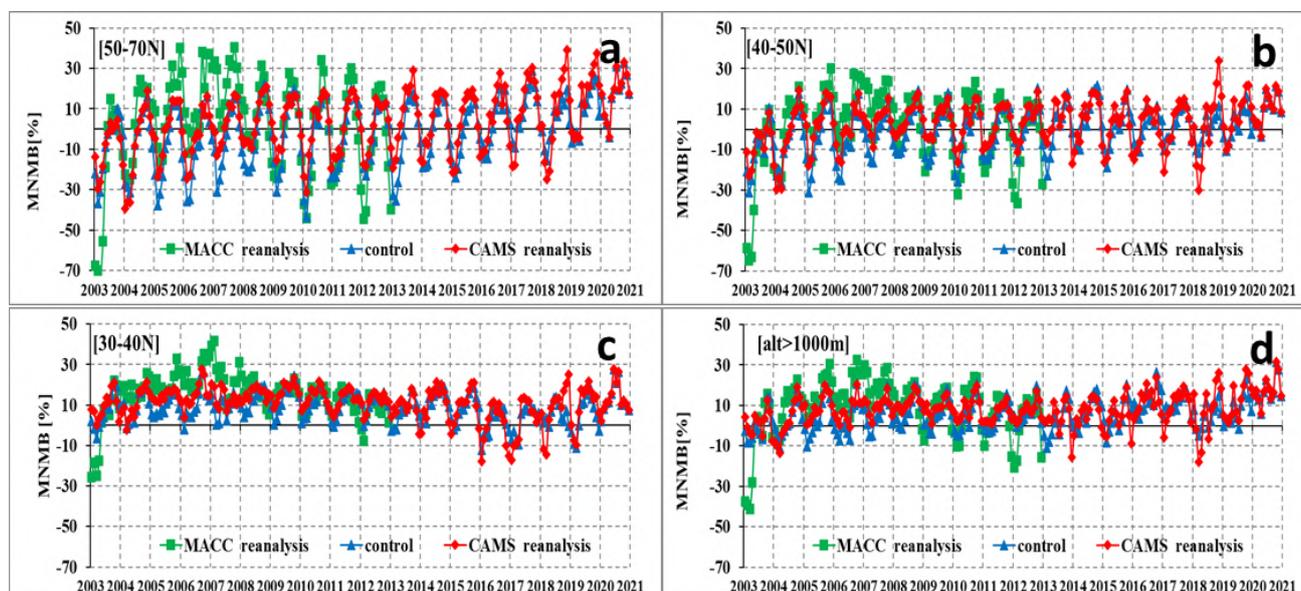


Figure 4.3.4: Monthly MNMBs (in %) for the CAMS reanalysis (red), the associated CAMS control run (blue) and MACC reanalysis (green) for the years 2003-2020 of the model simulations and the EMEP observations (black) averaged over European latitudinal zones of 50N-70N (a), 40N-50N (b) and 30N-40N (c) as well as for the group containing stations above 1000m (d).

Fig. 4.3.5 shows the spatial distribution of the Modified Normalised Mean Biases (MNMBS) calculated for the CAMS Reanalysis experiment and the MACC Reanalysis over the 262 background rural European EMEP and AirBase stations for the four seasons (winter, spring, summer and autumn). The seasonality in MNMBS is evident for the majority of the stations, especially over higher latitudes.

Fig. 4.3.6 shows the spatial distribution of the temporal correlations between modelled and observed surface ozone values calculated for the CAMS Reanalysis and the MACC Reanalysis over the 262 individual EMEP and AirBase stations for the four seasons. Correlations between reanalysis, ozone values and observations are highly significant for the majority of the stations. The higher correlations are observed during summer and autumn seasons ( $0.5 < r < 0.9$ ); relatively lower correlations are observed during the winter season particularly for the stations in Southern Europe ( $0.0 < r < 0.4$  depending on the station) and during spring at stations north of  $60^{\circ}\text{N}$  ( $0.0 < r < 0.5$  depending on the station). It should be noted that the CAMS Reanalysis experiments perform better than the MACC reanalysis in terms of correlation during all seasons.

#### 4.3.2 Comparison of model simulations during the extreme heatwave event in August 2003

The reanalysis performance was validated for the major heat wave over central and north-western Europe, during the 1st half of August 2003 (Fig. 4.3.7). This event has been extensively discussed in connection to an increased number of deaths (Koppe et al., 2004). Fig. 4.3.8 shows surface ozone times series (observed and modelled) during August 2003 at four stations located in France, Germany and Switzerland where very high ozone concentrations (up to 120 ppb) were observed during 1-14 August. From these figures it is evident that the reanalysis underestimates the peak values in surface ozone although there is a significant improvement over the MACC reanalysis in

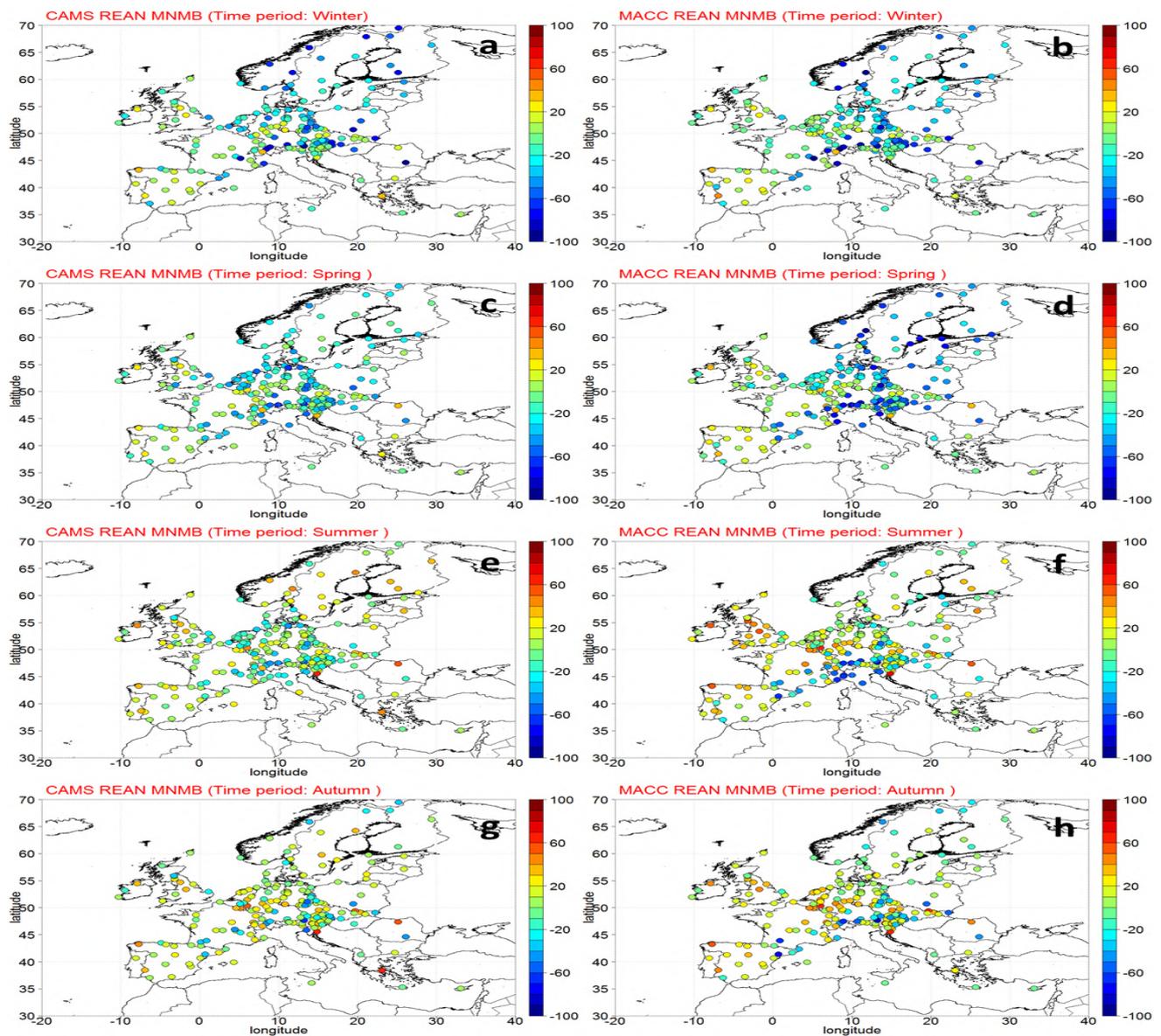


Figure 4.3.5. Modified Normalized Mean Biases (MNMBs) for EMEP data during Winter (1st row, a and b), Spring (2nd row, c and d), Summer (3rd row, e and f) and Autumn (4rd row, g and h) for the CAMS reanalysis 2003-2020 (left) and the MACC reanalysis 2003-2012 (right).

both the CAMS control and the CAMS reanalysis in terms of bias and correlations. The August 2003 event provided a unique extreme case to test the performance of the reanalysis. As an overall conclusion we observe that the CAMS control and reanalysis are a significant improvement compared to MACC, in particular over central and northern Europe. The differences between models and observations have slightly improved in the case of the extreme event in 2003, noting that this last analysis was made based on daily observations, in contrast to the monthly comparisons discussed above.

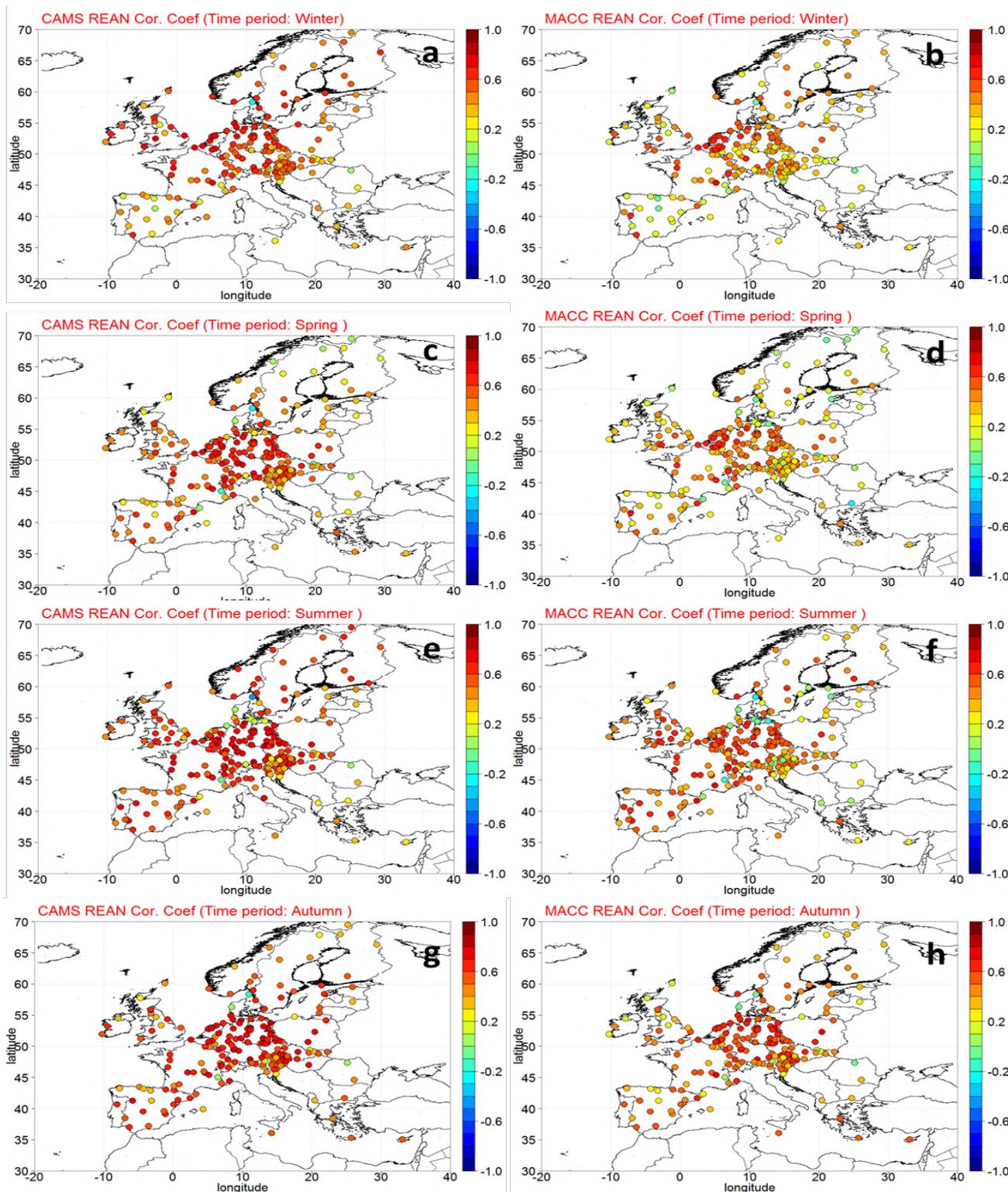


Figure 4.3.6. Correlation Coefficients (r) for EMEP data during Winter 1st row, a and b), Spring (2nd row, c and d), Summer (3rd row, e and f) and Autumn (4rd row, g and h) for the CAMS reanalysis 2003-2020 (left) and the MACC reanalysis 2003-2012 (right).

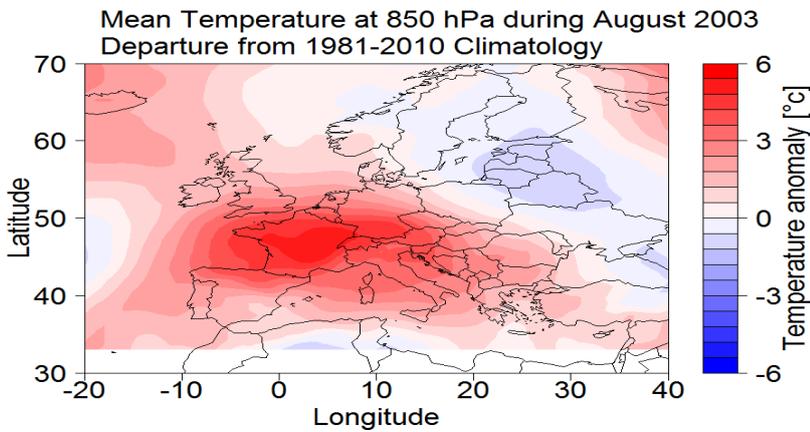


Figure 4.3.7. Spatial distribution of the air temperature anomaly at 850 hPa in August 2003, compared to the 1981-2010 climatology. Source of the data: ECMWF ERA-Interim reanalysis.

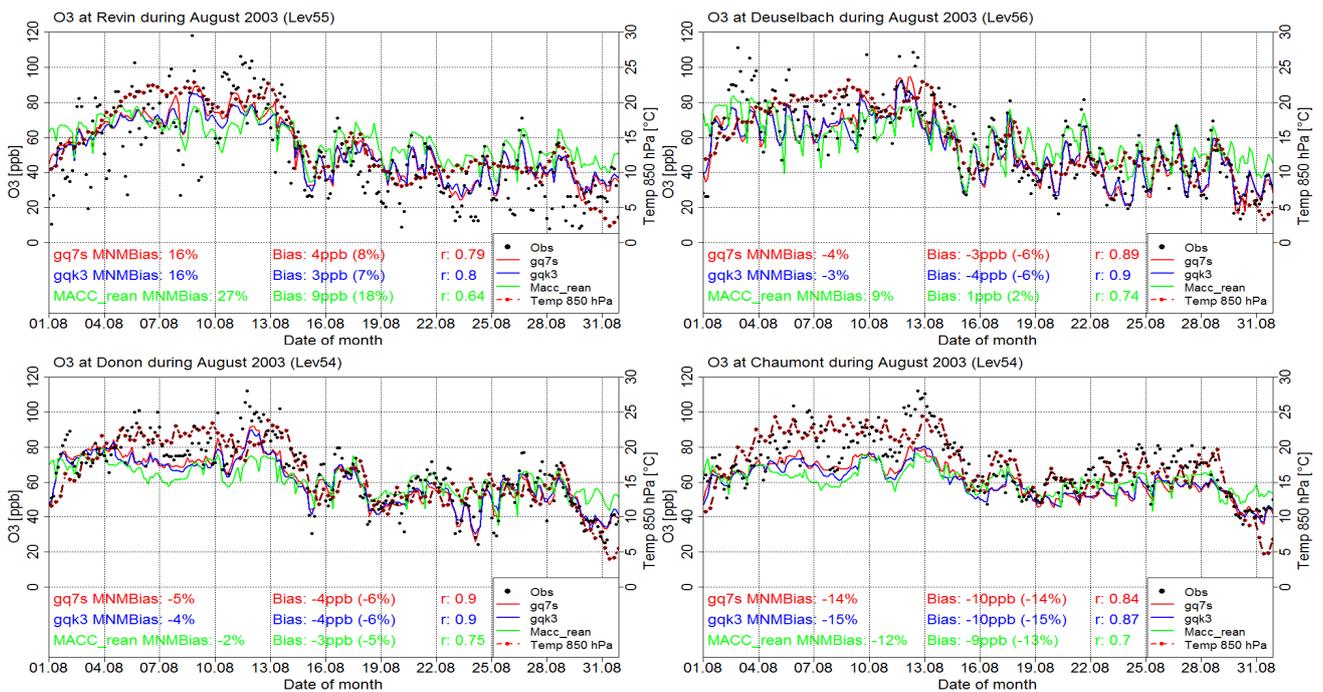


Figure 4.3.8. Time series of ozone over Revin, France (49.90°N, 4.63°E, top left), Donon, France (48.50°N, 7.13°E, top right), Deuselbach, Germany (49.76°N, 7.05°E, down left) and Chaumont, Switzerland (47.05°N, 6.98°E, down left) for the period of the intense heat wave over Europe 1 August - 31 August 2003. Observations are in black, the CAMS reanalysis is in red, the associated CAMS control run in blue and the MACC reanalysis in green. The dotted red line shows the temperature at 850 hPa.



#### 4.4 Verification with IAGOS ozone observations

The results for the evaluation of the CAMS and MACC reanalysis with IAGOS ozone observations are presented by regions namely: Europe, North America, India, Eastern Asia and Africa (Southern and Western). The two reanalysis datasets have been evaluated for the period January 2003 - December 2020 for CAMS and January 2003 - December 2012 for MACC. It should be noted that because of the availability delay of Level 2 data (about one year), the data used here are Level 2 except for the year 2020. Nevertheless, it might be noted that Level 2 data is not fully completed for year 2020. Using Level 2 data here for years 2017-2019 might explain some of the differences in the following results compared with those of the previous report where Level 1 was used (same for IAGOS CO, see CO section). Besides, it should be noted that the drastic reduction of air traffic related to the COVID-19 crisis has strongly impacted IAGOS operations, which explains the lack of IAGOS data for the year 2020 as shown hereafter. The monthly time series and monthly scores are presented at a number of airports providing long time series. The scores (MNMB and FGE) have been calculated using the ozone profiles on a daily basis and using the model profiles coincident with IAGOS (MOZAIC) observations. In the figures below only MNMB results are presented. Score results were also calculated for different layers of the atmosphere (SL: Surface Layer, BL: Boundary Layer, FT: Free troposphere, UT: Upper Troposphere, LS: Lower stratosphere) and different regions by grouping airports. The comparisons at cruise level of the flight are also presented at the end of each region subsection.

##### Europe

Between 2003 and 2020, there have been almost continuous IAGOS (MOZAIC) observations at Frankfurt. Time series of a few years are also available at other airports in Europe, with Vienna, Munich, London and Paris.

Fig. 4.4.1 shows the time series of the monthly profiles for ozone at Frankfurt during the period 2003 and 2020, for IAGOS observations, for the CAMS reanalysis and associated control run, as well as for the MACC reanalysis. The time series of the monthly scores (MNMB) associated to these models are presented in Fig. 4.4.2. The annual cycle is clearly depicted on Fig. 4.4.1 and is well reproduced by both models (see also Fig. 4.4.3).

Fig. 4.4.2 and 4.4.4 show that the CAMS reanalysis clearly represents an improvement as compared to the MACC reanalysis. Although the behaviour is similar for both models, the bias of the CAMS reanalysis is always smaller than that of MACC, especially in the lowest layers and in the free troposphere, while in the UTLS the results of the two models are more similar. In these figures, the CAMS reanalysis also clearly appears to be more stable than the MACC reanalysis.

The best performance of the CAMS reanalysis is found in the mid-troposphere as shown by the absolute values of MNMB (<15%) throughout the entire period. At the beginning of the period, MNMB results for both reanalyses present a seasonality feature from the surface to mid-troposphere (Fig. 4.4.2) which shows that ozone is underestimated during winter and spring and overestimated in summer and fall with the largest underestimations in the lowest layers. This seasonal pattern is stronger for the MACC than for the CAMS reanalysis with underestimations in winter that reach up to 40 % in the first year for MACC (Fig. 4.4.2). The seasonality of the bias is

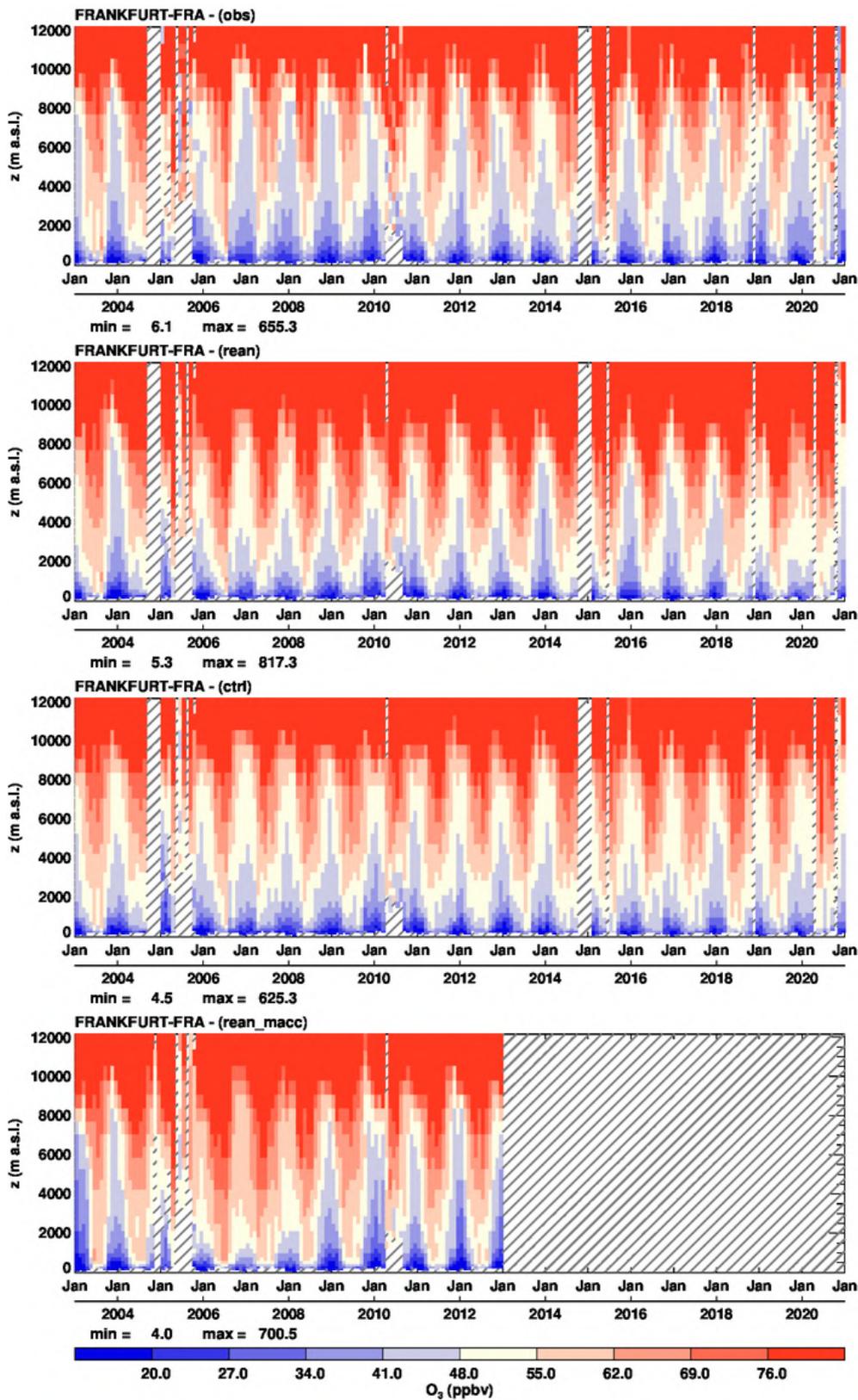


Figure 4.4.1: Time series of the monthly profiles for ozone at Frankfurt during the period 2003-2020. From top to bottom panel the time series correspond to: IAGOS (MOZAIC), CAMS reanalysis, control run and MACC reanalysis. Units: ppbv.

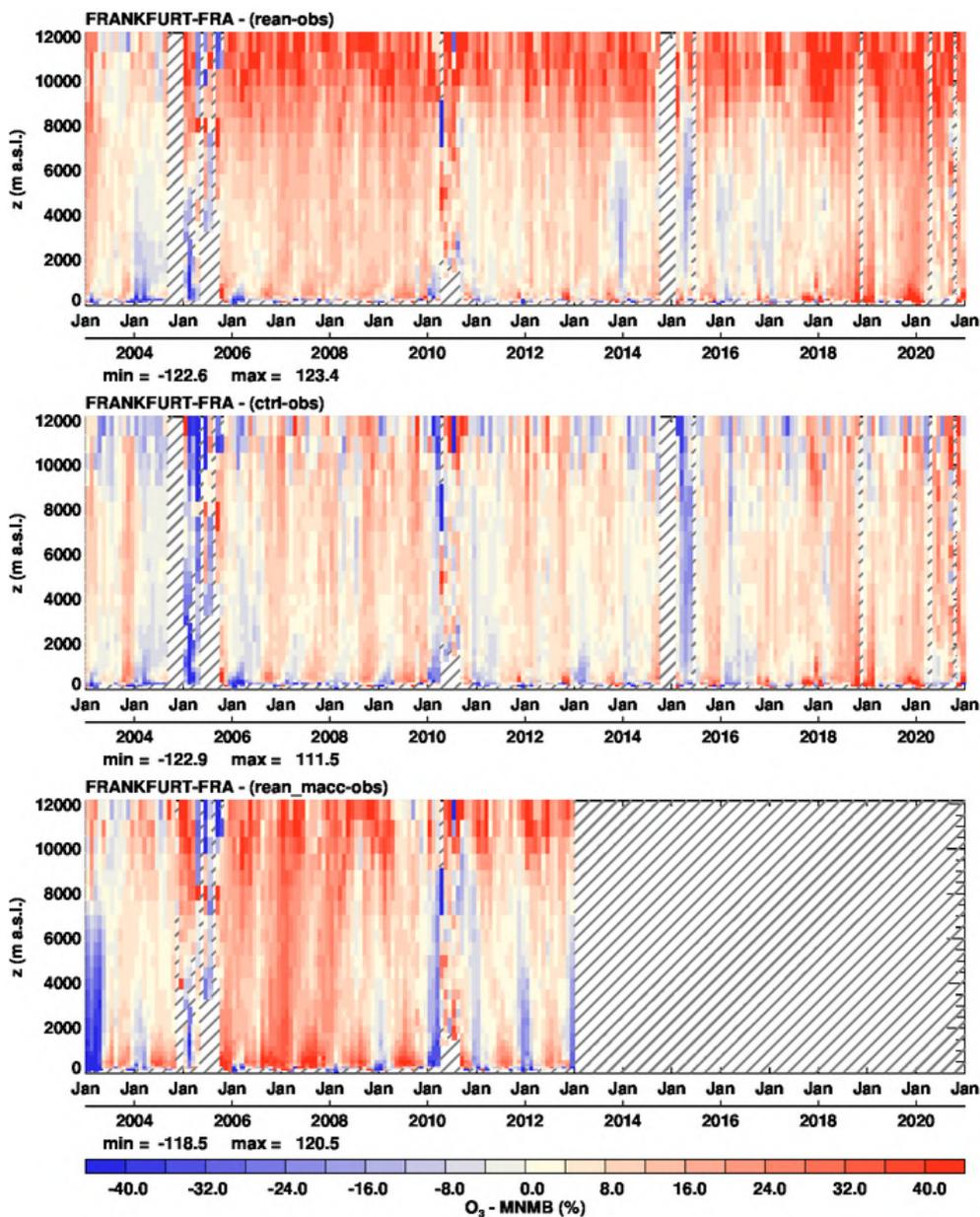


Figure 4.4.2: Time series of the monthly scores (MNMB) for ozone at Frankfurt during the period 2003-2020. From top to bottom panel, the scores are for: CAMS reanalysis, control run and MACC reanalysis. Units: %.

attenuated with the assimilation of new satellite data in late 2004 (MLS and OMI) for both reanalyses and ozone is mostly overestimated even in winter with rare underestimations close to the surface. While the bias is small for the CAMS reanalysis from the surface to mid-troposphere, the MACC reanalysis shows large overestimations during the period from 2006 to 2008 (note poor sampling in 2005) with summer overestimations mostly larger than 30% (Fig. 4.4.2). Negative biases appear again in the MACC reanalysis starting from 2009, while the behaviour of the CAMS reanalysis remains stable until 2013 when a seasonal pattern of the bias appears but much weaker than that observed in the MACC results between 2009 and 2012. Indeed, from 2014 on, mid-tropospheric ozone is often underestimated in winter (Fig. 4.4.2 and 4.2.4) with MNMB values sometimes

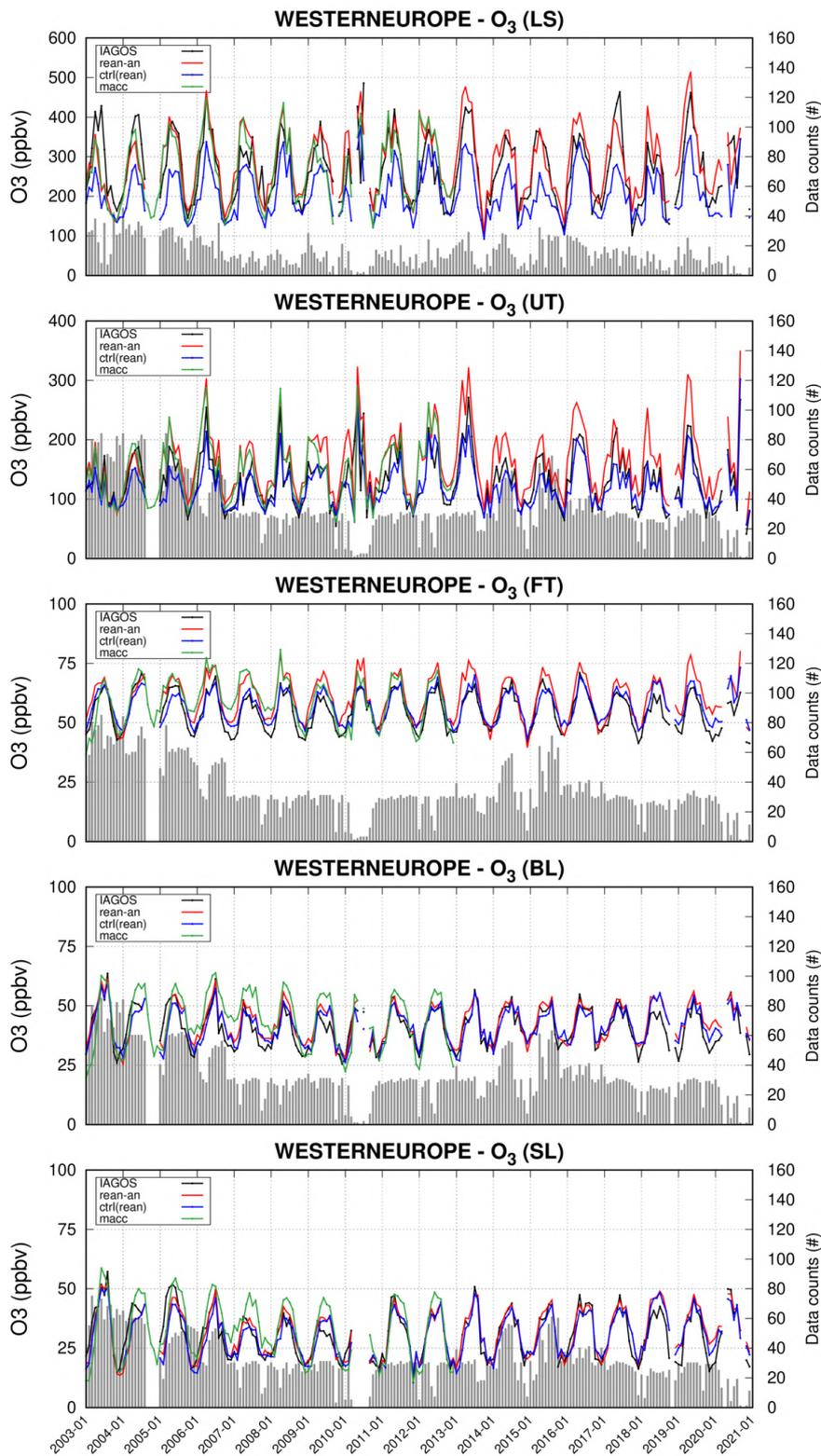


Figure 4.4.3: Monthly time series for ozone in different atmospheric layers over Western Europe during the period 2003-2020 (black: IAGOS, red: CAMS reanalysis, blue: Control run, green: MACC reanalysis). From top to bottom for SL: Surface Layer, BL: Boundary Layer, FT: Free troposphere, UT: Upper troposphere, LS: Lower Stratosphere. Units: ppbv.

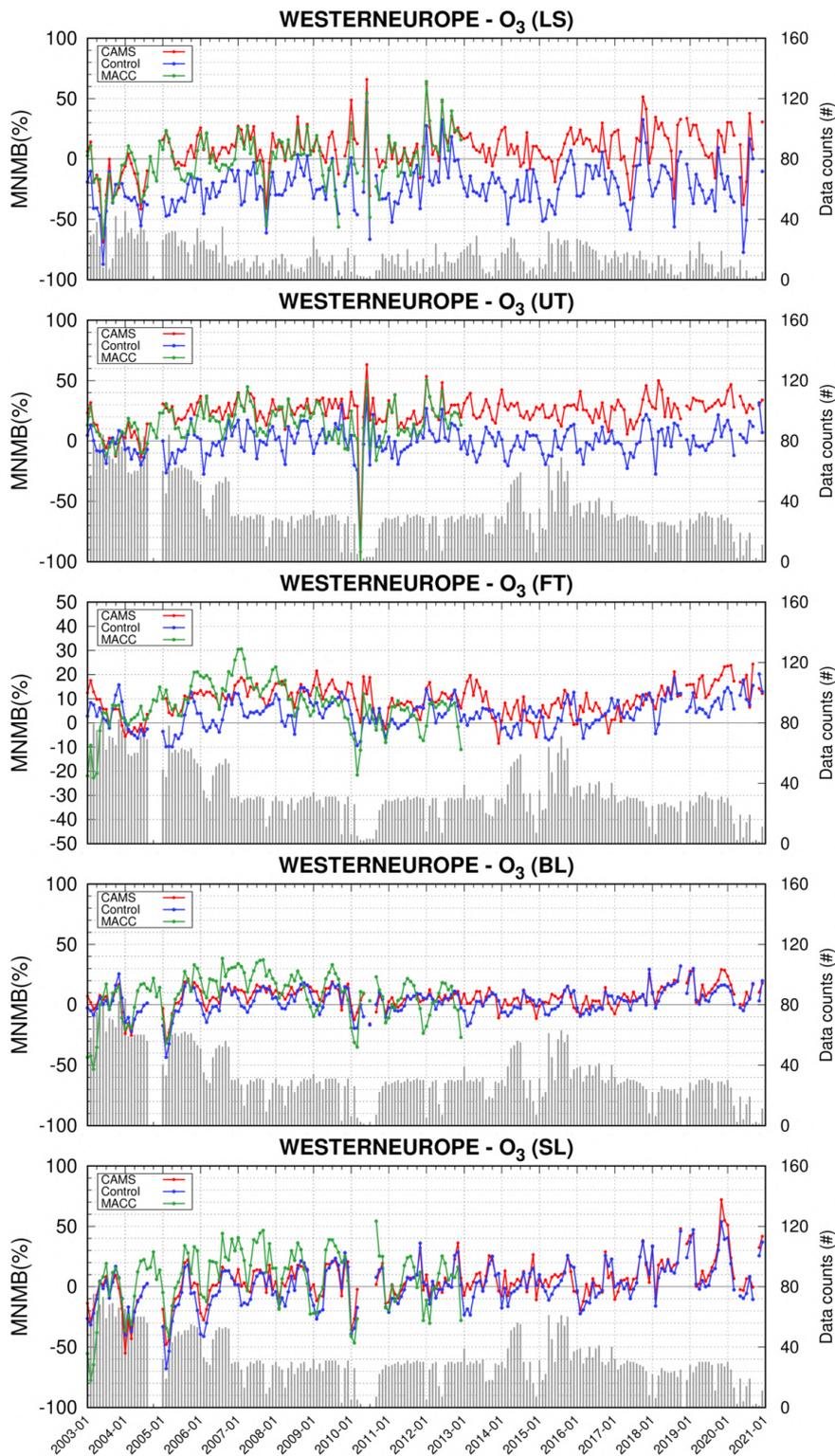


Figure 4.4.4: Time series of the monthly scores (MNMB) for ozone in different atmospheric layers over Western Europe during the period 2003-2020 (black: IAGOS, red: CAMS reanalysis, blue: Control run, green: MACC reanalysis). From top to bottom for SL: Surface Layer, BL: Boundary Layer, FT: Free troposphere, UT: Upper troposphere, LS: Lower Stratosphere. Units: %.

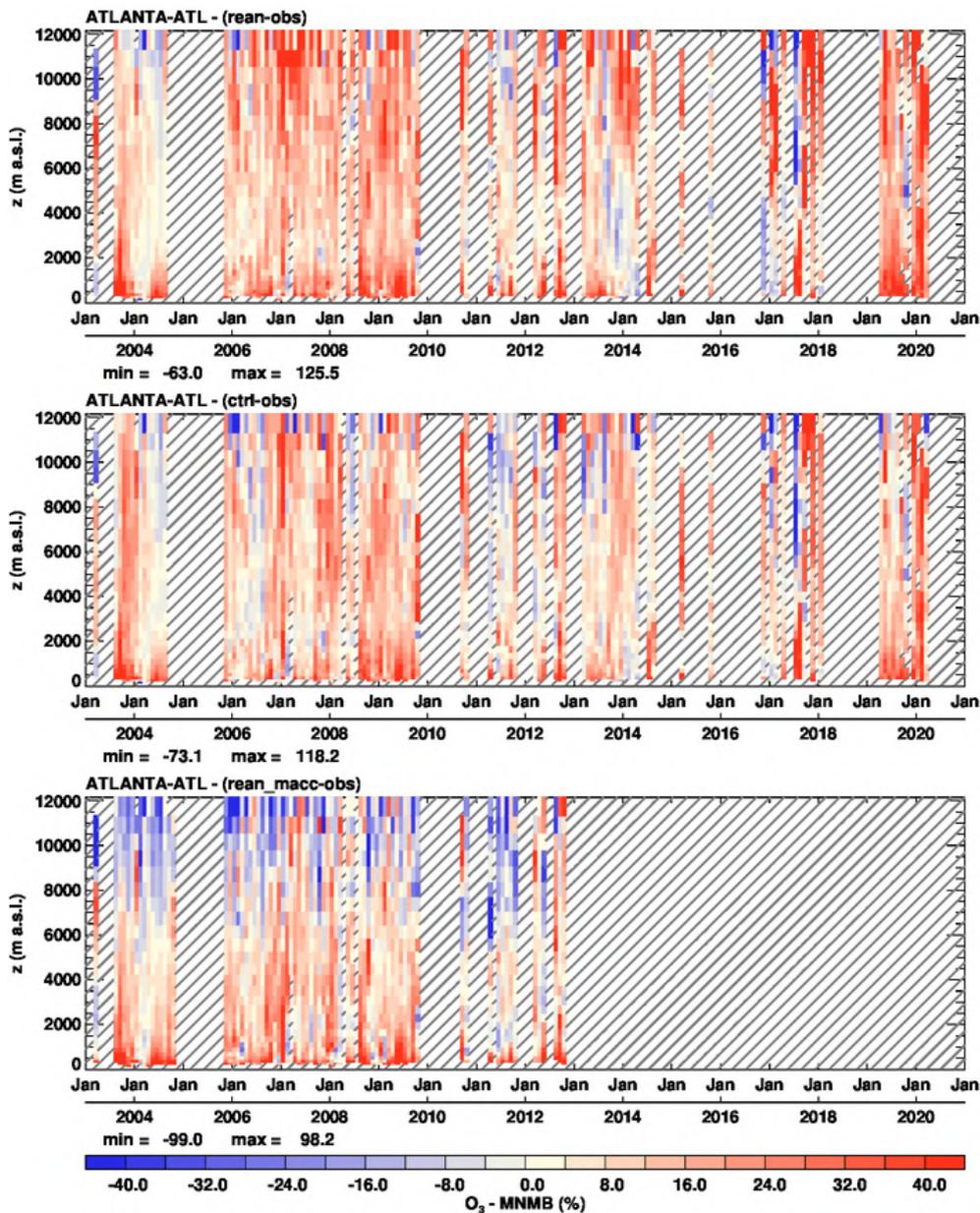


Figure 4.4.5: Time series of the monthly scores (MNMB) for ozone at Atlanta during the period 2003-2020. From top to bottom panel, the scores are for: CAMS reanalysis, control run and MACC reanalysis. Units: %.

reaching -20% at Frankfurt. This seasonality pattern might be linked to the change in the assimilated SBUV/2 data as also mentioned in comparisons with ozone sondes and GAW stations (see corresponding sections).

In the UTLS the results for the CAMS reanalysis show good agreement prior to fall 2004, with MNMB of about  $\pm 15\%$  (Fig. 4.4.2 and 4.4.4). At the end of 2004 the results of the CAMS reanalysis become worse with large overestimations, as shown by the large values of the scores (Fig. 4.4.2, MNMB > 40%). Like for CAMS, large overestimations start to appear in the UTLS in late 2004 for the MACC reanalysis. The occurrence of underestimations in MNMB from both reanalyses in 2005 and 2010 are likely related to poor data sampling. However, the CAMS reanalysis shows generally a



slightly better agreement than the MACC reanalysis in the UTLS which is clearly seen in FGE values (not shown here).

Regarding the control run results over Europe, the bias shows a systematic seasonal pattern in the lowest layers throughout the full period, with winter-time underestimations that are larger than for the CAMS reanalysis (Fig. 4.4.2 and 4.4.4). In the case of the control run, this seasonality pattern extends from the surface to the mid-troposphere. This behaviour of the control run in the low to mid-tropospheric layers underlines the benefit of assimilation in the CAMS reanalysis where the bias seasonality has mostly disappeared or is much weaker. In the UTLS while the CAMS reanalysis overestimates, no systematic behaviour is found for control run presenting both positive and negative MNMB (Fig. 4.4.2 and 4.4.4).

As expected, the increase of the bias observed in 2019 for both CAMS reanalysis and control run mentioned in the previous report (2003-2019) was an artefact that has disappeared here with the use of Level 2 data (i.e. post-calibrated data). This known problem has already occurred for years 2017 and 2018 (see former reports).

The results obtained at the other European airports are similar to those of Frankfurt. The seasonal feature in the lowest layers is more pronounced than at Frankfurt for some locations and extends to the mid-troposphere, as it is the case for example in Vienna. For these airports the agreement in the UTLS for both reanalyses seem to be better prior to 2004.

#### North America

Many North American airports have been frequently visited by IAGOS (MOZAIC) between 2003 and 2019 including: Atlanta, Washington, Philadelphia, New York, Montreal, Vancouver, Boston and Detroit. As shown in Fig. 4.4.5, 4.5.6 and 4.5.7, at North American airports, the results of the CAMS reanalysis and control run are very similar to those found over Europe (see previous comments). In the lowest layers the bias is found to be slightly larger over North America than over Europe with larger overestimations for CAMS and associated control run as shown for example at Atlanta (Fig. 4.4.5 and 4.4.6). Another difference as compared to Europe is the presence of large underestimations reaching 40% at the Western US airports of Los Angeles and Vancouver, where data are available mostly in winter. In the UTLS, although ozone is mostly overestimated like over Europe, underestimations are also observed for many airports. Large negative biases of about -30% are found for example at Toronto, New York and Philadelphia (not shown). The results of the MACC and CAMS reanalyses are very similar from the surface to mid-troposphere in North America whereas more differences were found over Europe (Fig. 4.4.5 and 4.4.7). In the UTLS, the MACC reanalysis often underestimates over North American airports (Fig. 4.4.7) with mostly underestimations at Atlanta (Fig. 4.4.5), which was not the case over Europe.

#### India

IAGOS (MOZAIC) has been providing nearly continuous time series at Hyderabad during the period 2006-2011 (Fig. 4.4.8). The seasonal pattern at Hyderabad contrasts with that of mid-latitude locations where surface ozone is usually high in summer and spring due to both enhanced photochemistry and stronger stratosphere-troposphere exchanges. The monsoon regime regulates the ozone in the lowest layers with high values in the winter monsoon and low values in the

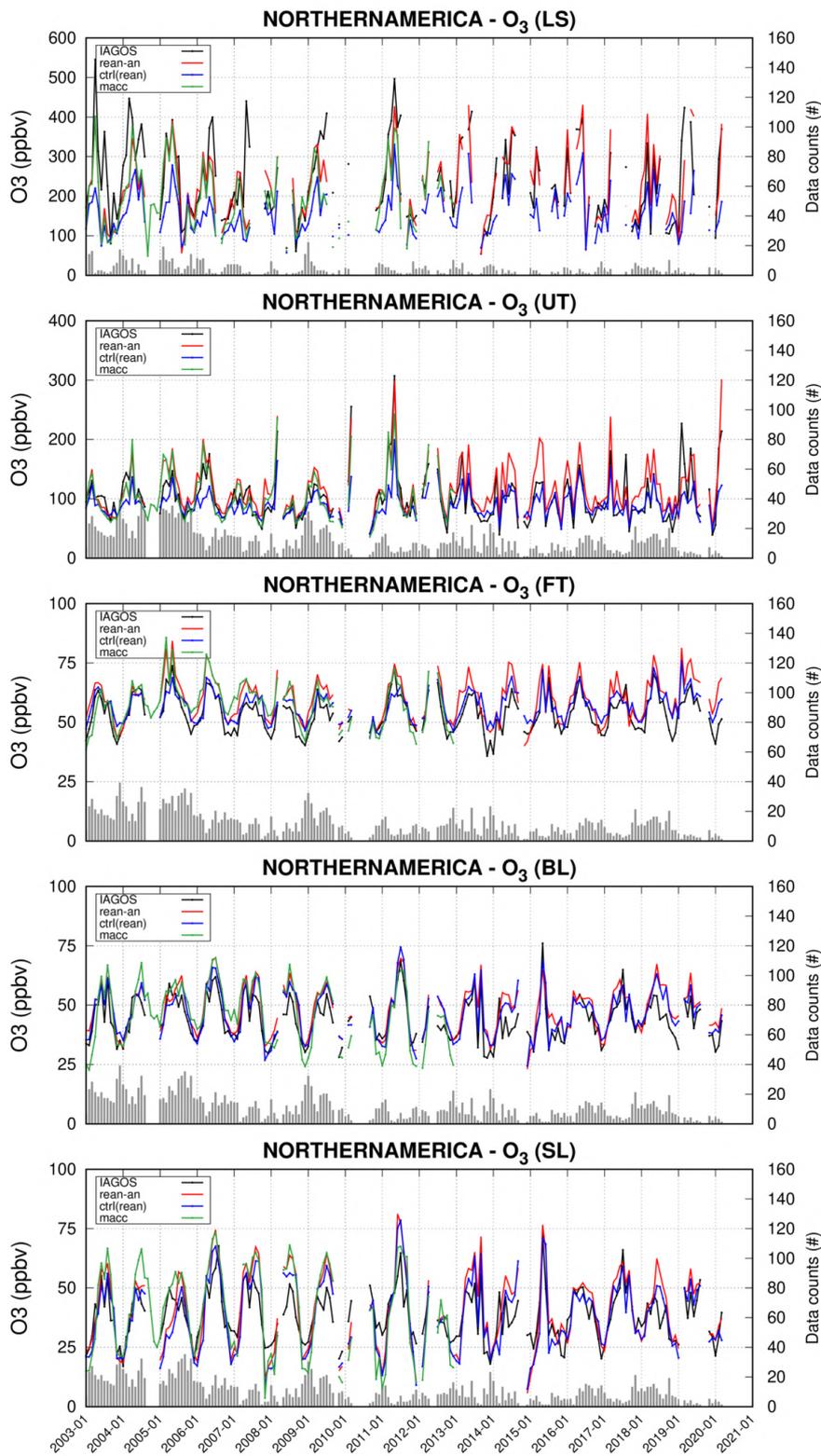


Figure 4.4.6: Monthly time series for ozone in different atmospheric layers over Northern America during the period 2003-2020 (black: IAGOS, red: CAMS reanalysis, blue: Control run green: MACC reanalysis). From top to bottom for SL: Surface Layer, BL: Boundary Layer, FT: Free troposphere, UT: Upper troposphere, LS: Lower Stratosphere. Units: ppbv.

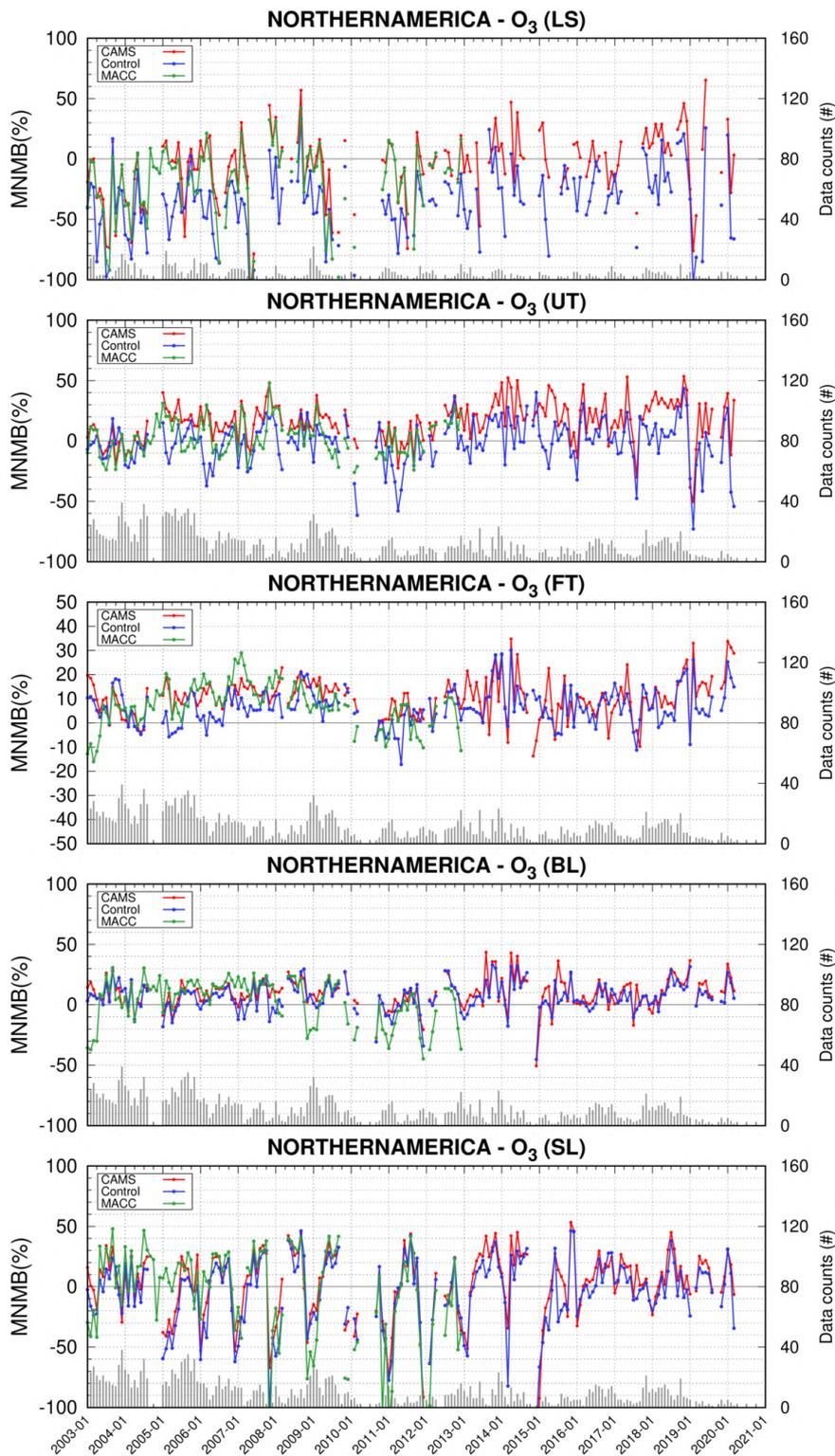


Figure 4.4.7: Time series of the monthly scores (MNMB) for ozone in different atmospheric layers over Northern America during the period 2003-2020 (black: IAGOS, red: CAMS reanalysis, blue: Control run, green: MACC reanalysis). From top to bottom for SL: Surface Layer, BL: Boundary Layer, FT: Free troposphere, UT: Upper troposphere, LS: Lower Stratosphere. Units: %.

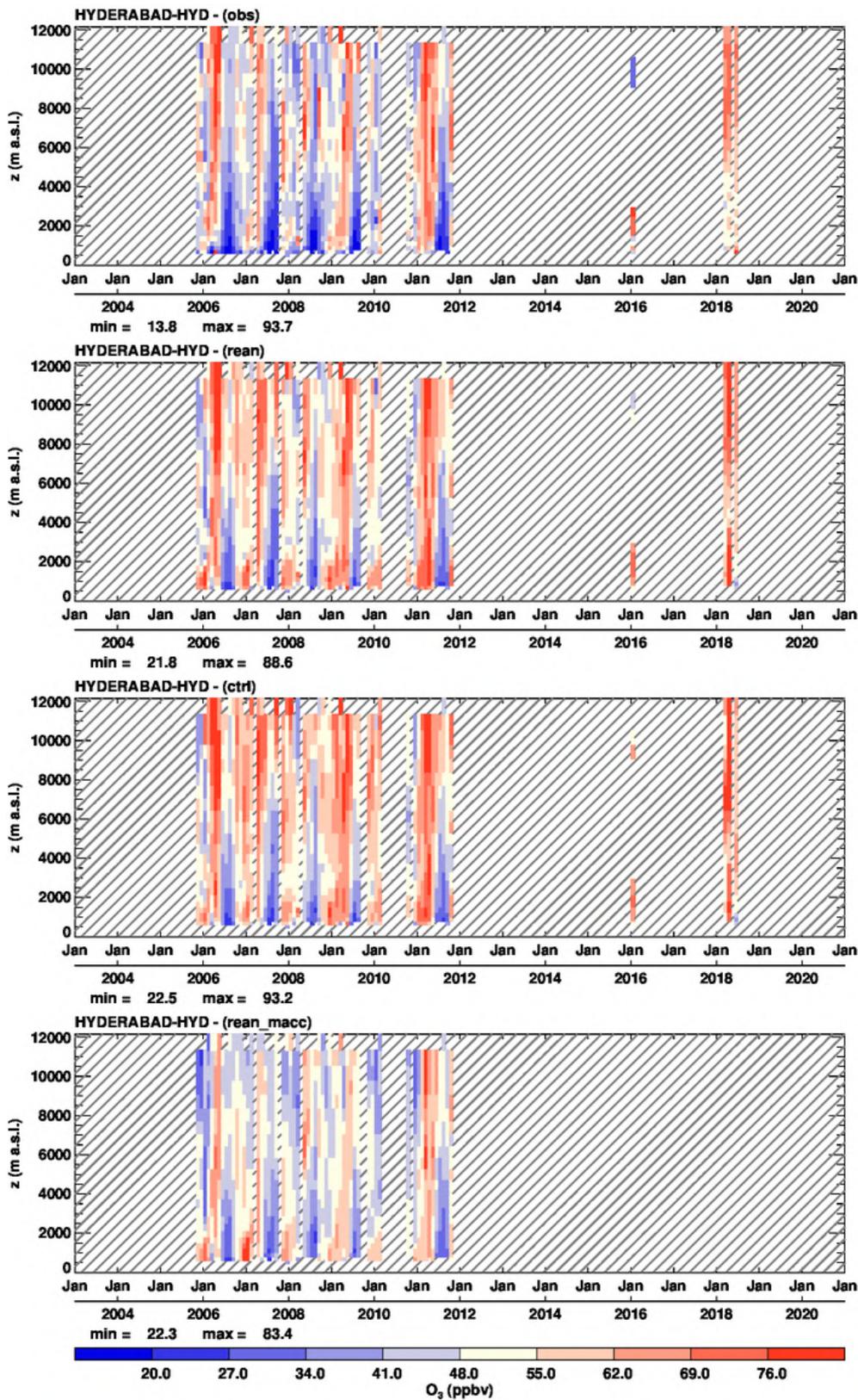


Figure 4.4.8: Time series of the monthly profiles for ozone at Hyderabad during the period 2003-2020. From top to bottom panel the time series correspond to: IAGOS (MOZAIC), CAMS reanalysis, control run and MACC reanalysis. Units: ppbv.

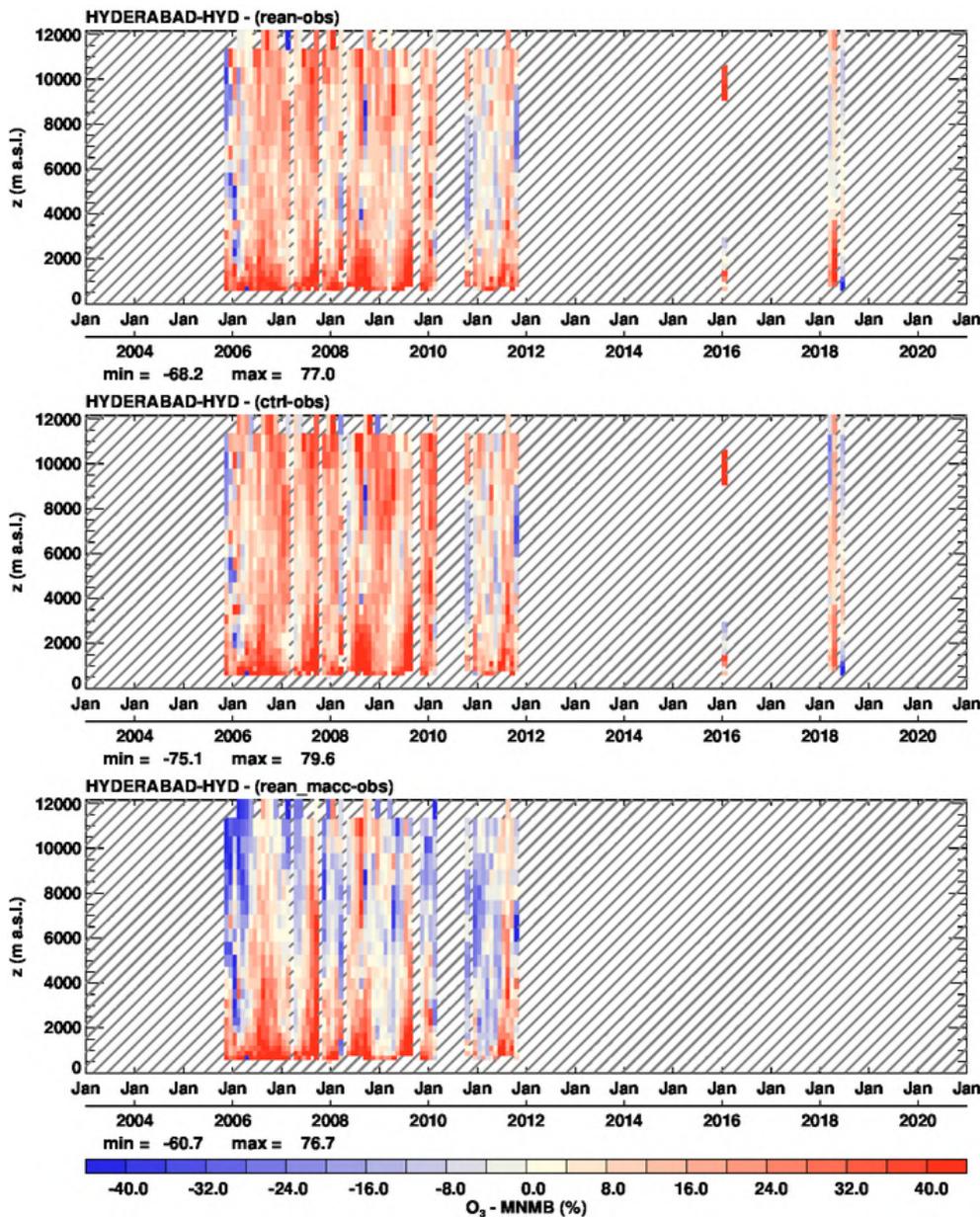


Figure 4.4.9: Time series of the monthly scores (MNMB) for ozone at Hyderabad during the period 2003-2020. From top to bottom panel, the scores are for: CAMS reanalysis, control run and MACC reanalysis. Units: %.

summer monsoon. The seasonality is well reproduced by both reanalyses, although the seasonal cycle appears to be stronger for the observations than for the models especially in the lowest layers (Fig. 4.4.8).

As shown by MNMB values in Fig. 4.4.9, the results of the two reanalyses are similar in the lowest layers and differ mostly in the mid-troposphere and UTLS. For the CAMS reanalysis, the ozone is mostly overestimated in the entire atmospheric column. The largest overestimations are found in the lowest layers during the summer (MNMB reaching more than 40%), and also sometimes in the UTLS. The results are slightly better in the mid-troposphere, but usually worse than at the European airports with positive MNMB often greater than 20%.

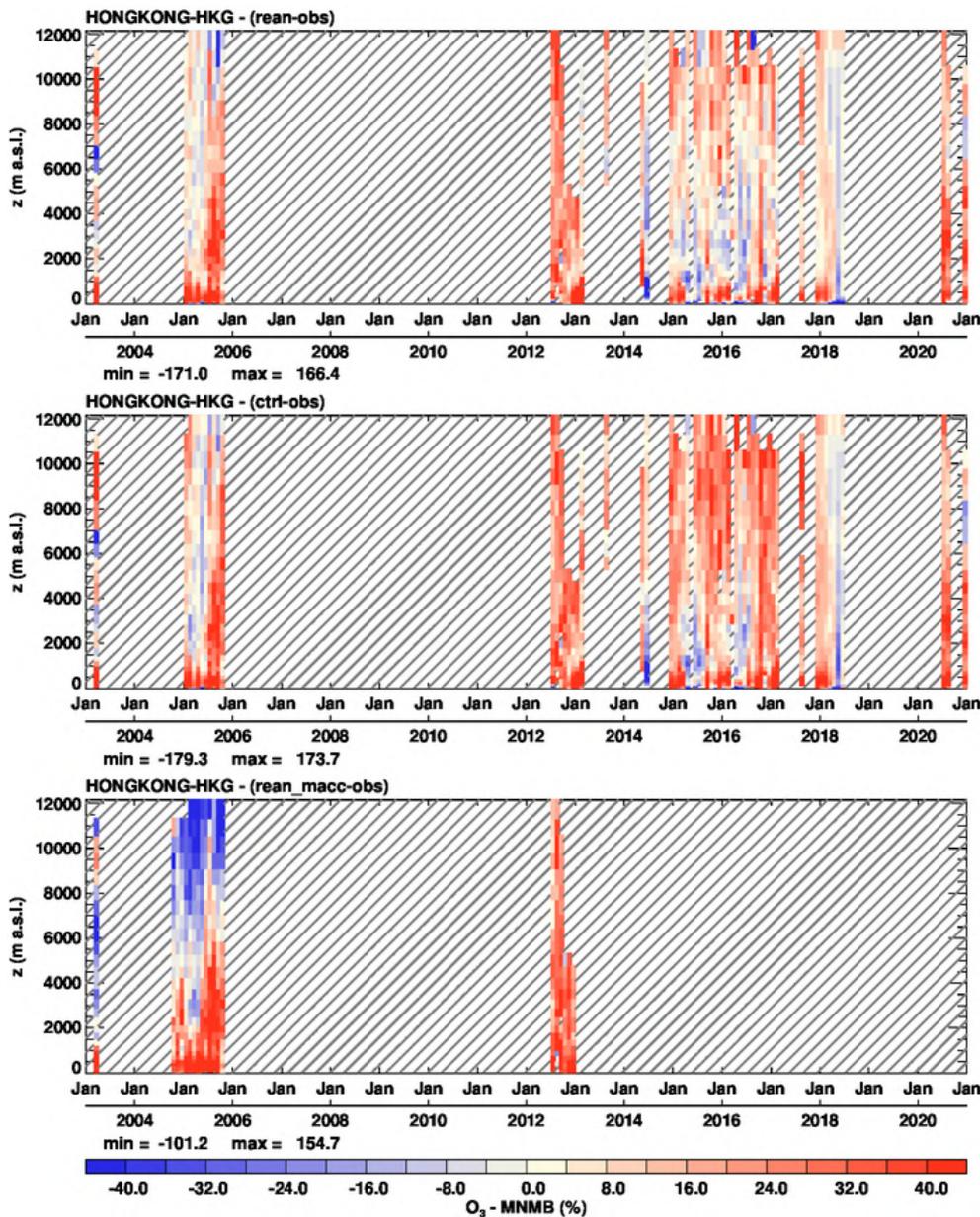


Figure 4.4.10: Time series of the monthly scores (MNMB) for ozone at Hong Kong during the period 2003-2020. From top to bottom panel, the scores are for: CAMS reanalysis, control run and MACC reanalysis. Units: %.

In the surface and boundary layer, MACC performs slightly worse than the CAMS (Fig. 4.4.9). Higher up in the atmosphere from the low troposphere to the UTLS, the MACC reanalysis presents a seasonal pattern, with underestimations in the winter monsoon and overestimations in the summer monsoon, a seasonality which is not present in the CAMS reanalysis (Fig. 4.4.9).

The CAMS reanalysis seems to generally provide slightly better performance than the control run which has slightly larger positive MNMB in all layers and throughout the period.

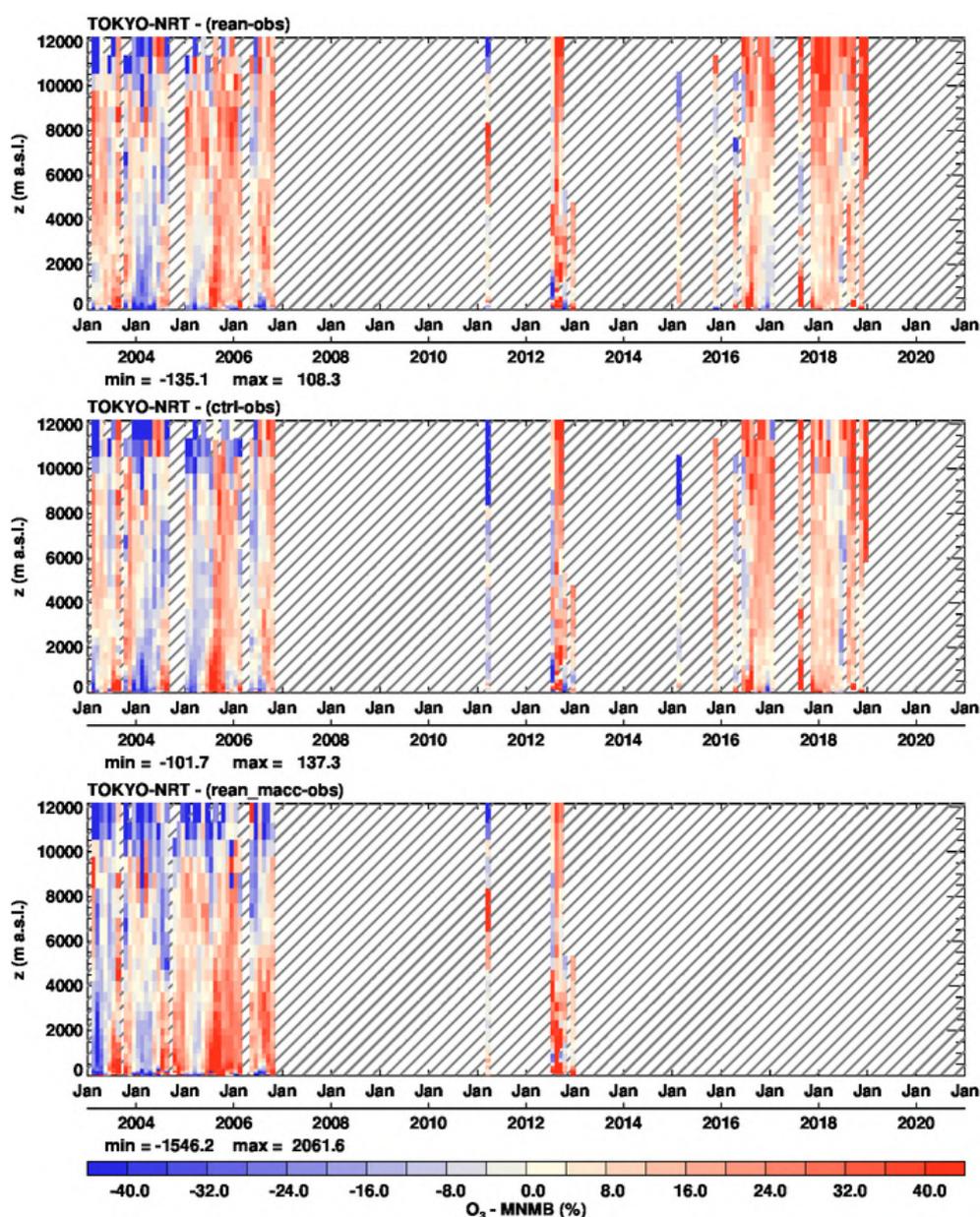


Figure 4.4.11: Time series of the monthly scores (MNMB) for ozone at Tokyo during the period 2003-2020. From top to bottom panel, the scores are for: CAMS reanalysis, control run and MACC reanalysis. Units: %.

### Eastern Asia

During the period 2003-2019, IAGOS (MOZAIC) provided time series of more than several years at many different locations across Eastern Asia: Taipei Hong Kong (Fig. 4.4.10), Ho Chi Minh City, Beijing, Nagoya Nanjing, Shenyang, Tokyo (Fig. 4.4.11), Osaka, and Bangkok. As shown on Fig. 4.4.12, on average the seasonality is well reproduced by all models over Eastern Asia.

In the lower troposphere, ozone is again mostly overestimated by the CAMS reanalysis, and for most airports this bias is larger than at the European airports (Fig. 4.4.10, 4.4.11). This is the case at Hong Kong with MNMB often exceeding 40% throughout the year (Fig. 4.4.10). At some airports like Tokyo (Fig. 4.4.11) the bias also presents a seasonal pattern, as found in other regions of the world,

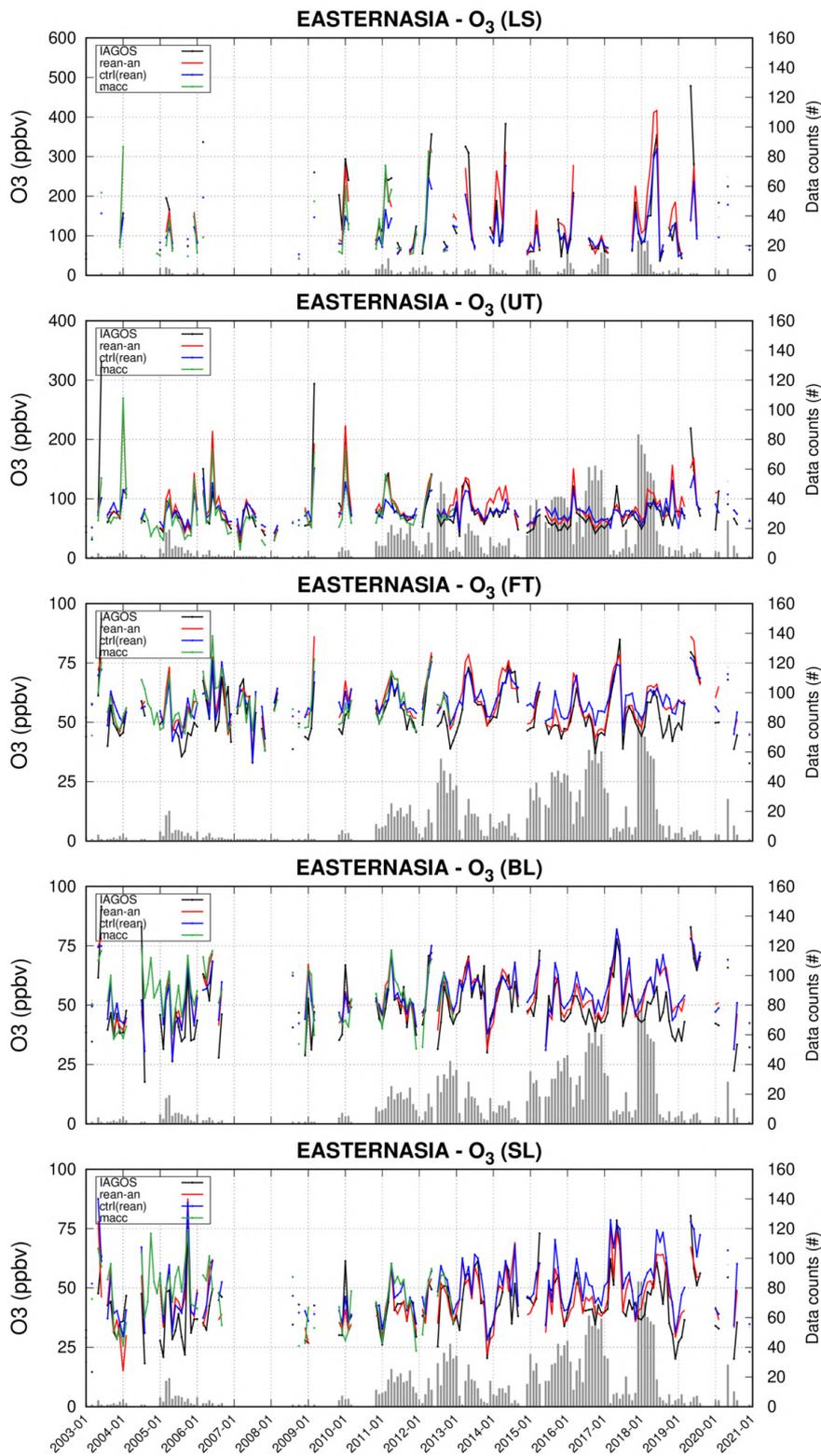


Figure 4.4.12: Monthly time series for ozone in different atmospheric layers over Eastern Asia during the period 2003-2020 (black: IAGOS, red: CAMS reanalysis, blue: Control run green: MACC reanalysis). From top to bottom for SL: Surface Layer, BL: Boundary Layer, FT: Free troposphere, UT: Upper troposphere, LS: Lower Stratosphere. Units: ppbv.

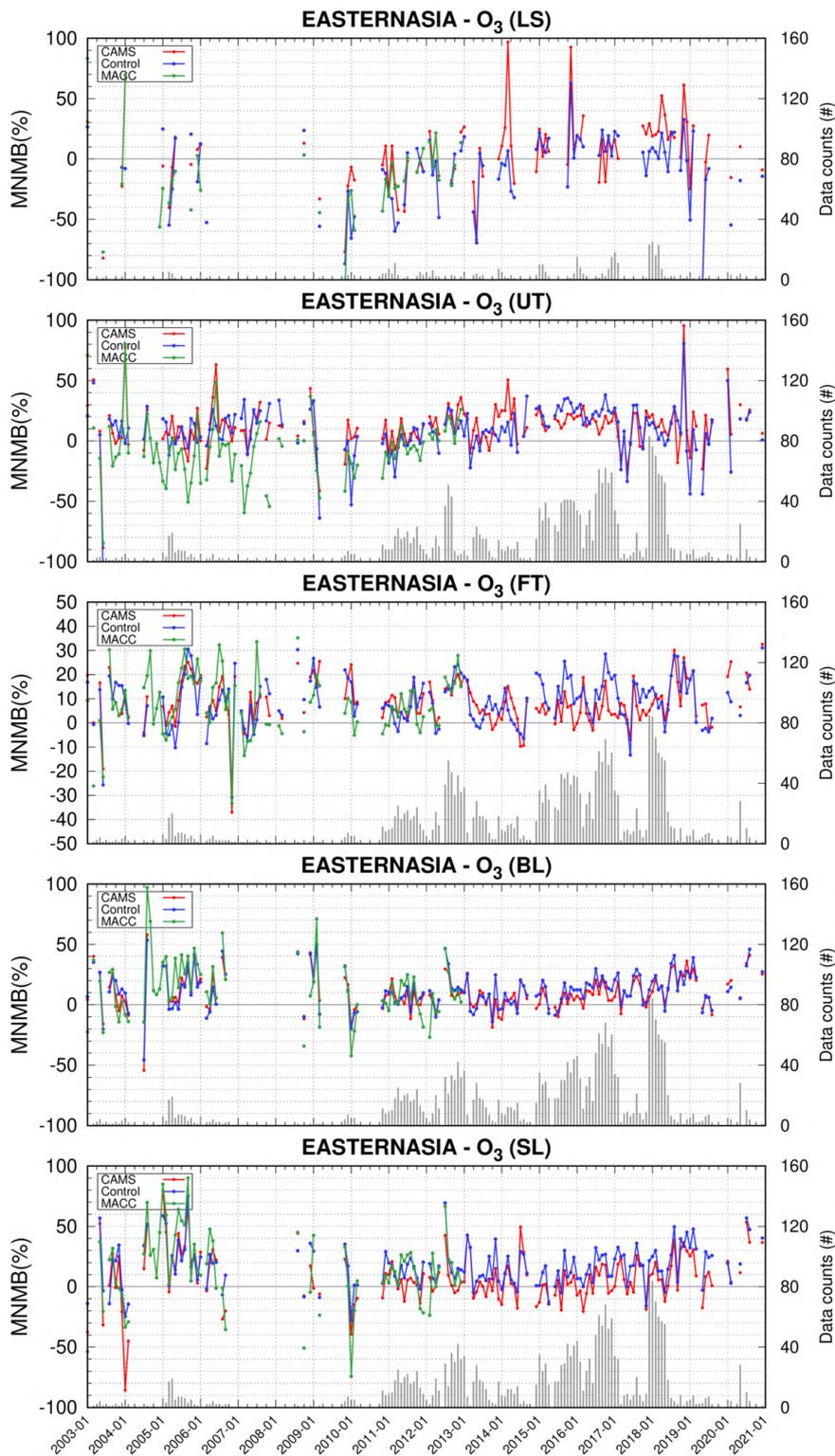


Figure 4.4.13: Time series of the monthly scores (MNMB) for ozone in different atmospheric layers over Eastern Asia during the period 2003-2020 (black: IAGOS, red: CAMS reanalysis, blue: Control run, green: MACC reanalysis). From top to bottom for SL: Surface Layer, BL: Boundary Layer, FT: Free troposphere, UT: Upper troposphere, LS: Lower Stratosphere. Units: %.



which sometimes extends to the mid-troposphere. Moreover, considering all airports in this region, the seasonality of the bias seems to be more pronounced in the first years of the evaluation period than in the most recent years (Fig. 4.4.11 and 4.4.13), similar to what was observed in Europe at the Frankfurt airport. For the MACC reanalysis the results are similar to those of CAMS in the low troposphere, but with sometimes larger bias especially during the summer (Fig. 4.4.13). This is the case in Tokyo at the beginning of the evaluation period (Fig. 4.4.11).

In the mid-troposphere, like for most regions of the world, the agreement of the CAMS reanalysis with observations is in general better, with MNMB absolute values often below 15% (Fig. 4.4.10, 4.4.11 and 4.4.13). No notable difference is found for the bias when comparing with MACC results in the mid-troposphere at Eastern Asia airports (Fig. 4.4.11 and 4.4.13).

In the UTLS both reanalyses present a large bias. At the beginning of the evaluation period, no systematic under- or overestimations are found. In the most recent years (when observations are available) the CAMS reanalysis mostly overestimates in the UTLS, as also noted at airports from other regions. The MACC results are similar to those of CAMS in the UTLS with the exception of large underestimations in the upper troposphere before 2008 (Fig. 4.4.10, 4.4.11 and 4.4.13).

Like for India, for most airports in Eastern Asia the results for ozone from the CAMS reanalysis and control run are in general very similar throughout the full period 2003-2020 (Fig. 4.4.10, 4.4.11 and 4.4.13). However, at some airports control run presents larger biases than the CAMS reanalysis in the mid troposphere and UTLS, a result which is found for instance in Hong Kong in the period 2015-2017 (Fig. 4.4.10).

#### Southern and Western Africa

Ozone has been continuously measured by IAGOS (MOZAIC) at the southern hemisphere airport of Windhoek in Namibia (1700 m.a.s.l) during 2006-2009 and 2011-2012. Although less continuous, time series are also available at the coastal site of Luanda (also located in the southern hemisphere) and at Lagos located at the Gulf of Guinea (not shown here). The seasonality is well reproduced by both reanalyses at these airports, with a six-month difference in timing of the African dry season between regions north and south of the Equator (not shown). Some differences in the behaviour of the models, especially between the CAMS and MACC reanalysis, can be noted at Windhoek on Fig. 4.4.14.

At Windhoek, for the CAMS reanalysis, there is a clear seasonality in MNMB which extends from the surface up to the mid-troposphere, corresponding to large underestimations in the wet season (November-April), and small overestimations in the dry season (May-October) as shown in Fig. 4.4.14. In the lowest layers the agreement is better during the dry season (i.e. when fires are common) than during the wet season, with MNMB in the range of  $\pm 16\%$  during the dry season against MNMB of in the range of  $-40\%$  and  $-20\%$  during the wet season.

For the MACC reanalysis the agreement with the observations in the lower troposphere is in general slightly better than that of the CAMS reanalysis especially during the wet season, while in the dry season the performance of both runs are more similar (Fig. 4.4.14).

In the mid-troposphere both reanalyses perform well, with a slightly better performance by the CAMS reanalysis especially during the dry season. The largest biases are found in the UTLS for both

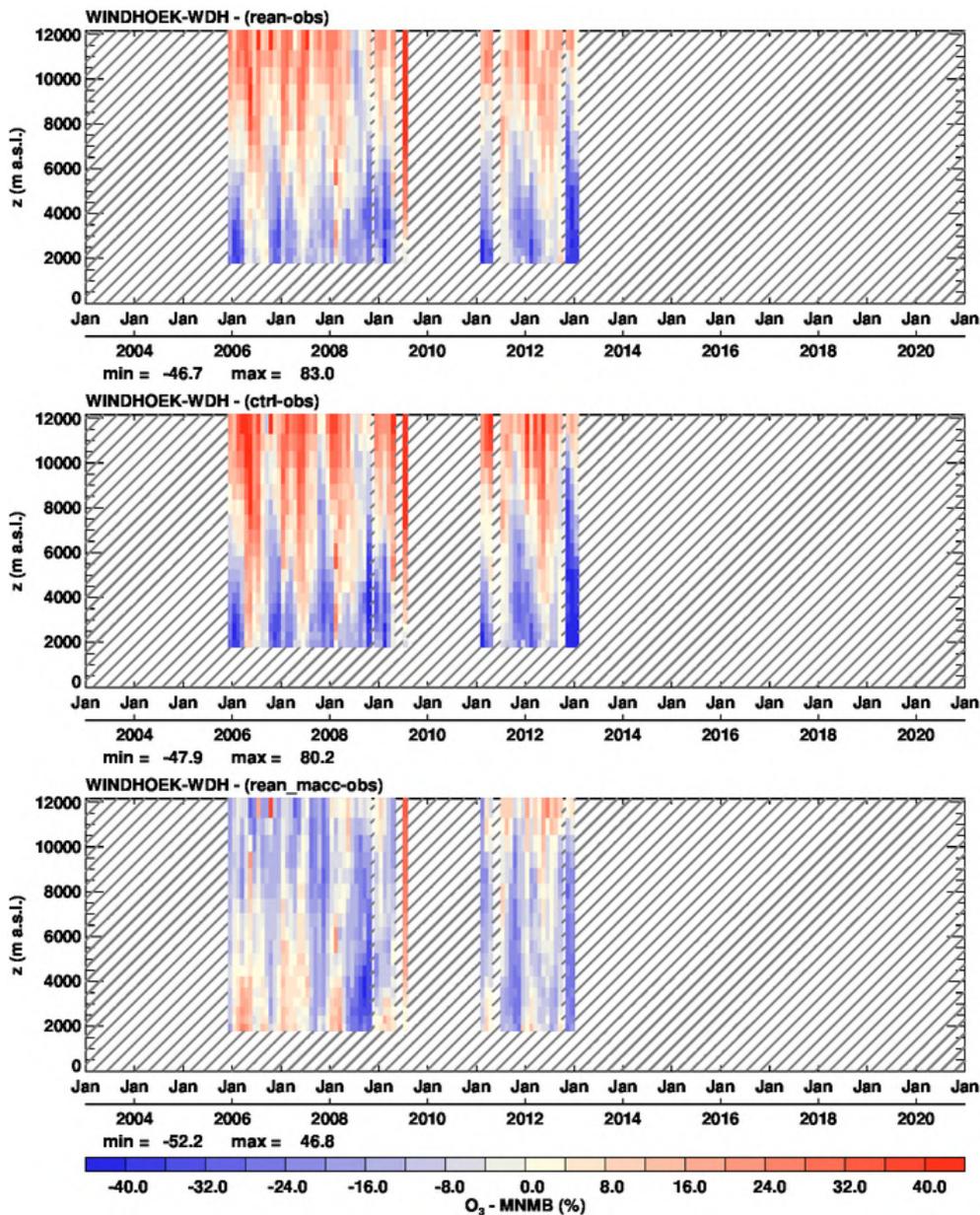


Figure 4.4.14: Time series of the monthly scores (MNMB) for ozone at Windhoek during the period 2003-2020. From top to bottom panel, the scores are for: CAMS reanalysis, control run and MACC reanalysis. Units: %.

CAMS and the MACC reanalysis. In the UTLS, the models behave differently until 2008 and seem to provide more similar results in the last years of available observations (2011-2012).

The control run presents a similar behaviour to that of the CAMS reanalysis for all layers but with larger biases in the UTLS as shown at Windhoek (4.4.14).

The bias results at Luanda (not shown) and Lagos are similar to those of Windhoek, moreover in the lowest layers (sea level) the bias is mostly positive and often with large values (>36%) throughout the year.

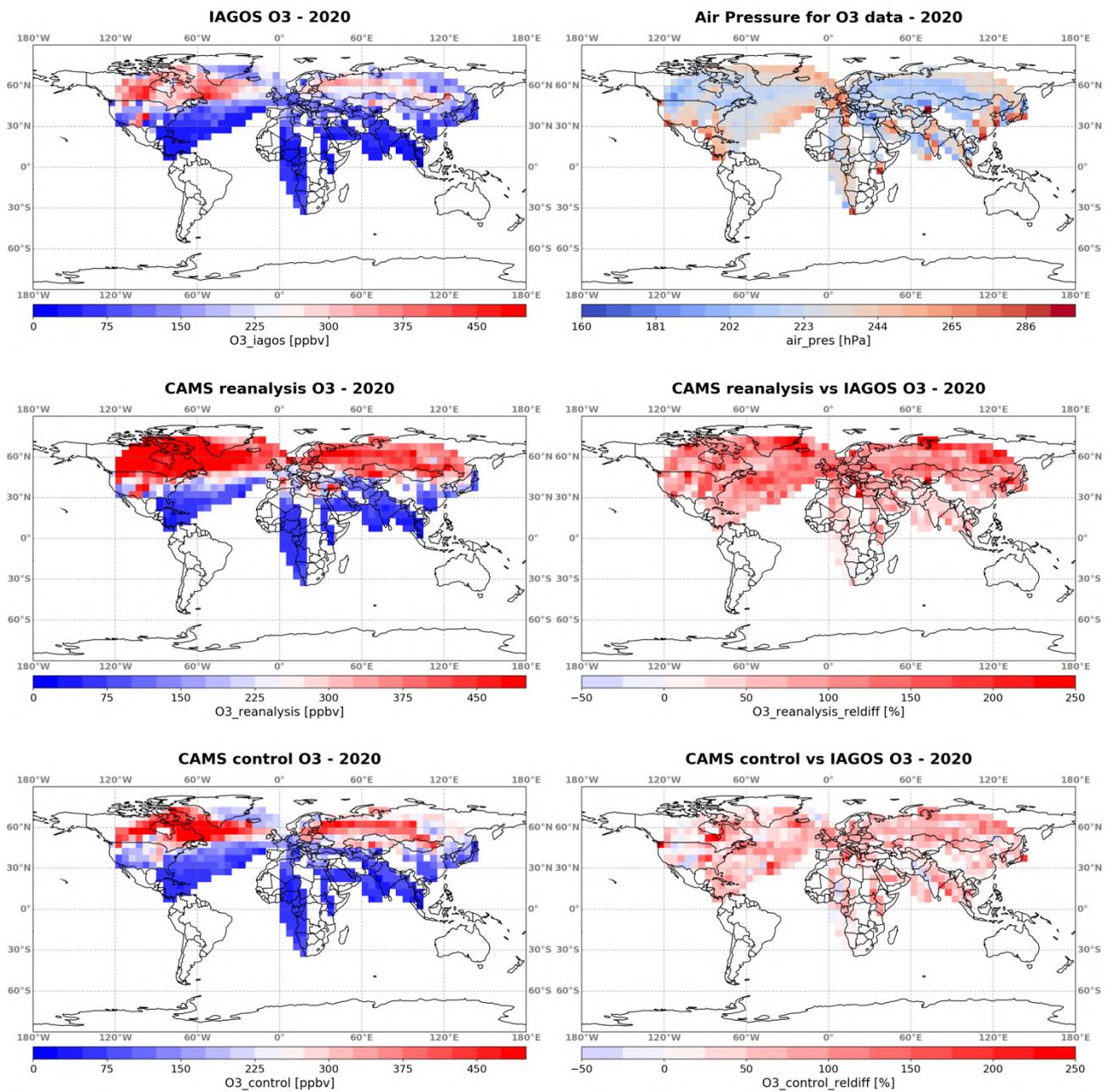


Figure 4.4.15: Global maps of gridded averages ( $5^{\circ} \times 5^{\circ}$ ) for ozone comparison with IAGOS cruise data for the year 2020. From left to right, top panel: IAGOS CO (in ppbv) and air pressure (in hPa). Bottom panel: CO from the CAMS reanalysis (in ppbv) and associated relative differences (in %) with respect to IAGOS.

Fig. 4.4.15 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 configurations. The CAMS reanalysis overestimates ozone in the midlatitudes of northern hemisphere where most observations are available. The largest overestimations appear mostly over the highest latitudes with bias values reaching 250 % while it is mostly below 150% at midlatitudes. The smallest values of the bias are found over western and southern Africa, as well as over India and Eastern Asia, with a bias in the range of  $\pm 50\%$ .

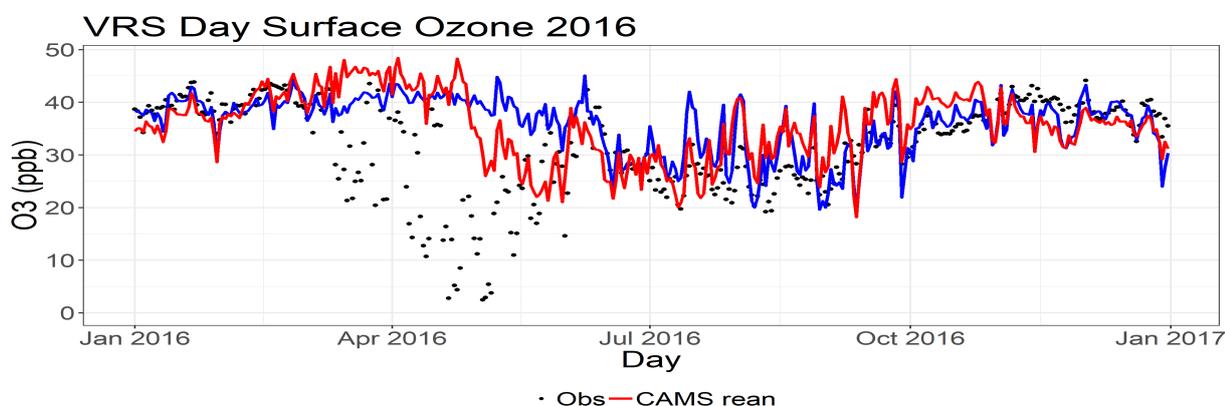


Figure 4.5.1: Daily averaged surface O<sub>3</sub> concentrations for 2016 for the high Arctic site Villum Research Station, Northern Greenland. Red: reanalysis; blue: control run; black: observations.

In the northern hemisphere midlatitudes, similarly to the CAMS reanalysis the bias from control run is positive but clearly smaller than for the CAMS reanalysis with values below 100%. Moreover, in the case of control run high northern latitudes do not present higher bias pattern than northern midlatitudes, although it was the case in the CAMS reanalysis. Over western and southern Africa the bias from control run is small like for the CAMS reanalysis but with some negative values. Over India and Eastern Asia, the bias from control run show larger positive bias than that of the CAMS reanalysis with values mostly between 50 and 150%, as well as some negative values on the western part of India.

#### 4.5 Verification with observations in the Arctic

The predicted reanalysis and control surface O<sub>3</sub> concentrations are compared to measurements from 11 arctic sites, six from the high Arctic and five from the European Arctic. Data for parts of 2020 are not quality controlled.

Four of the high Arctic sites are coastal stations (Alert, Barrow and Villum Research Station) or relatively close to the coast (Svalbard), and the measurements from these sites are influenced by ozone depletion events in spring related to halogen chemistry reactions (Fig. 4.5.1). These reactions are not represented in the model simulations and the model is not able to capture the low

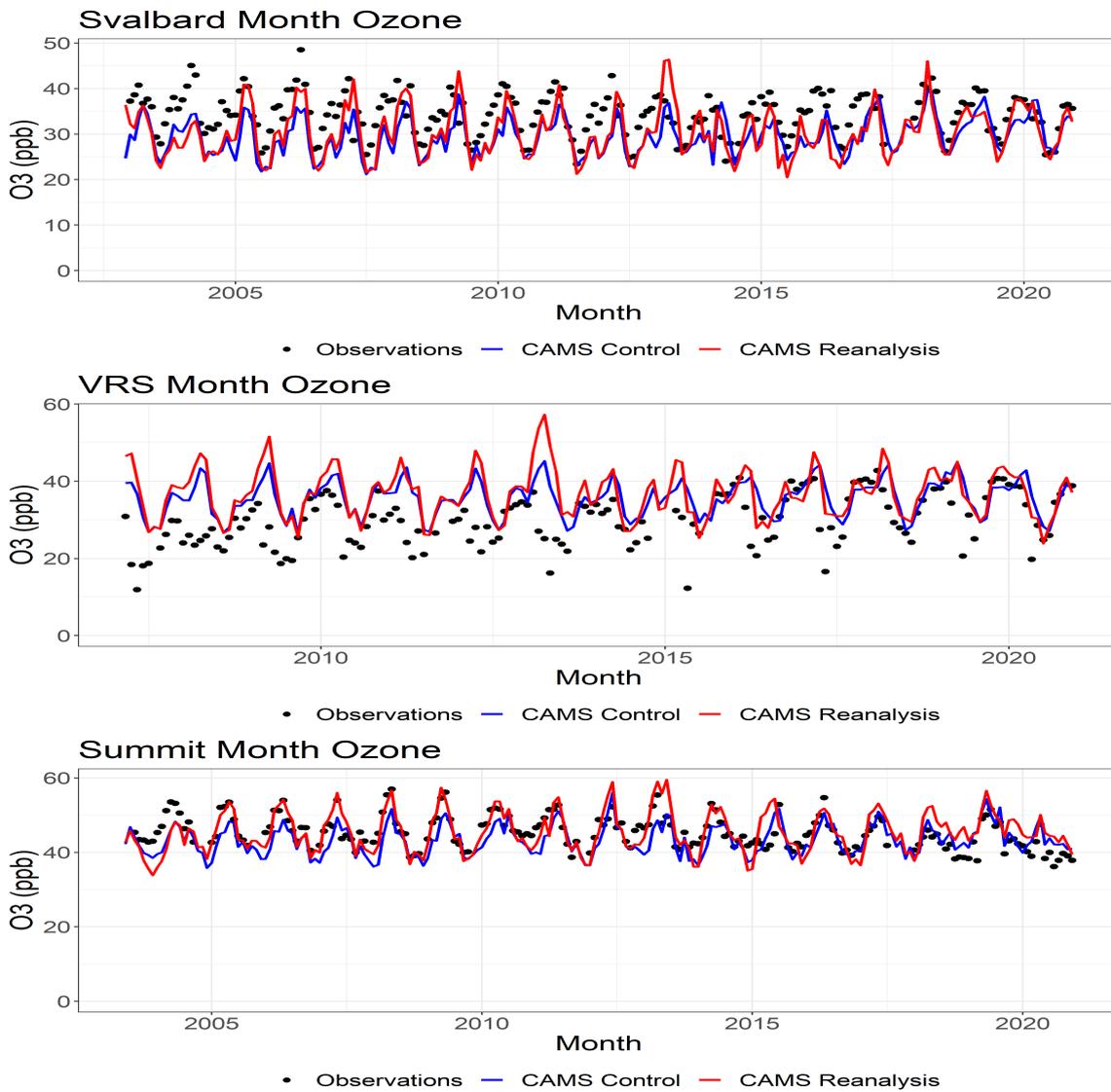


Figure 4.5.2: Time series for surface O<sub>3</sub> in the period 2003-2020 for the high Arctic sites Svalbard (Level 53), Villum Research Station (Level 60) and Summit (Level 58). Red: reanalysis; blue: control run; black: observations.

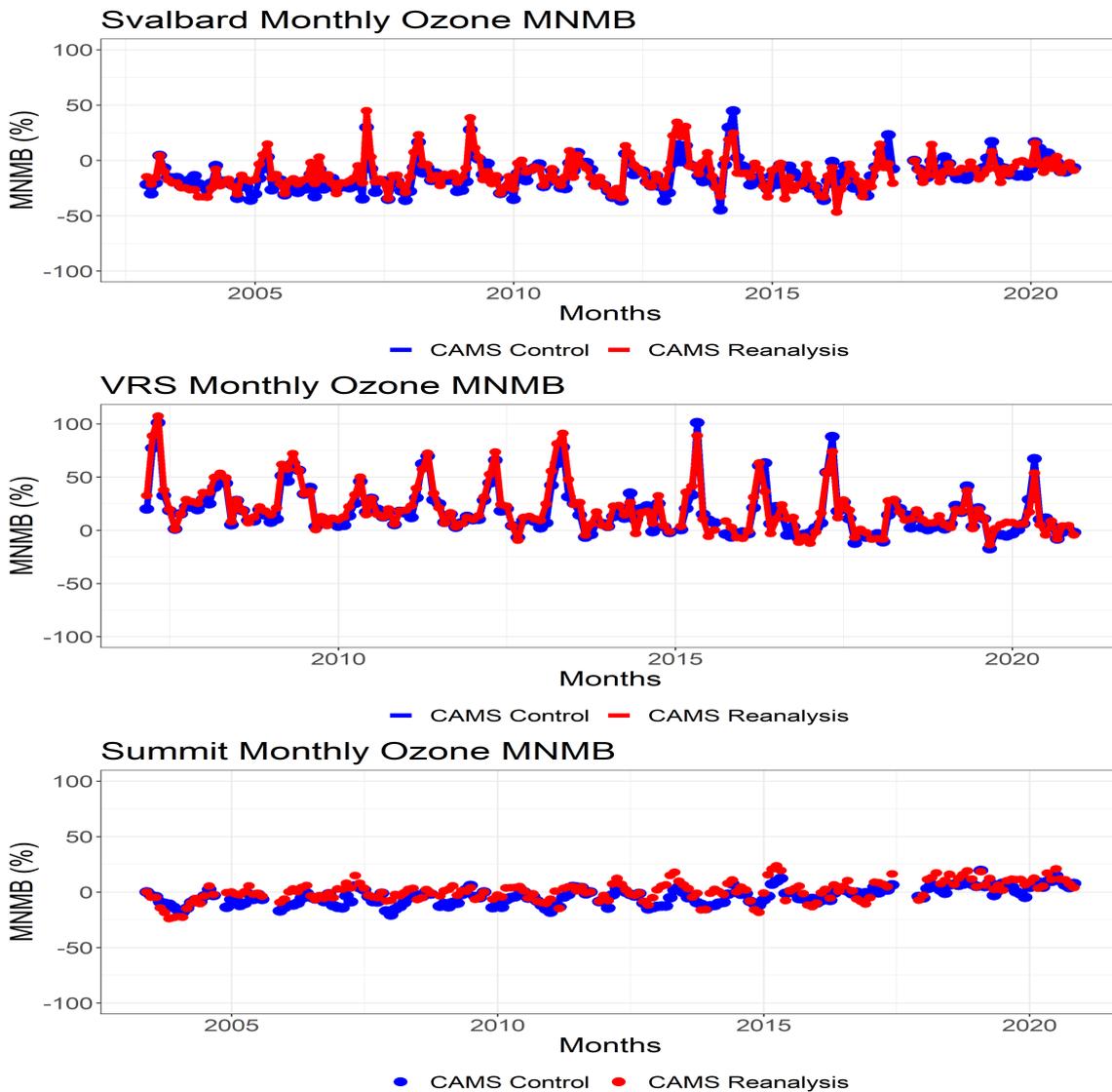


Figure 4.5.3: Time series of monthly values of MNMB for the period 2003-2020 for the high Arctic sites Svalbard, Villum Research Station and Summit. Red: reanalysis; blue: control run.

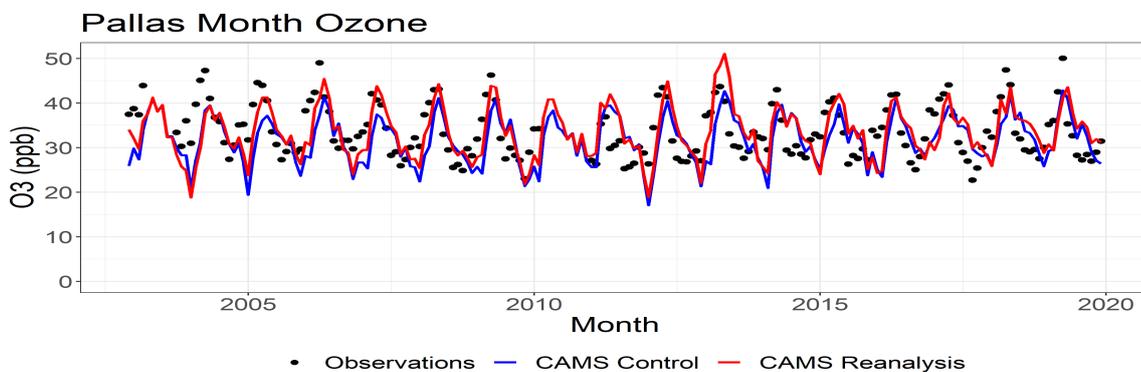


Figure 4.5.4: Time series for O<sub>3</sub> in the period 2003-2020 for the European Arctic site Pallas, Finland. Red: reanalysis; blue: control run; black: observations.

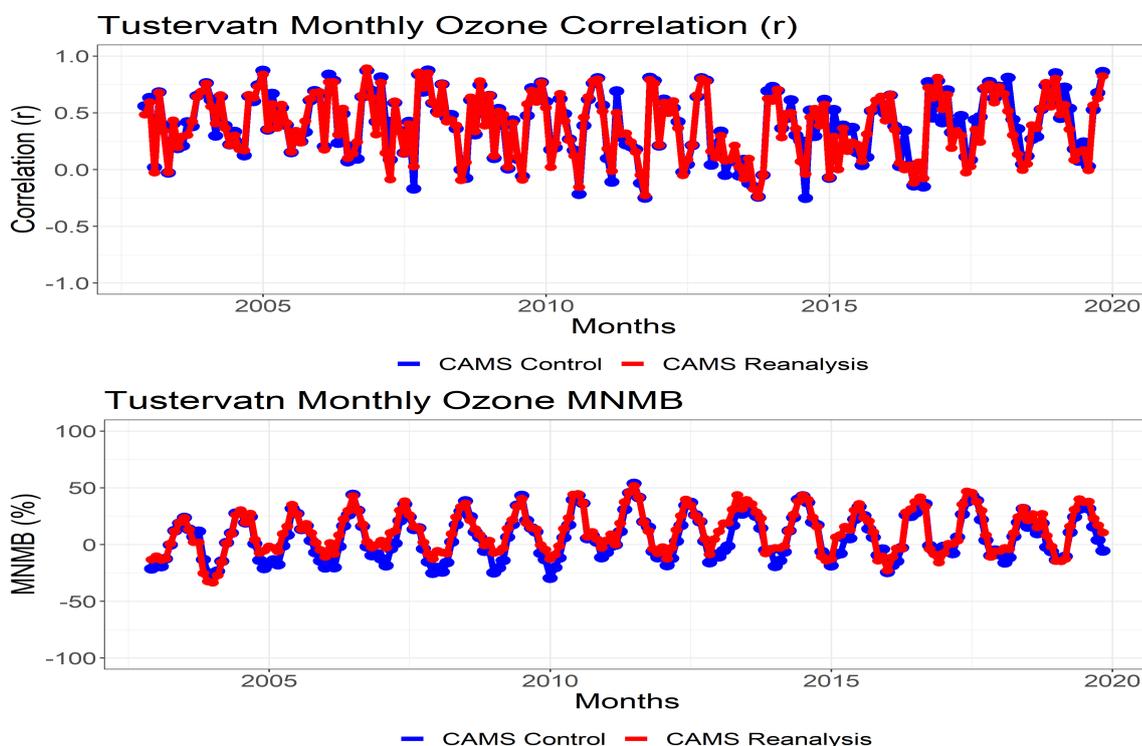


Figure 4.5.5: Time series of monthly values of  $r$  (top) and NMB (bottom) for Tustervatn, Norway in the period 2003-2020. Similar results are found for the other European Arctic sites. Red: reanalysis; blue: control run.

concentrations measured in spring at these sites. The episodes are less pronounced at Svalbard because the station is placed at a mountain some distance from the coast. The levels and variability are well captured in the other seasons, but the resulting bias is high, and the correlation is low (Table 4.5.1). At Summit, central Greenland, the level (MNMB = 0.01) and the annual variation is well captured ( $r = 0.64$ ) apart from an underestimation of the measurements in the first half year. The performance of the model improves over the simulated period with positive trends in the bias at Svalbard and Summit and a negative trend in the bias at VRS (Figure 4.5.3). No trends are seen at the other three high Arctic sites. Differences between the reanalysis and the control simulations are generally small, but the control run performs slightly better in terms of bias and correlation than the reanalysis at the high Arctic sites.



Table 4.5.1. Normalised Mean Bias (NMB) and correlation coefficient (r) of the reanalysis and the control simulations for the sites Alert, Nunavut, Canada; Barrow, Alaska, US; Svalbard, Norway; Villum Research Station (VRS) and Summit, Greenland; Esrange, Sweden; Karasjok, Norway; Oulanka and Pallas, Finland, and Tustervatn, Norway for the period 2003 –2020.

	Data coverage	Run	NMB	R
<b>Alert</b>	2003, 2005-2012, 2016-2018	reanalysis	0.15	0.36
		control	0.12	0.37
<b>Barrow</b>	2003-2005,2007-2015,2017-2020	reanalysis	0.07	0.20
		control	0.04	0.23
<b>Eureka</b>	2016-2020	reanalysis	0.12	0.14
		control	0.09	0.09
<b>Svalbard</b>	2003-2020	reanalysis	-0.12	0.41
		control	-0.13	0.41
<b>VRS</b>	2007-2020	reanalysis	0.20	0.31
		control	0.18	0.34
<b>Summit</b>	2004-2020	reanalysis	0.01	0.64
		control	-0.03	0.68
<b>Esrange</b>	2003-2018	reanalysis	0.09	0.56
		control	0.04	0.50
<b>Karasjok</b>	2003-2010	reanalysis	0.00	0.61
		control	-0.05	0.55
<b>Oulanka</b>	2003-2018	reanalysis	0.11	0.45
		control	0.06	0.40
<b>Pallas</b>	2003-2009, 2011-2019	reanalysis	-0.01	0.61
		control	-0.05	0.58
<b>Tustervatn</b>	2003-2019	reanalysis	0.11	0.45
		control	0.07	0.41

The pattern of the O<sub>3</sub> concentrations is the same at all European Arctic sites. The reanalysis tends to underestimate the measured levels in late winter and spring and to overestimate them in summer and autumn at the sites (Fig. 4.5.4). The average O<sub>3</sub> mixing ratios for the whole period are overestimated (-1%-11%). The model generally captures high concentration episodes in spring and summer although the peaks are underestimated. The correlation varies between 0.45 and 0.61 for the European Arctic sites. The correlations are slightly better for the reanalysis run than for the control run for the European Arctic sites, while the bias is better at three of the five sites but worse at the other sites. The performance of the model to predict O<sub>3</sub> concentrations over the European Arctic tend to deteriorate with increasing bias over the full simulated period. The Correlation decreases over the first part of the simulated period and increases in recent years (Fig. 4.5.5).

### **Transport episodes**

In April and May 2006, a large number of fires occurred in the Baltic countries, western Russia, Belarus, and the Ukraine. Meteorological conditions with low pressure over Iceland favored atmospheric transport towards the Arctic in the end of April and beginning of May. Two episodes, around the 27 April and 3 May respectively were identified with increased measured levels of halocarbons, black carbon, CO<sub>2</sub>, O<sub>3</sub> and CO, as well as levoglucosan and potassium at Svalbard (Stohl

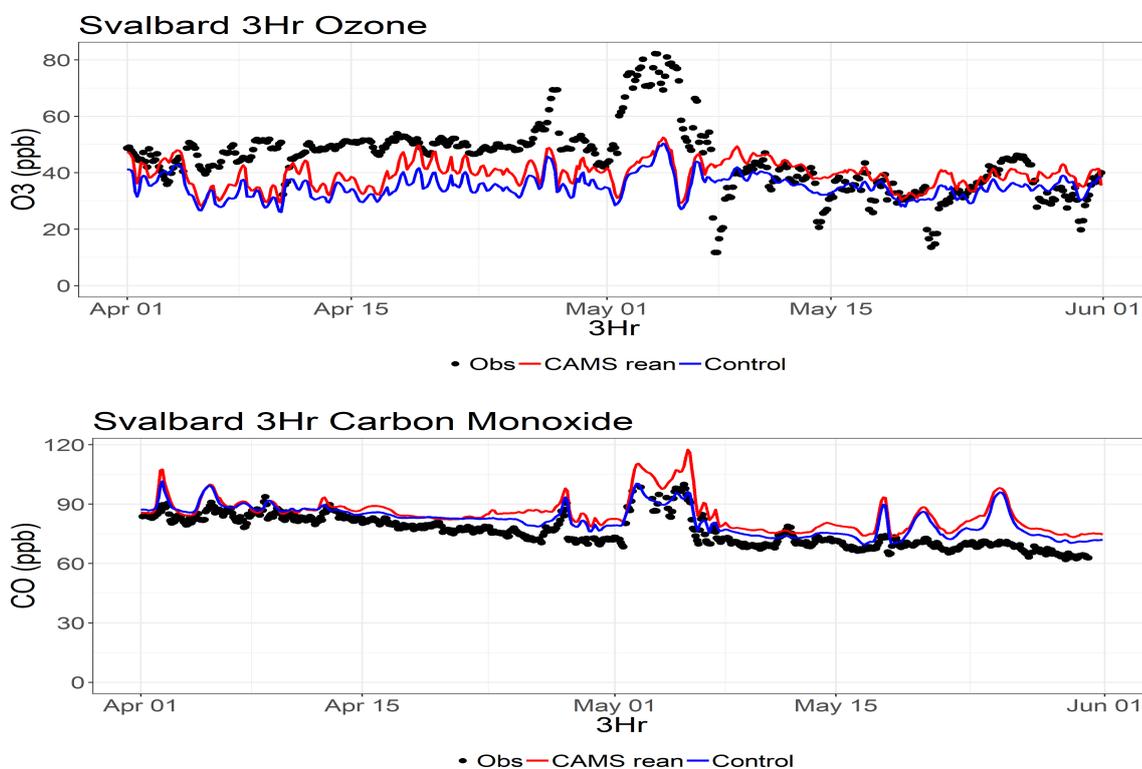


Figure 4.5.6: Surface O<sub>3</sub> (top) and CO (bottom) measurements (black dots) reanalysis simulation (red) and control (blue) for the Zeppelin Mountain in Svalbard for the period April-June 2006.

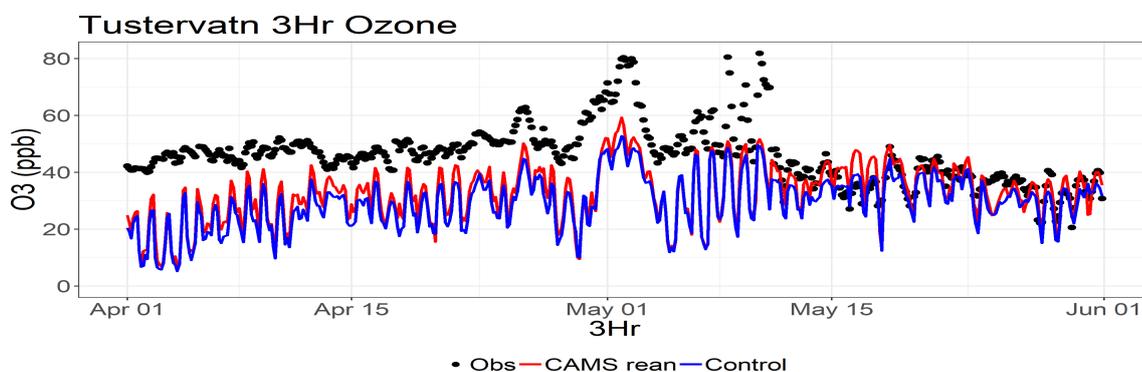


Figure 4.5.7 Surface O<sub>3</sub> (top) and CO (bottom) measurements (black dots) reanalysis simulation (red) and control (blue) for Tustervatn in Norway for the period April-June 2006.

et al., 2006). Peaks are seen in the simulated ozone concentrations for the two episodes, but the model fail to capture the increased O<sub>3</sub> concentrations. For CO, both episodes are captured well by the reanalysis simulation (Figure 4.5.6).

The episodes passed over the northern Scandinavia approximately one day before they reached Svalbard. Both episodes were captured by monitoring stations in Norway and Finland, where a peak in ozone concentrations were measured e.g., at Tustervatn (Figure 4.5.7). These high ozone episodes are also seen at the sites of Esrange, Karasjok, Oulanka and Pallas. The transport episodes are captured by the model although the peak concentrations are underestimated.

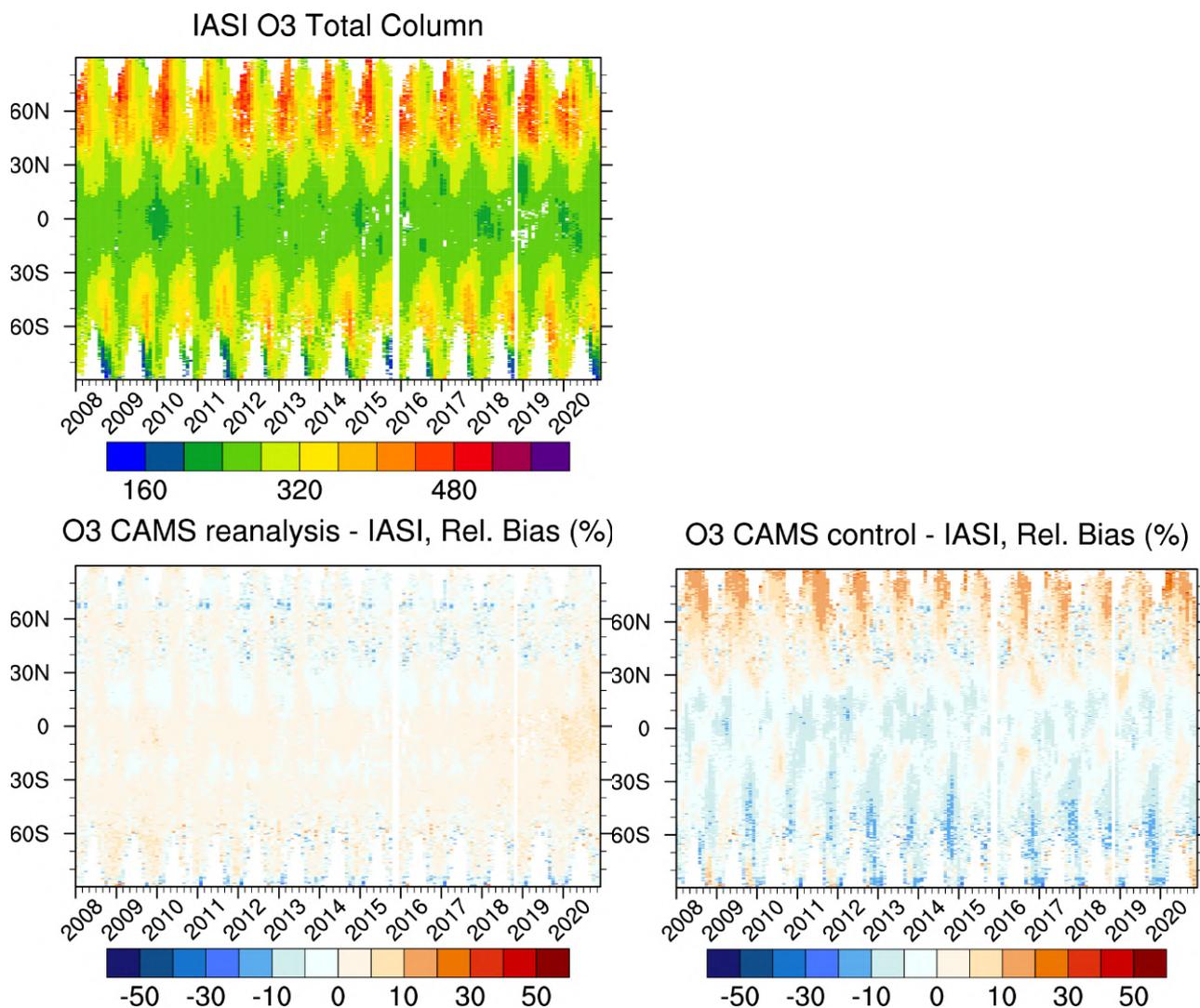


Figure 4.6.1: IASI Metop-A O<sub>3</sub> total column (top) as function of latitude and time from January 2008 to December 2020. Relative difference between CAMS reanalysis and IASI (bottom left) and control and IASI (bottom right).

#### 4.6 Comparison of the CAMS reanalysis with IASI O<sub>3</sub>

CAMS reanalysis and control O<sub>3</sub> total column data are compared with IASI Metop-A satellite observations (Clerbaux et al., 2009). For the comparison with the IASI data, the vertically integrated model O<sub>3</sub> data were transformed using IASI averaging kernels (Rodgers, 2000).

Fig. 4.6.1 shows total column of O<sub>3</sub> as a function of latitude and time from IASI and the relative difference between the model runs and IASI. The reanalysis captured well the high O<sub>3</sub> values at middle and high northern latitudes during winter/spring seasons and low values over the Southern Hemisphere polar region during ozone hole seasons in autumns. The reanalysis bias is within 5%. The control run shows an overestimation of high O<sub>3</sub> values over the mid- and high northern latitudes up to 20% and a slight overestimation of the low values over the southern polar region in autumns within 10%. Underestimation within 10% can be seen over the southern midlatitudes in autumn seasons. In year 2020, the reanalysis indicates slightly higher positive bias over the low-

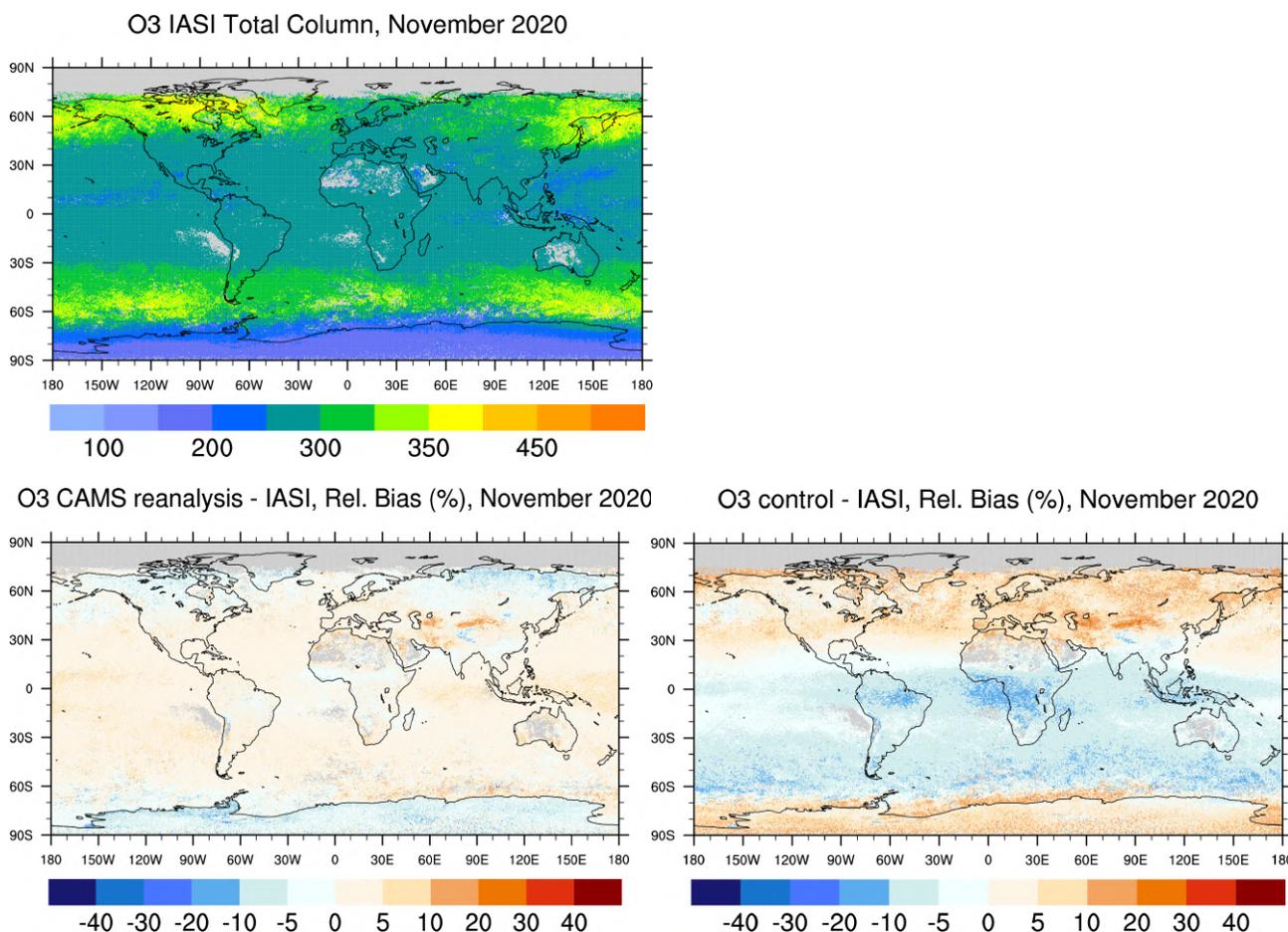


Figure 4.6.2: Ozone total column for IASI Metop-B (top) satellite observations and relative difference between the reanalysis and IASI (bottom left) and control and IASI (bottom right) for November 2020. Grey colours indicate missing values.

latitudes (within 10%). The control run shows slightly higher overestimation of low O<sub>3</sub> values over the high northern latitudes in spring (up to 30%) and over Antarctica (up to 20%) in autumn season. Note, that from 4 December 2019 onwards, the number of processed IASI O<sub>3</sub> observations was reduced.

Global distribution of O<sub>3</sub> total column obtained from IASI, as well as the relative difference between the CAMS runs and IASI, are shown in Fig. 4.6.2 for November 2020. IASI shows relatively high O<sub>3</sub> over the north of Canada and north-east of Russia and over southern mid-latitudes and low values over Antarctica and some equatorial regions. The reanalysis captures the high as well as low O<sub>3</sub> values and is in good agreement with IASI observations, showing general biases within 10%. The control run shows a positive bias over the mid- and high northern latitudes and Antarctica (up to 20%) and a negative bias in other parts of the Globe (up to 20%).



## 5. Carbon monoxide

### 5.1 Validation with Global Atmosphere Watch (GAW) Surface Observations

In the following, an evaluation of model performance for the CAMS reanalysis run has been carried out for the period January 2003 to December 2020. GAW hourly data from 21 stations from the World Data Centre for Greenhouse Gases (WDCGG) has been used for model verification for the period 2003-2020, see Table 5.1.1. For the years 2016 -2020 near-real-time station data has been used.

Figure 5.1.1 shows monthly MNMBs (calculated from 3-hourly data) averaged over stations in Europe, North America and East Asia.

For North America, monthly MNMBs range between 30% and -15%. Episodes with large positive MNMBs are the years 2005 and 2006, during spring and summer, likely related to overestimations of fire emissions in the model (Fig. 5.1.2). From 2011 onwards, MNMBs show an increasing trend up until 2018, likely related to the emissions in the model. Especially during the end of 2013 (November) there seems to be a problem with the fire emissions, see Fig. 5.1.4.

Surface CO for GAW stations in Europe is mostly slightly underestimated by the model with annual MNMBs between 10% and -20%, see Fig. 5.1.1. Discontinuities in the time –series of MNMBs show during 2006 and 2010, where larger negative biases appear. From 2011 onwards, especially for the control run, MNMBs show an increasing trend up until 2018. This might be related to the anthropogenic emissions but also biogenic and fire emissions. The seasonal evaluation reveals that larger biases appear mostly during summer and autumn, with 2010 being the year with the largest negative MNMBs. The seasonal variation of the MNMB is similar to the seasonal variation in biogenic and fire emissions and hints to emission related issues. The scaling of winter traffic emissions over North America and Europe (Stein et al., 2014) applied in the CAMS reanalysis, however, has largely improved the winter biases in the CAMS model compared with the MACC reanalysis (Fig. 5.1.1).

The MNMBs maps (Fig. 5.1.2) for annual MNMBs show that especially the stations in Eastern Europe (exposed to larger CO mixing ratios) show larger negative annual MNMBs.

For stations located in Asia CO, mixing ratios are mostly slightly underestimated by the model with monthly MNMBs between 13% and -18%. Positive MNMBs appear during the period 2006 to 2010 (Fig. 5.1.1). From 2011 onwards, same as for North America and Europe, there is a positive trend in the control run up to 2016, during spring and summer (Fig. 5.1.1), likely related to an overestimation of emissions in the model. Starting from the year 2014 the discrepancy between the reanalysis and control runs increased.

Correlation coefficients (see Fig. 5.1.3) are on average 0.7. Especially the North American stations (LAC, FRA, EGB, ETL) show larger standard deviations in the model than in the observations, likely due to an overestimation of fire signals in the model.

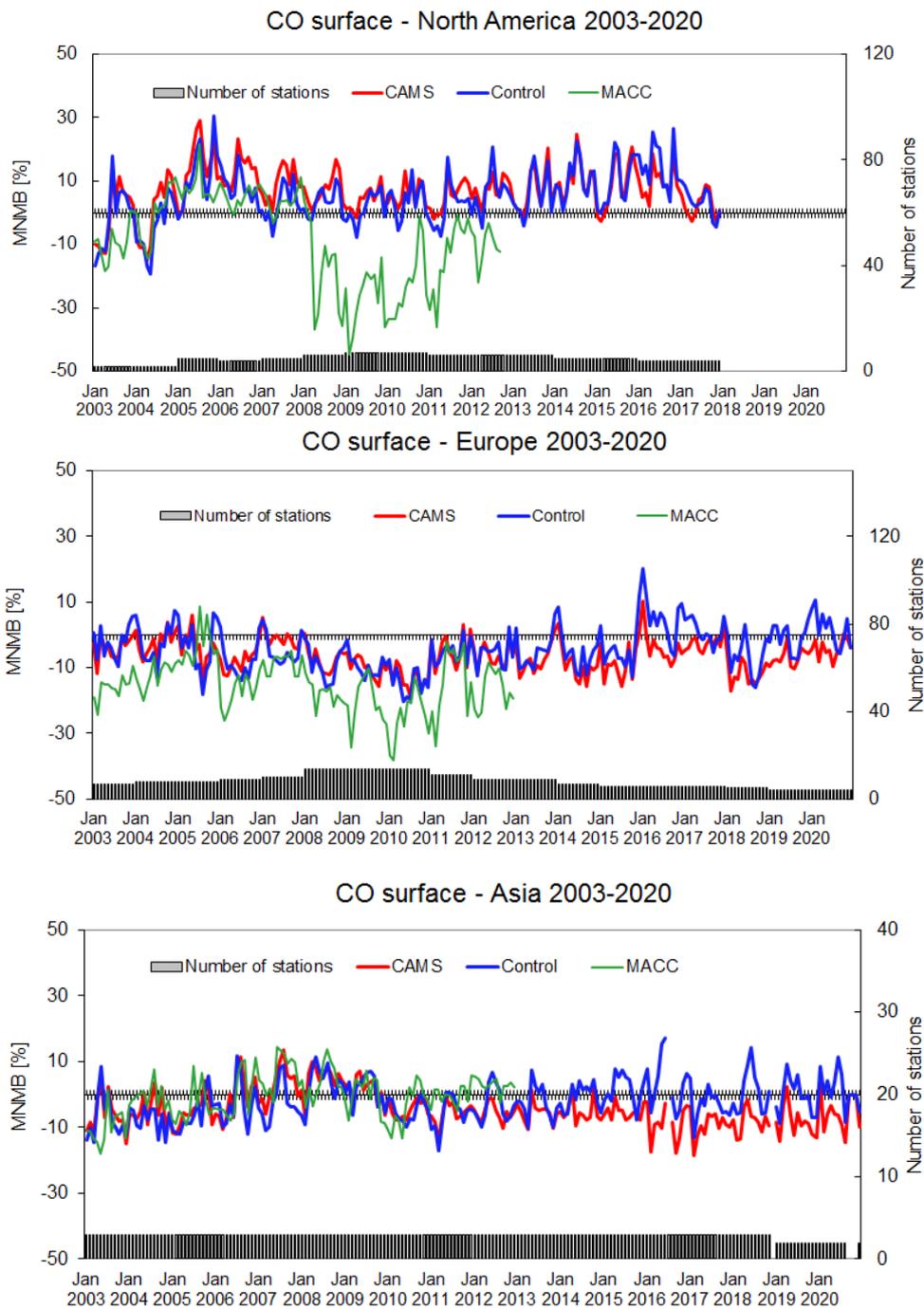


Figure 5.1.1: Mean Normalised bias (MNMB) for GAW stations calculated for the years 2003-2020 from 3-hourly data, averaged on a monthly basis, averaged over different areas (first row left: North America, first row right: Europe, second row: Asia). Red: CAMS reanalysis, green: MACC reanalysis, blue: Control run. The availability of observations reduces towards the end of the period.

Comparison with the MACC reanalysis

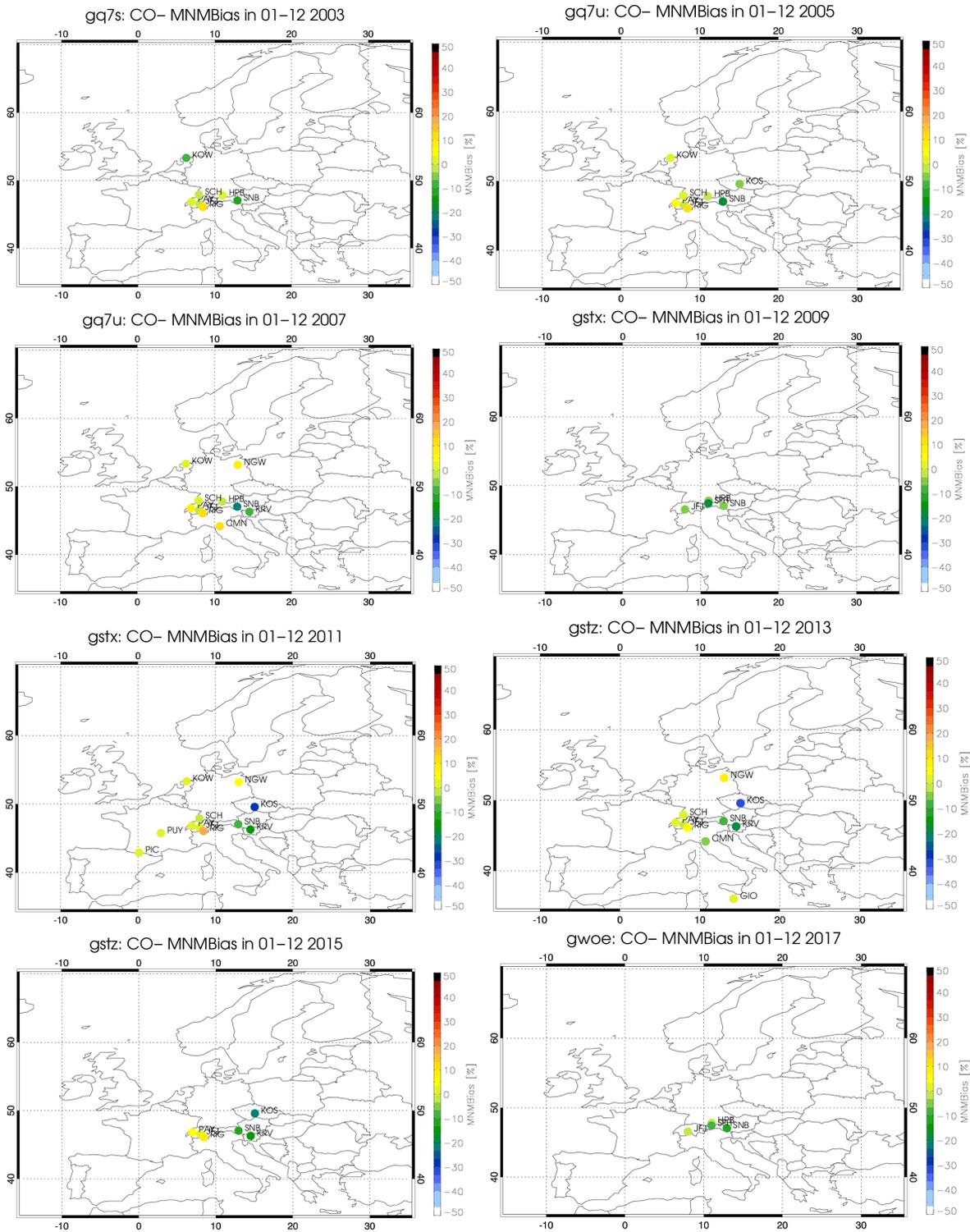
The MACC reanalysis in comparison shows much larger negative MNMBs for Europe and North America after 2008. The negative peak in bias relates to the assimilation of IASI data from April 2008 onwards.



Table 5.1.1: List of GAW stations used in the validation (CO)

Station	Name	Latitude	Longitude	Altitude	Year
Alert	ALT	82.45	-62.52	210	2003-2010
Cape Point	CPT	-34.35	18.48	230	2003-2020
Cape Verde	CVO	16.848	-24.871	10	2008, 2010-2015,2020
Chibougamau	CHI	49.6827	-74.342	393	2007, 2008, 2010
Monte Cimone	CMN	44.18	10.7	2165	2007, 2008,2010,2013,2014,2020
Egbert	EGB	44.23	-79.780	253	2005-2008, 2010-2015
Estevan Point	ESP	49.38	-126.55	39	2010-2015
East Trout Lake	ETL	54.3501	-104.9834	492	2005-2008, 2010-2015
Fraserdale	FRA	49.88	-81.57	210	2003-2015
Hohenpeissenberg	HPB	47.8	11.02	985	2003-2010,2016,2020
Izana	IZO	28.3	-16.5	2367	2008, 2009-2015
Jungfrauoch	JFJ	46.55	7.98	3578	2003-2020
Kollumerwaard	KOW	53.33	6.28	0	2003-2008, 2010-2012
Kosetice	KOS	49.58	15.08	534	2004-2006, 2008, 2010-2015
Krvavec	KRV	46.3	14.53	1720	2006-2008, 2010-2015
Lac La Biche	LAC	54.95	-112.45	540	2008, 2010-2015
Minamitorishima	MNM	24.28	153.98	8	2003-2020
Neuglobsow	NGW	53.17	13.03	65	2007,2008, 2010-2013
Payerne	PAY	46.82	6.95	490	2003-2008, 2010-2015
Pic du Midi	PIC	42.9372	0.1411	2877	2008, 2010,2011
Puy de Dome		45.7719	2.9658	1465	2008, 2010,2011
Rigi	RIG	46.07	8.45	1031	2003-2008, 2010-2015
Ryori	RYO	39.03	141.82	260	2003-2020
Sable Island	SAB	43.93	-60.02	5	2003-2008, 2010-2014
Schauinsland	SCH	47.92	7.92	1205	2003-2008, 2010-2014
Schneefernerhaus	SFH				2008-2010, 2016-2020
Sonnblick	SNB	47.05	12.95	3106	2003-2019
Ushuaia	USH	-54.83	-68.3	18	2007-2011, 2013-2020
Yonagunijima	YON	24.47	123.02	30	2003-2020

The Improvements in the CAMS reanalysis thus relate to changes in the data assimilation (only MOPITT data is assimilated) but also the difference in the models and the scaling of winter road traffic emissions could improve the negative biases seen for MACC for Europe (Fig. 5.1.1).



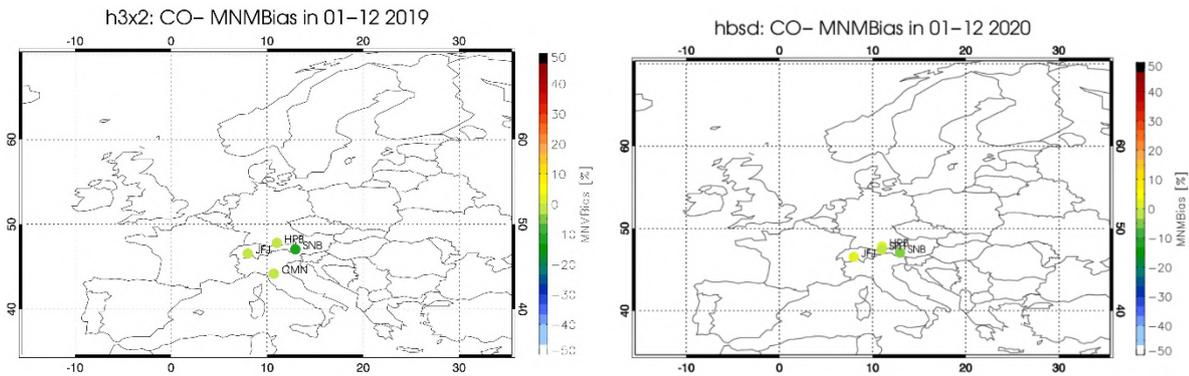
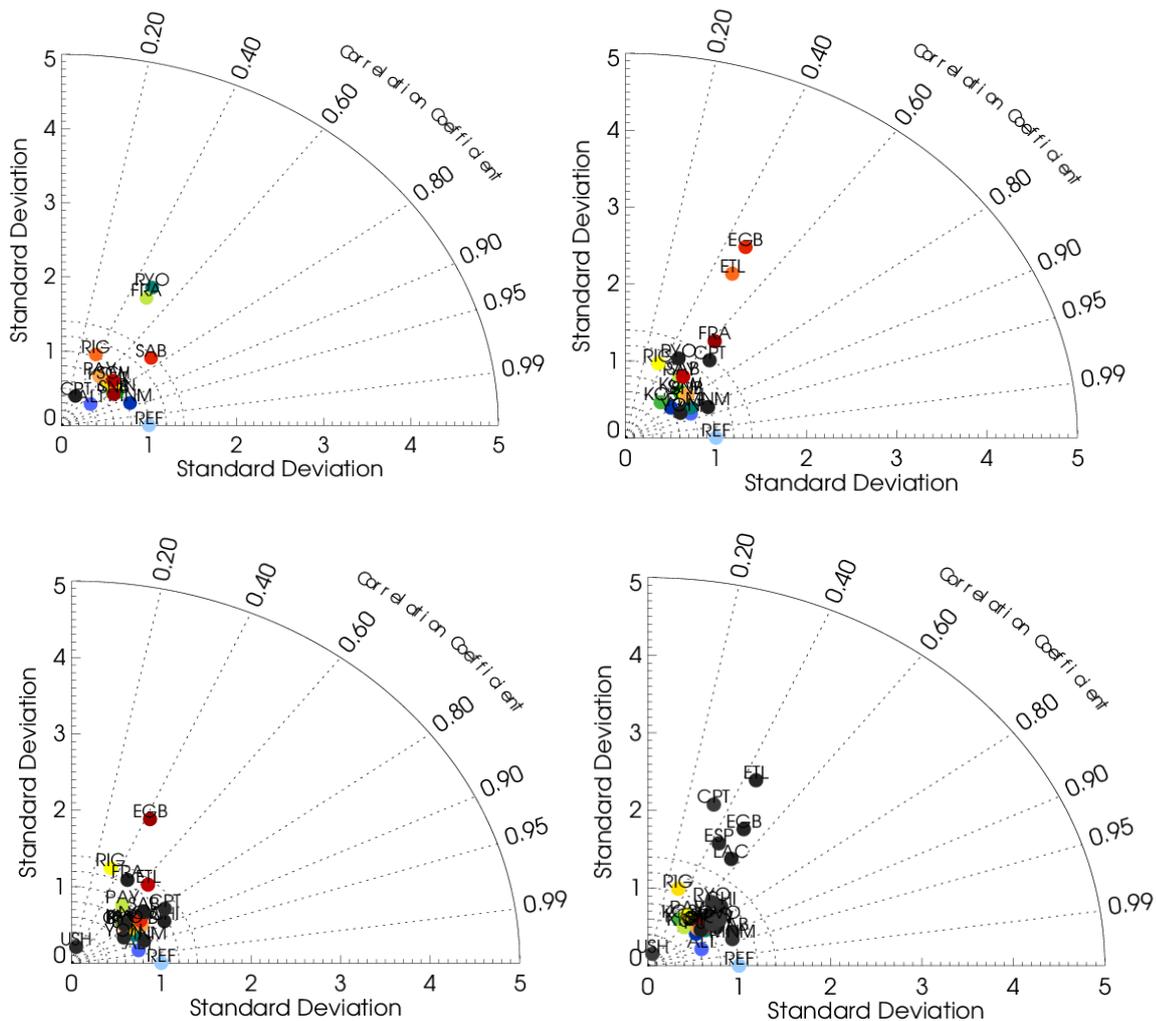


Figure 5.1.2: Distribution of overall MNMBs for the CAMS reanalysis against GAW for CO for the years 2003 (first row left), 2005 (first row right), 2007 (second row left), 2009 (second row right), 2011 (third row left), 2013 (third row right), 2015 (fourth row left), 2017 (fourth row right) and 2019 (last row left) and 2020 (last row right).



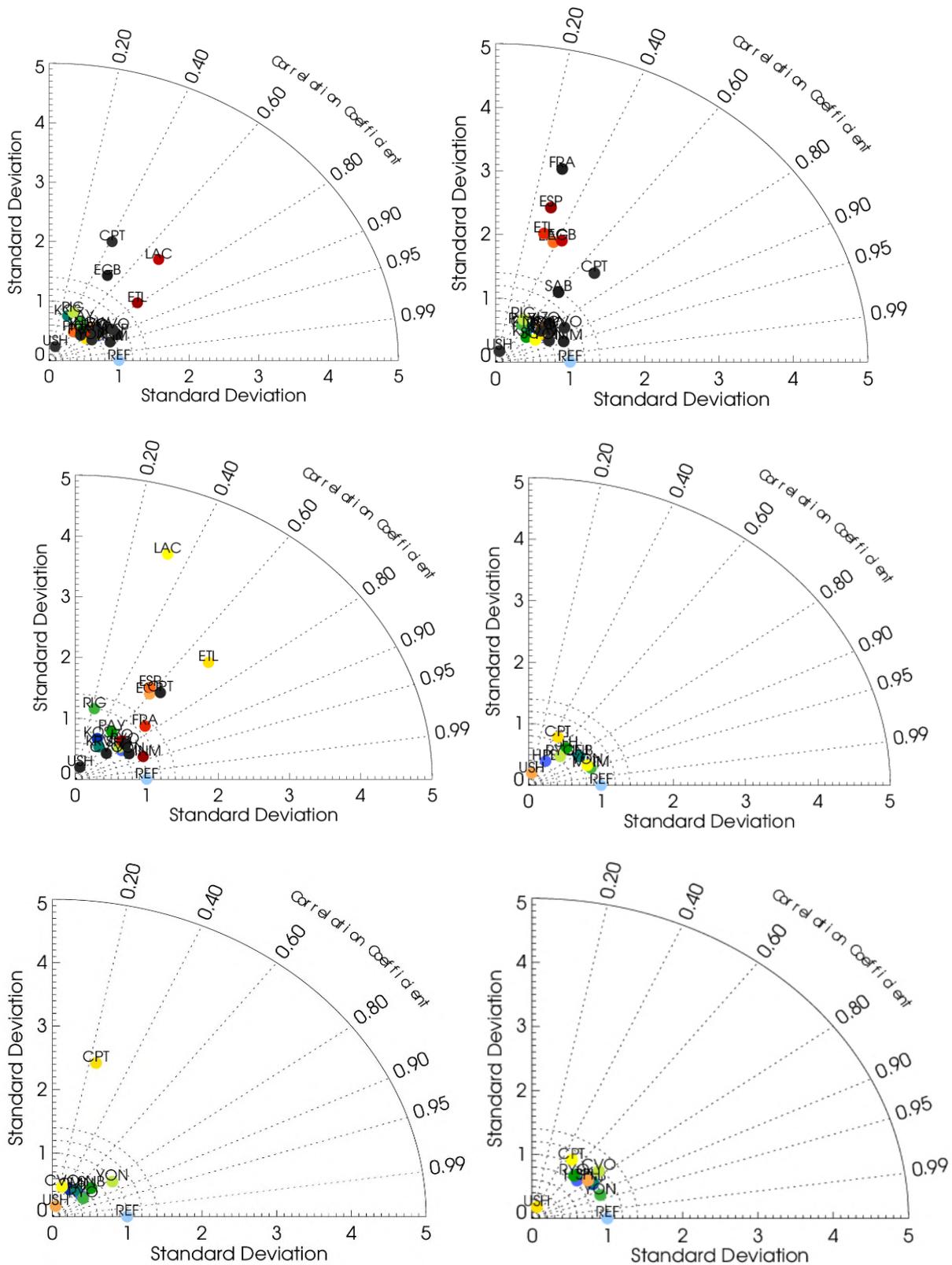


Figure 5.1.3: Taylor diagram with standard deviation and correlation for CO in 2003 (first row left), 2005 (first row right), 2007 (second row left), 2009 (second row left), 2011 (third row left), 2013 (third row right), and 2015 (fourth row left), 2017(fourth row right), and 2019 (last row left) and 2020 (last row right).

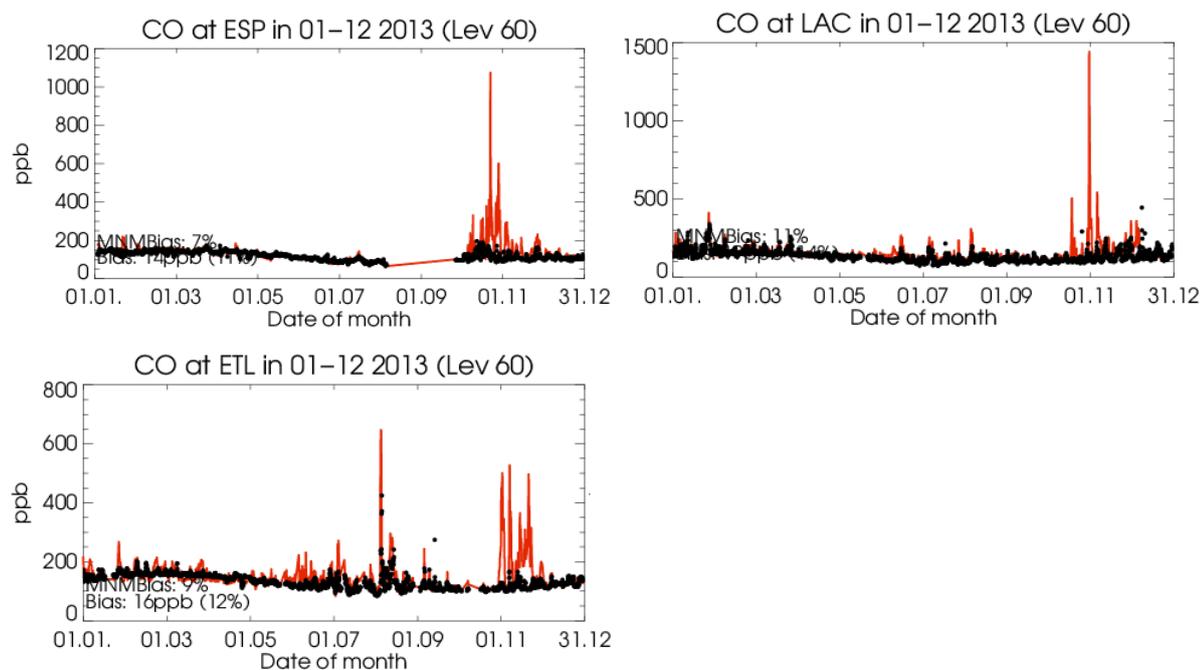


Figure 5.1.4: Time series plots for CO (in ppb) for GAW stations Estevan Point (EVP), Lac la Biche (LAC), and East Trout Lake (ETL) in Canada during 2013. Red: CAMS reanalysis, black: observations.

## 5.2 IAGOS Aircraft observations

The comparison results with IAGOS CO are presented by region for available airports similarly to what was done in the ozone section (see details section 4.4). The same methodology, evaluation periods and data levels as for ozone are used here for the following CO comparisons.

### Europe

Fig. 5.2.1 shows the time series of the monthly profiles for CO at Frankfurt during the period 2003 to 2020, for the two reanalysis, CAMS control run and observations. The time series of the monthly scores (MNMB) associated to all models are presented in Fig. 5.2.2. The results calculated for the different atmospheric layer and for the Western European region include mostly observation at the airport of Frankfurt which provides the densest and longest time series. These results are also shown in Fig. 5.2.3 and Fig. 5.2.4.

The time series of the observations (Fig 5.2.1) show a reduction in CO values from the surface up to mid-troposphere. To some extent this behaviour is also reproduced by the models but with underestimated CO in the surface, boundary layer and low troposphere. The magnitude of these underestimations is more clearly shown in MNMB time series (Fig 5.2.2 and 5.2.3).

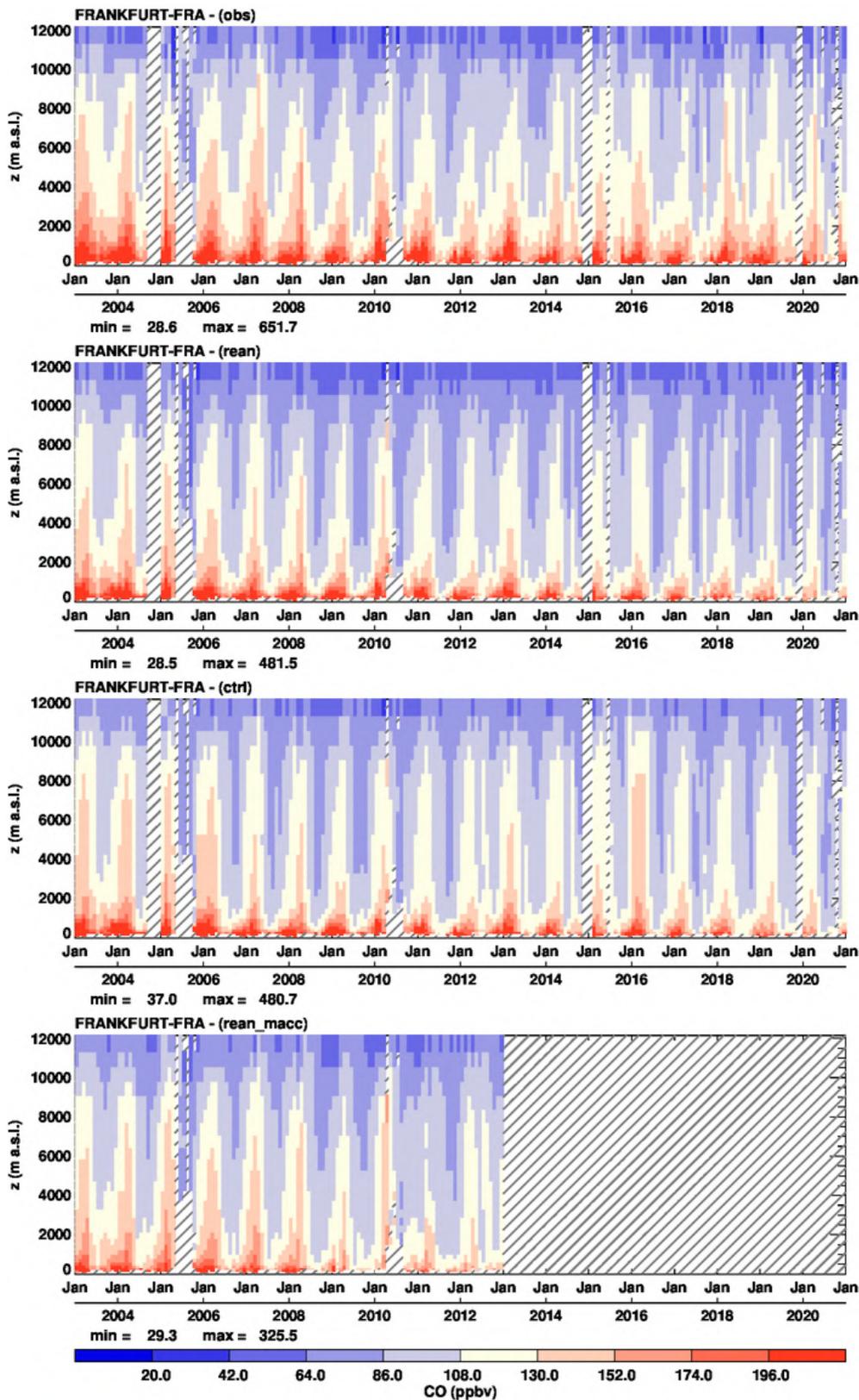


Figure 5.2.1: Time series of the monthly profiles for CO at Frankfurt during the period 2003-2020. From top to bottom panel the time series correspond to: IAGOS (MOZAIC), CAMS reanalysis, control run and MACC reanalysis. Units: ppbv.

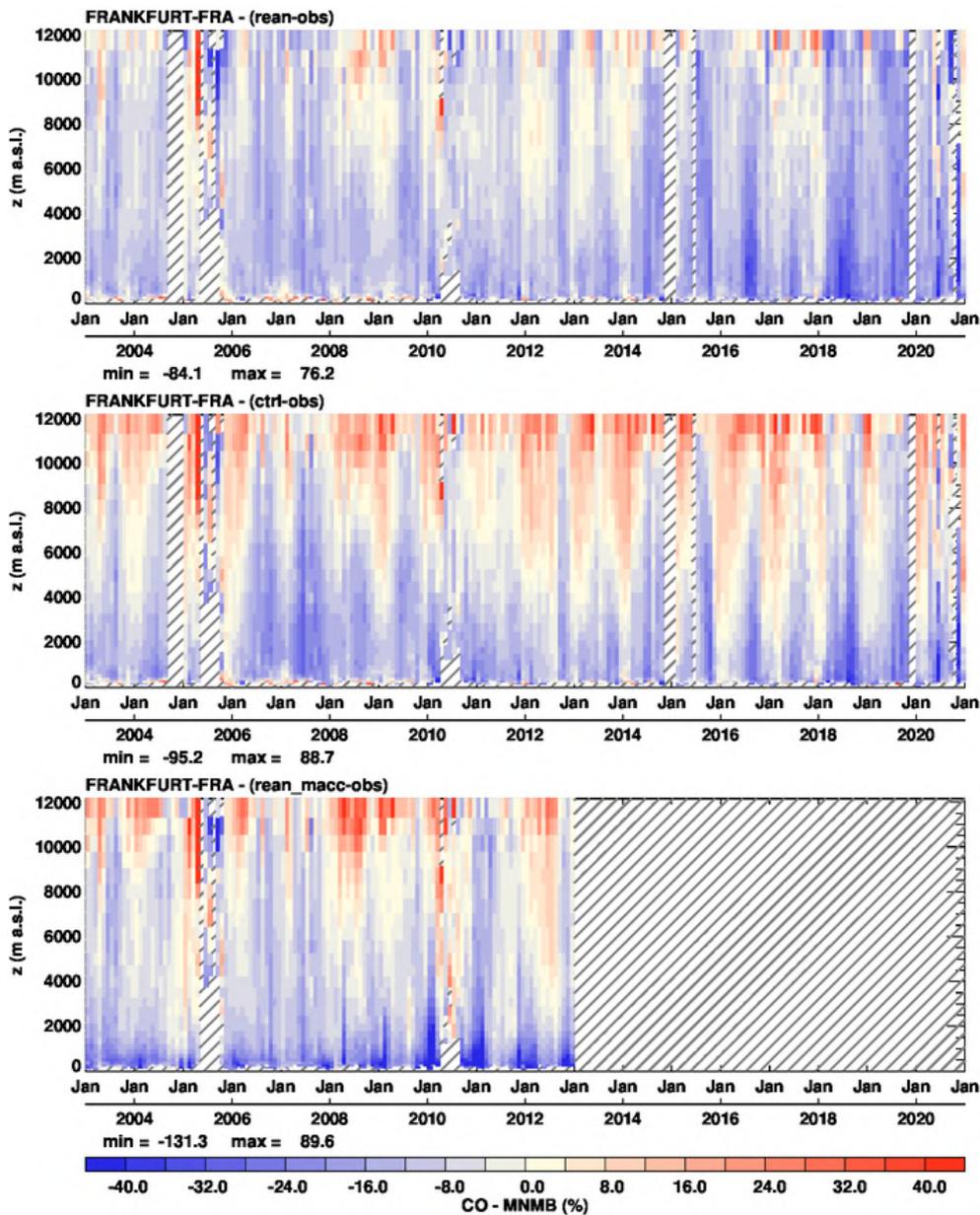


Figure 5.2.2: Time series of the monthly scores against IAGOS (MNMB) for CO at Frankfurt during the period 2003-2020. From top to bottom panel, the scores are for: CAMS reanalysis, control run and MACC reanalysis. Units: %.

According to Fig. 5.2.2 and 5.2.4, CO is underestimated by the CAMS reanalysis over nearly the whole atmospheric column and for all seasons with MNMB ranging between 15% and 25% at Frankfurt. In the mid-troposphere the agreement is good with biases smaller than 10%. This seems to be more systematic from 2009 on, and after the end of 2013 the performance seems to degrade again in the mid-troposphere. This could be related to the model assimilated MOPITT data. While in 2017 the bias seems to improve, in the years 2018-2019 the bias appears larger from the surface to the upper troposphere. This feature was already mentioned in the previous reanalysis report but here Level 1 data have been upgraded to Level 2 data for 2018 and 2019 and nevertheless the

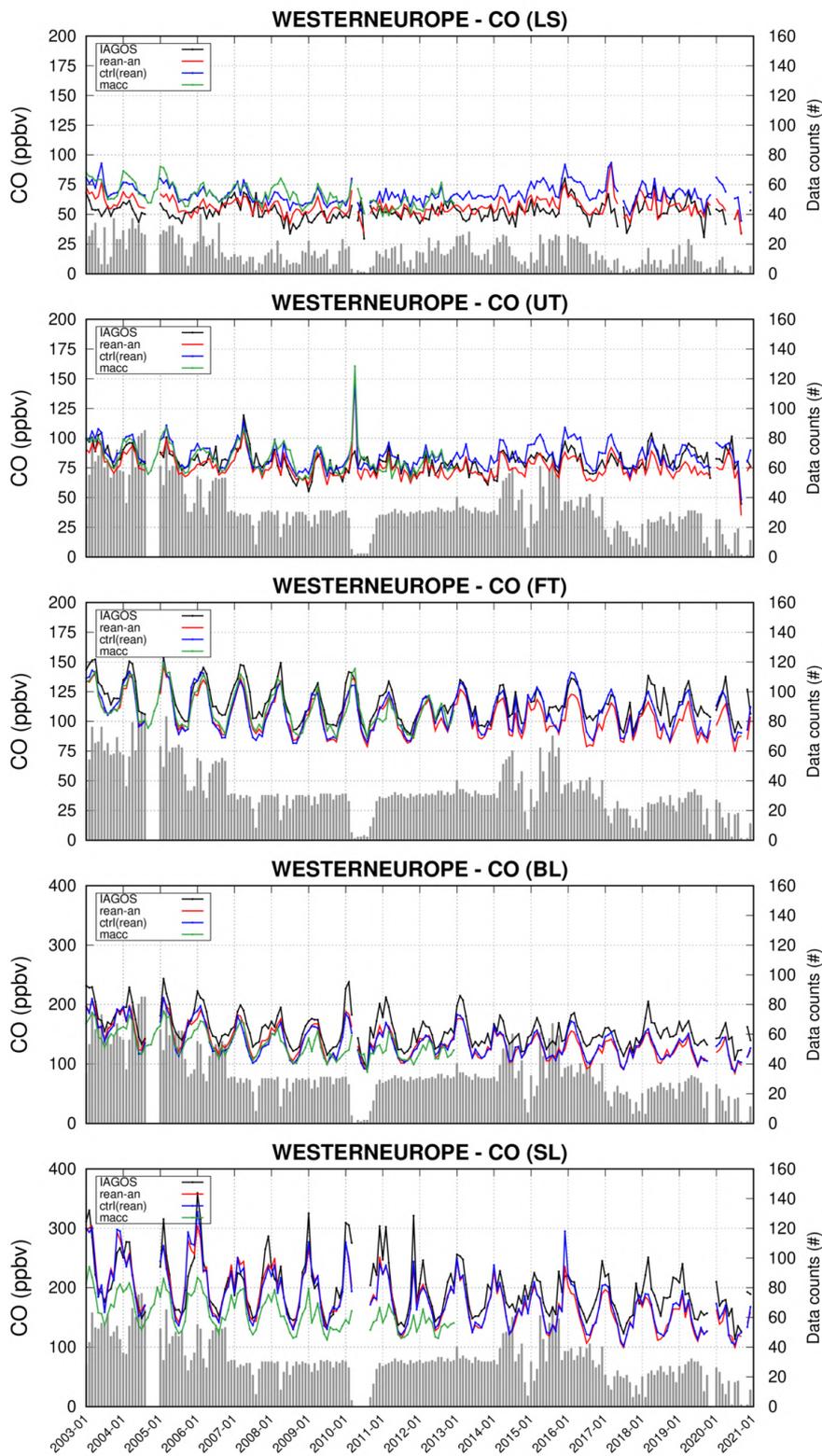


Figure 5.2.3: Monthly time series for ozone in different atmospheric layers over Western Europe during the period 2003-2020 (black: IAGOS, red: CAMS reanalysis, blue: Control run green: MACC reanalysis). From top to bottom for SL: Surface Layer, BL: Boundary Layer, FT: Free troposphere, UT: Upper troposphere, LS: Lower Stratosphere. Units: ppbv.

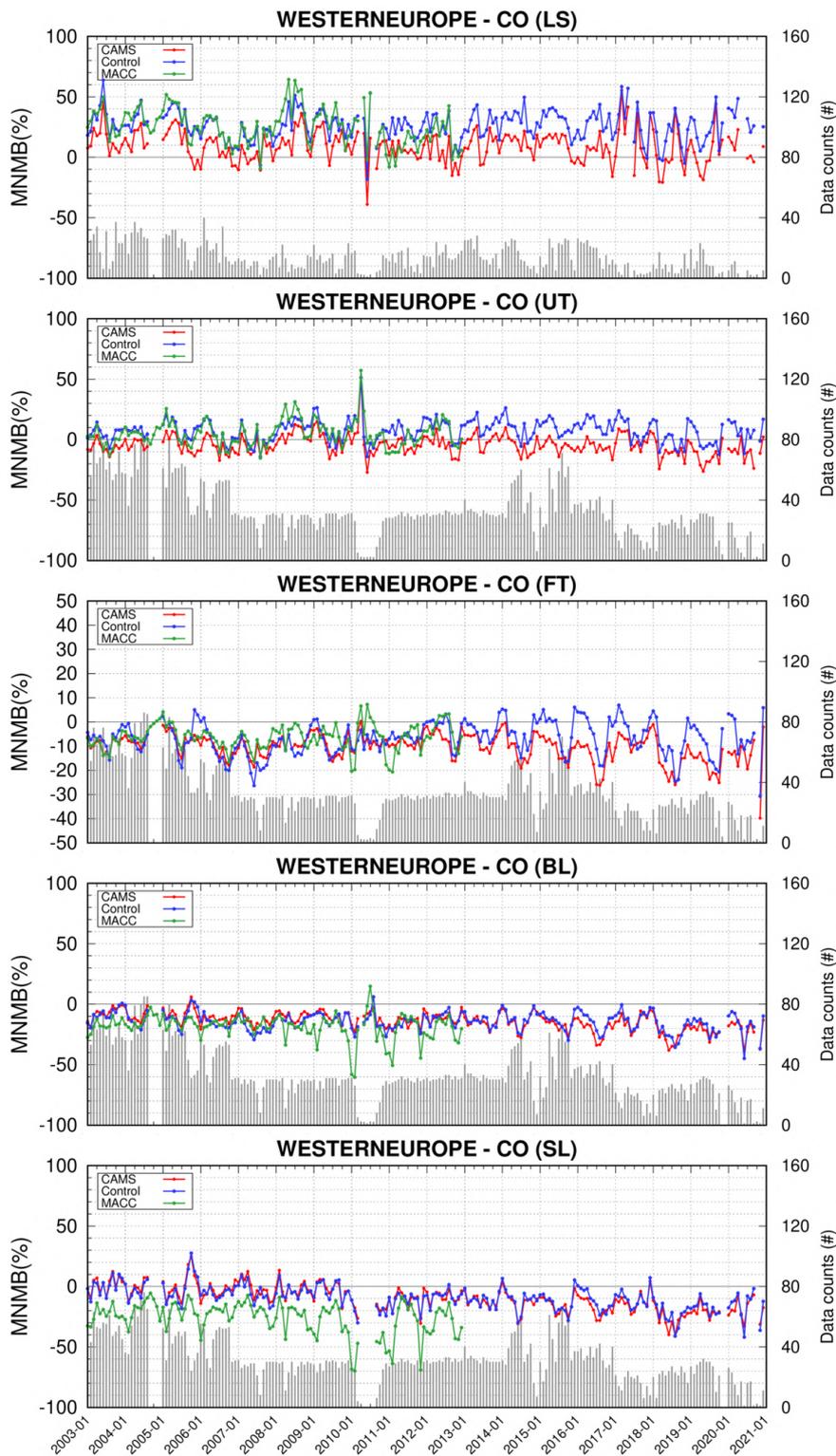


Figure 5.2.4: Time series of the monthly scores (MNMB) for ozone in different atmospheric layers over Western Europe during the period 2003-2020 (black: IAGOS, red: CAMS reanalysis, blue: Control run, green: MACC reanalysis). From top to bottom for SL: Surface Layer, BL: Boundary Layer, FT: Free troposphere, UT: Upper troposphere, LS: Lower Stratosphere. Units: %.

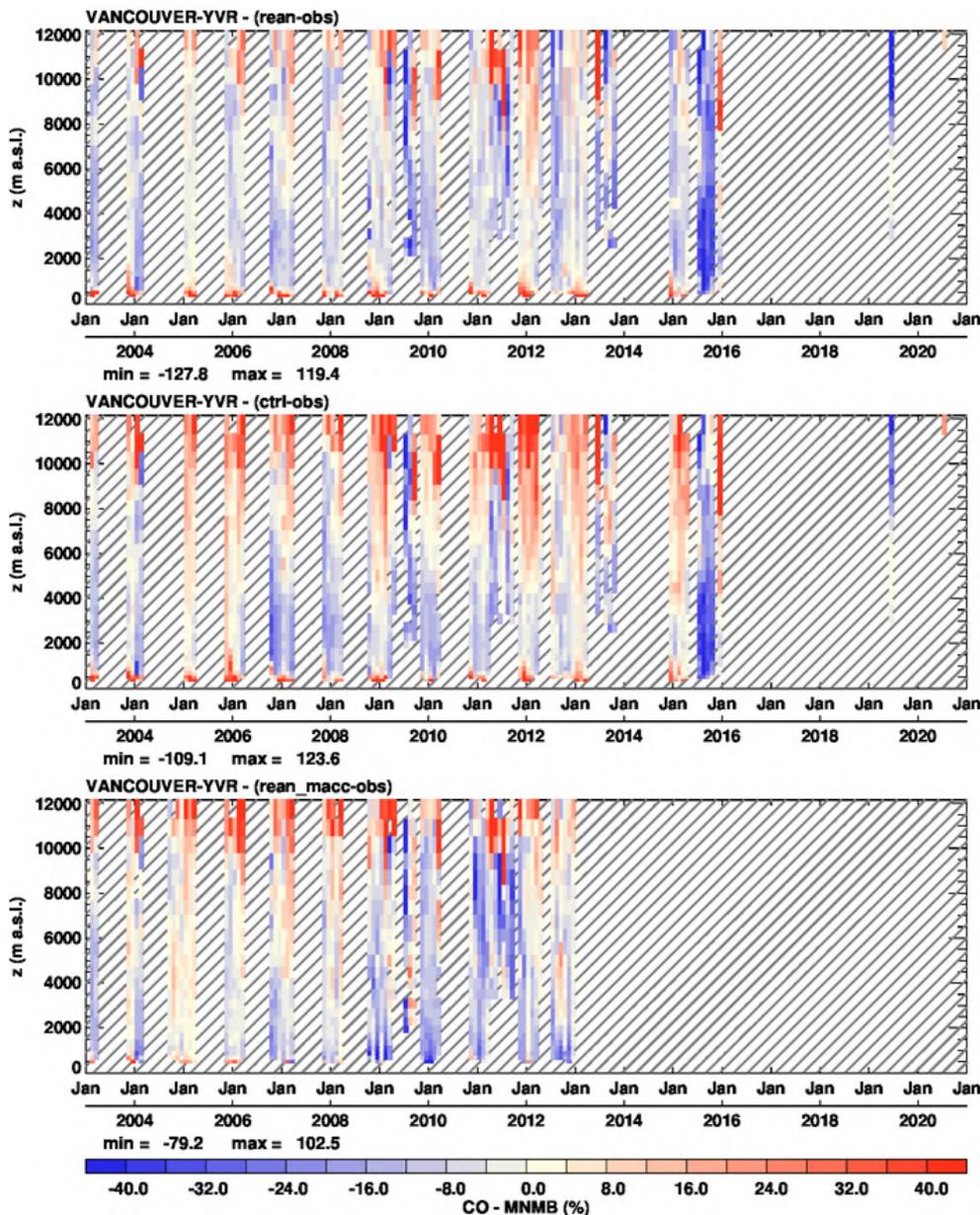


Figure 5.2.5: Time series of the monthly scores against IAGOS (MNMB) for CO at Vancouver during the period 2003-2020. From top to bottom panel, the scores are for: CAMS reanalysis, control run and MACC reanalysis. Units: %.

feature persists. Moreover, it should be noted that this increase is not present in control run comparisons. Therefore, this is likely related to the model, and maybe to assimilated MOPITT data. Like in the free troposphere, in the upper troposphere the bias is small, and in the LS the bias becomes mostly positive with large overestimations (40-50%) in the most recent years (Fig. 5.2.4).

The MACC results are very similar to those of the CAMS reanalysis from the surface to the mid-troposphere, although the CAMS reanalysis performs generally better than MACC in the lowest layers as shown especially after 2009 (Fig. 5.2.2 and 5.2.4). Indeed, the MACC reanalysis shows underestimations in the lowest layers which become largest in winter, starting from the end of 2009, with MNMB at Frankfurt reaching (in absolute value) more than 50% as compared to about



20-30% in the previous winters. In the UTLS, the MACC reanalysis performs worse than the CAMS reanalysis (Fig. 5.2.2 and Fig. 5.2.4), and CO is mostly overestimated with MNMB reaching more than 40% at Frankfurt.

The results for the other European airports are very similar to those of Frankfurt. It can be noted that CO underestimations are generally larger in the big cities like Frankfurt, and in particular in London where the MNMB often reaches more than 50% in the lower troposphere, while smaller biases are sometimes found in cities like Munich and Vienna, with even small overestimations.

Over Europe, the results from the control run differ mainly from those of the CAMS reanalysis in the upper layers, where the positive bias is much larger than for the CAMS reanalysis, as shown clearly in the time series at Frankfurt (Fig. 5.2.2 and 5.2.4). Moreover, in the free troposphere the control run presents a seasonal feature in the bias at all airports which appears here clearly starting from 2014 in Fig. 4.4.4.

### North America

As shown in Fig. 5.2.5-7, results at the North American airports are very similar to those over Europe (see previous comments). However, at some airports the CAMS reanalysis presents large overestimations near the surface, in particular at the western airports of Vancouver (Fig. 5.2.5) and Los Angeles (not shown) with large MNMB values often reaching 40%. These overestimations of CO near the surface are also present in the control run but not in the MACC reanalysis (Fig. 5.2.5).

### India

In addition to the airport of Hyderabad (see ozone section), Madras is also considered here because long CO time series are available at this airport too. The seasonality driven by the monsoon is well reproduced by both models as shown at Madras (Fig. 5.2.8).

For the CAMS reanalysis, in the lowest layers the bias presents a clear seasonality behaviour. CO is largely overestimated in winter monsoon (MNMB reaching more than 40%) and slightly underestimated in summer monsoon at both Hyderabad (not shown) and Madras (Fig. 5.2.9), with constant overestimations near the surface, in particular for Madras (Fig. 5.2.9). In the mid-troposphere, CO is mostly underestimated by the CAMS reanalysis, with a much larger bias than observed over Europe and North America in the mid-troposphere (Fig. 5.2.9) reaching up to 40%.

The MACC results are available only at Hyderabad (not shown), and present in general a lower bias than the CAMS reanalysis, especially the mid-troposphere, with an absolute value of MNMB often below 24% (not shown).

Unlike the CAMS reanalysis, the control run mostly overestimates in the mid-troposphere and UTLS as shown at Madras (Fig. 5.2.9). Moreover, this bias shows a seasonal pattern which is more clearly

seen at Hyderabad (not shown) where the performance of control run is found to be better than that of the CAMS reanalysis.

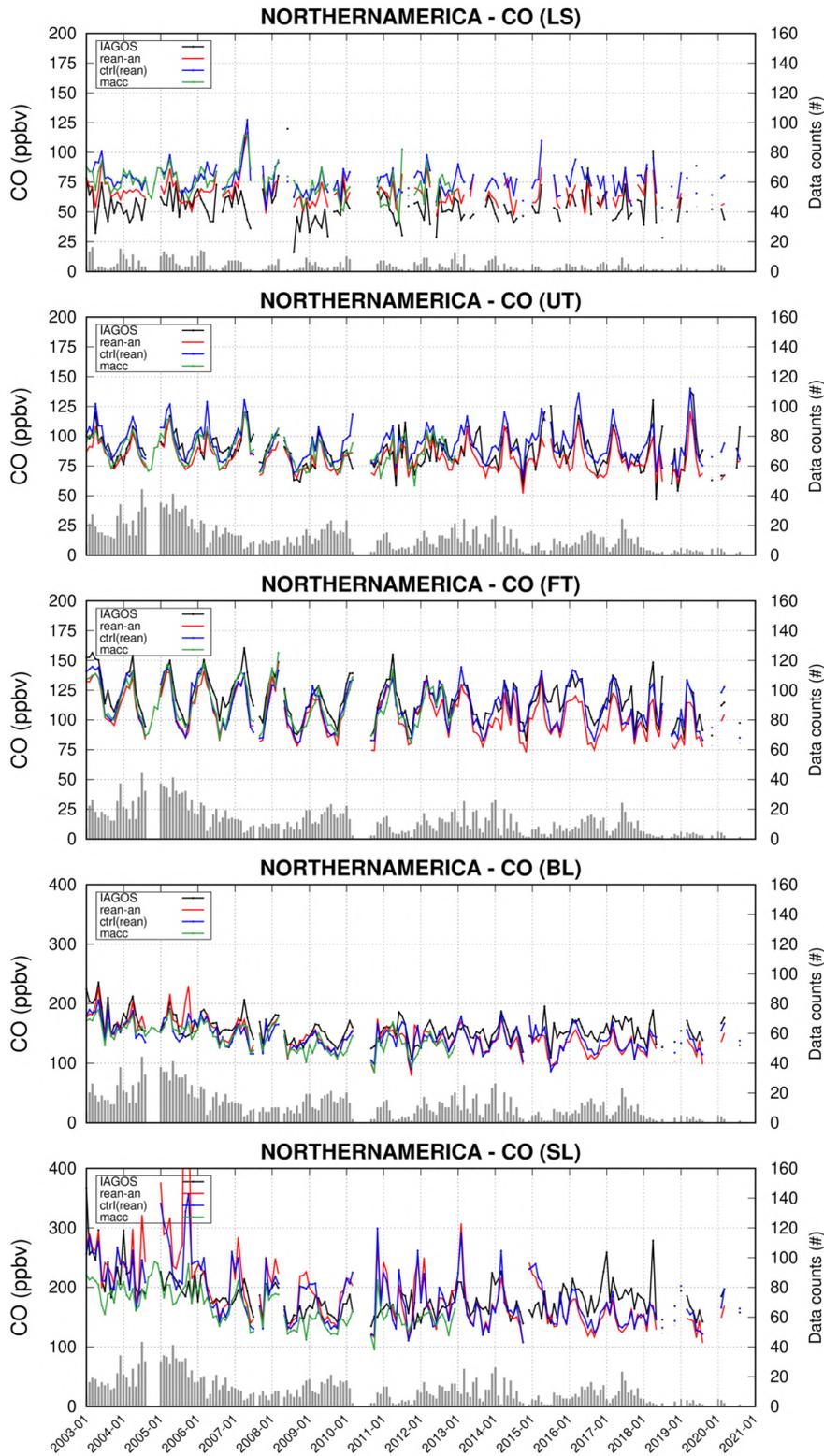


Figure 5.2.6: Monthly time series for ozone in different atmospheric layers over Northern America during the period 2003-2020 (black: IAGOS, red: CAMS reanalysis, blue: Control run green: MACC reanalysis). From top to bottom for SL: Surface Layer, BL: Boundary Layer, FT: Free troposphere, UT: Upper troposphere, LS: Lower Stratosphere. Units: ppbv.

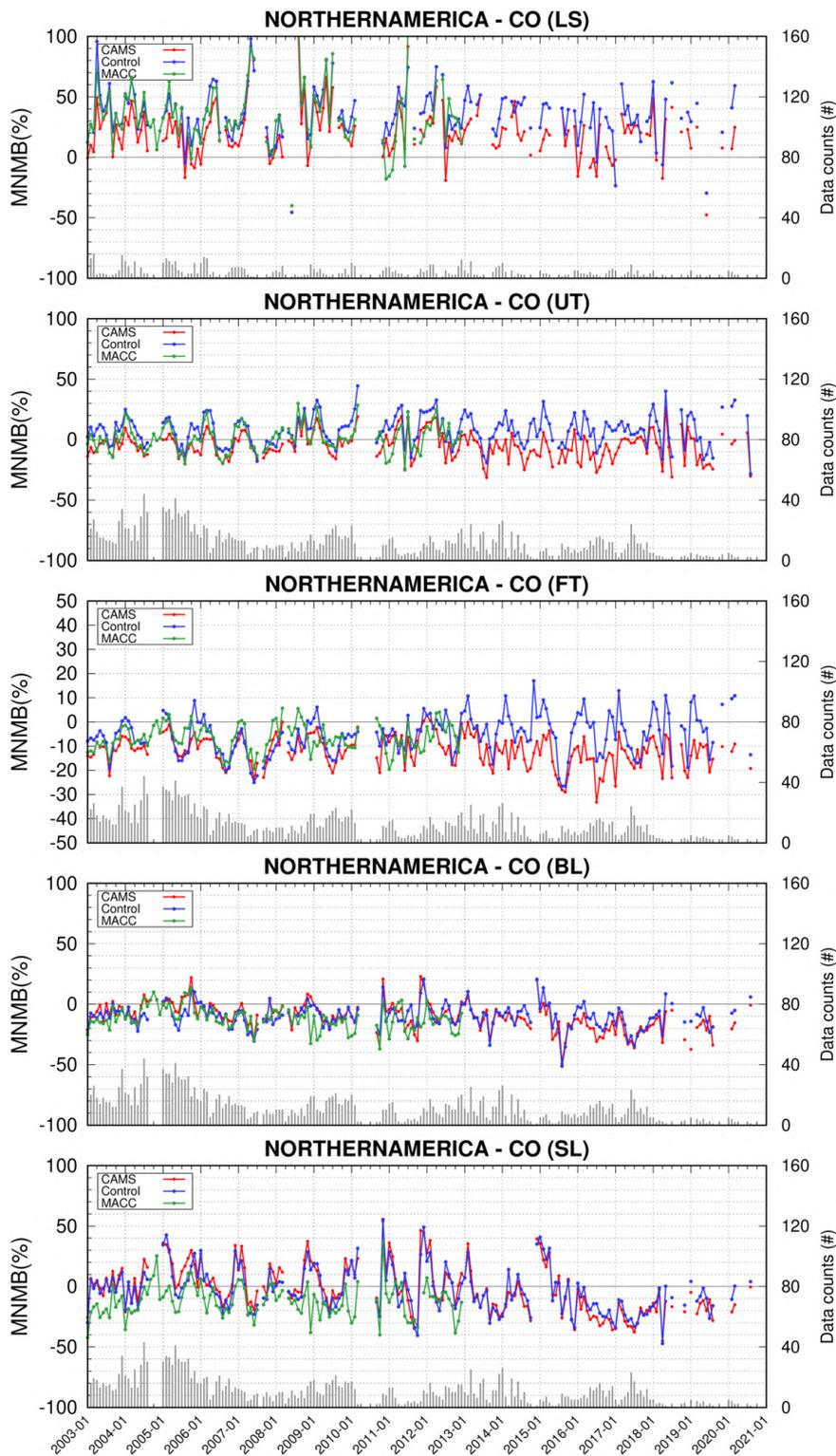


Figure 5.2.7: Time series of the monthly scores (MNMB) for ozone in different atmospheric layers over Northern America during the period 2003-2020 (black: IAGOS, red: CAMS reanalysis, blue: Control run, green: MACC reanalysis). From top to bottom for SL: Surface Layer, BL: Boundary Layer, FT: Free troposphere, UT: Upper troposphere, LS: Lower Stratosphere. Units: %.

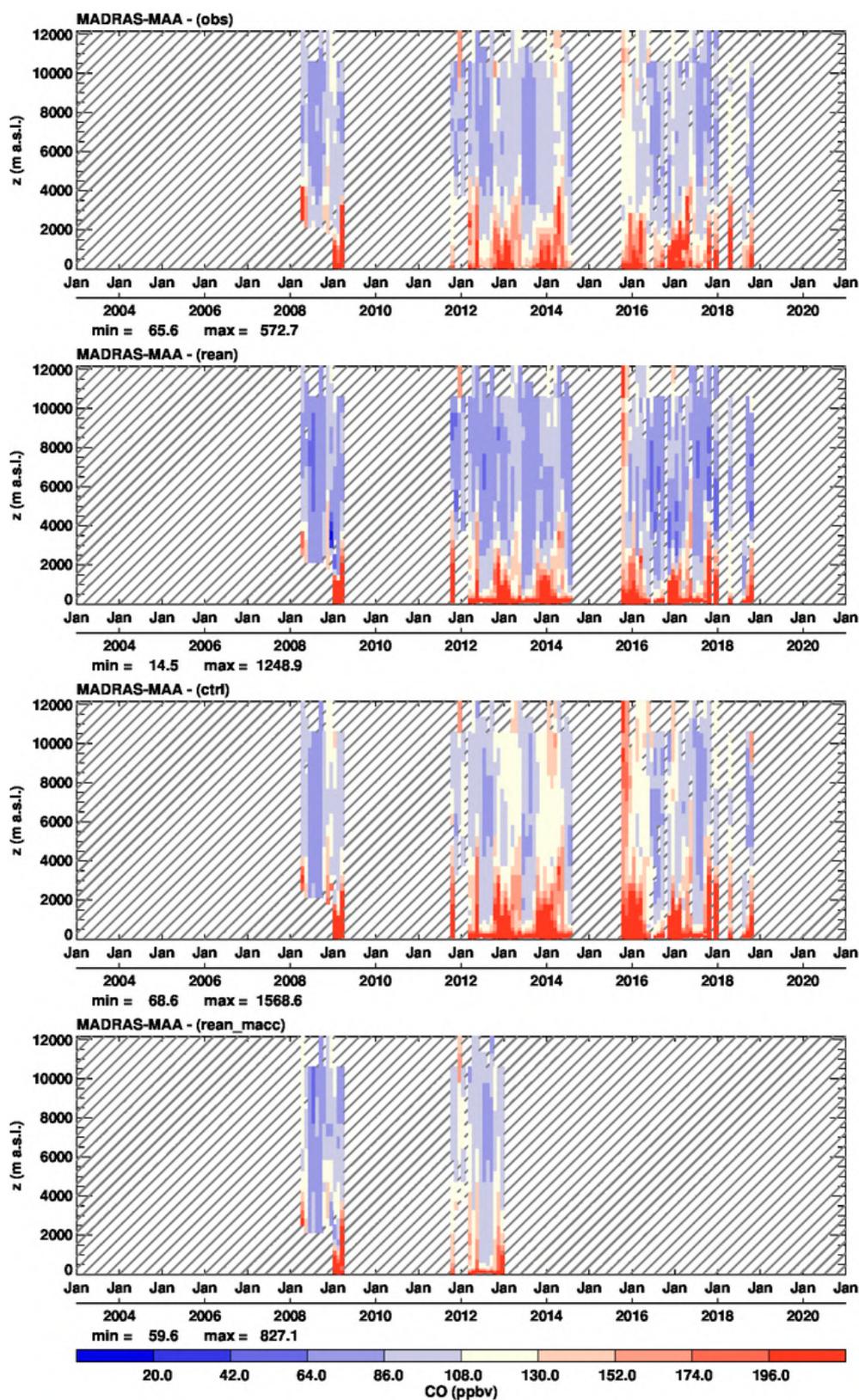


Figure 5.2.8: Time series of the monthly profiles for CO at Madras during the period 2003-2020. From top to bottom panel the time series correspond to: IAGOS (MOZAIC), CAMS reanalysis, control run and MACC reanalysis. Units: ppbv.

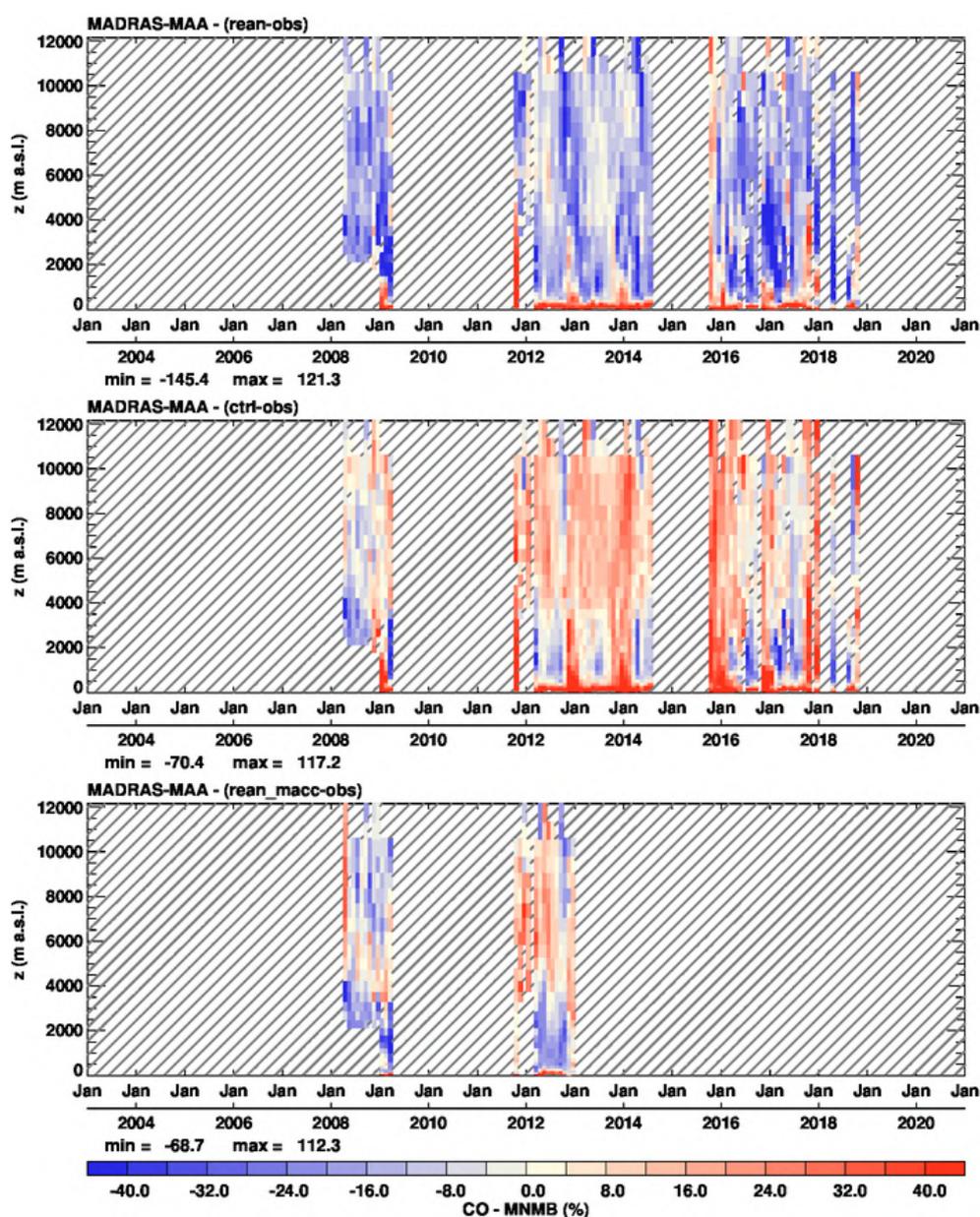


Figure 5.2.9: Time series of the monthly scores against IAGOS (MNMB) for CO at Madras during the period 2003-2020. From top to bottom panel, the scores are for: CAMS reanalysis, control run and MACC reanalysis. Units: %.

### Eastern Asia

Over Asia, CO is mostly underestimated by the CAMS reanalysis in the lowest layers, with generally much larger biases than over Europe (Fig. 5.2.10-13). At Taipei, these underestimations are generally between 16 to 32% (Fig. 5.2.10). This bias is larger at the other Chinese airports where CO values are much higher in the lower troposphere throughout the year. This is the case at the airports of Shenyang, Nanjing and Beijing but also at the Thai airport of Bangkok (not shown), with MNMB values often reaching 40% throughout the year. However, similarly to the Indian airport of

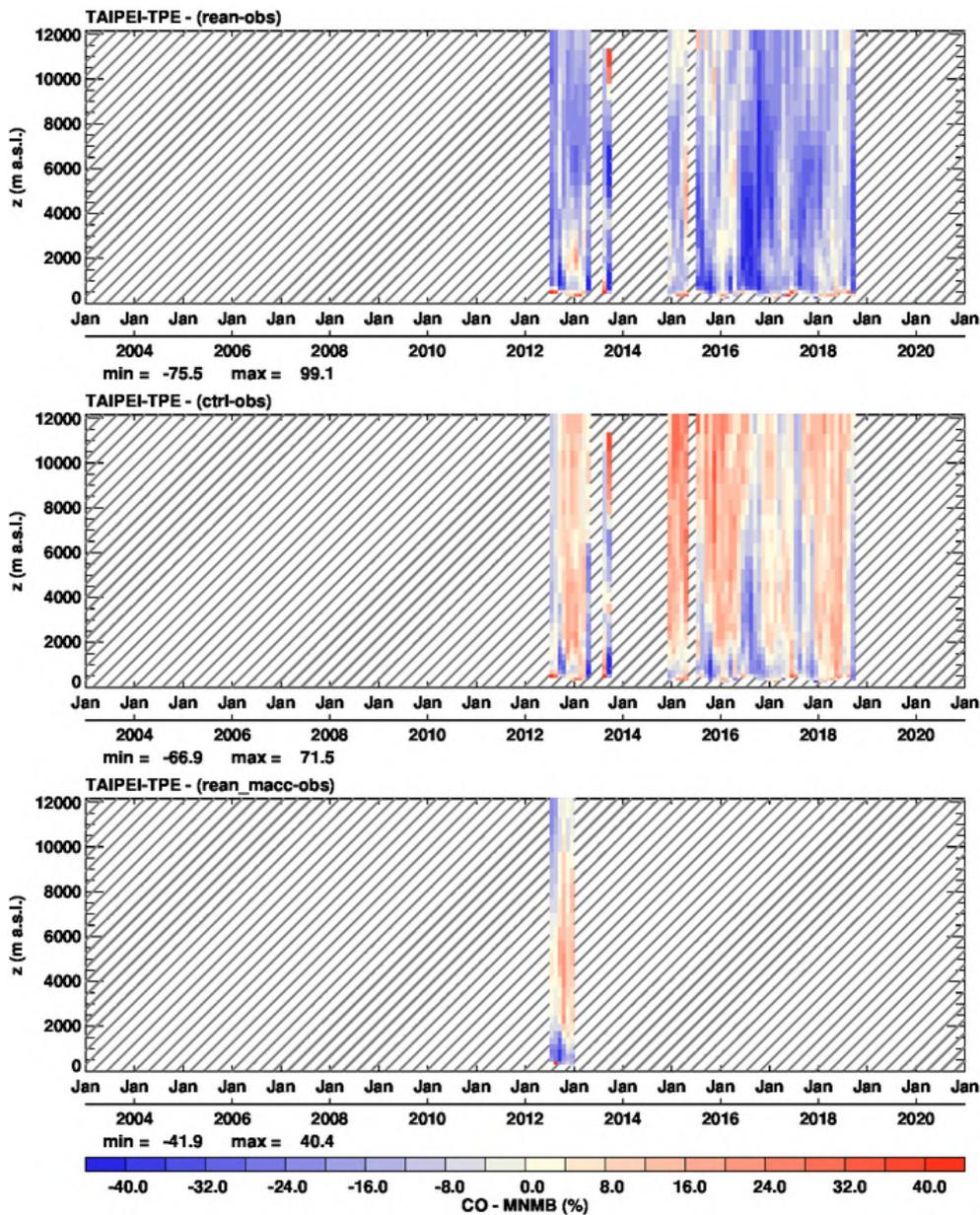


Figure 5.2.10: Time series of the monthly scores against IAGOS (MNMB) for CO at Taipei during the period 2003-2020. From top to bottom panel, the scores are for: CAMS reanalysis, control run and MACC reanalysis. Units: %.

Madras, very close to the surface the values are overestimated at some airports, as demonstrated by the observations over Hong Kong (Fig. 5.2.11) and also at Bangkok (not shown). At Taipei and the Japanese cities, these overestimations occur rarely and are much smaller (not shown).

In the mid-troposphere and UTLS, the results of the CAMS reanalysis over Asian airports (Fig. 5.2.10, 5.2.11 and 5.2.13) are similar to those obtained for India (see previous comments). The model shows an overall underestimation of CO although a seasonality of the bias is present in the mid-troposphere at some airports with usually better results during winter months as seen at Tokyo (not shown).

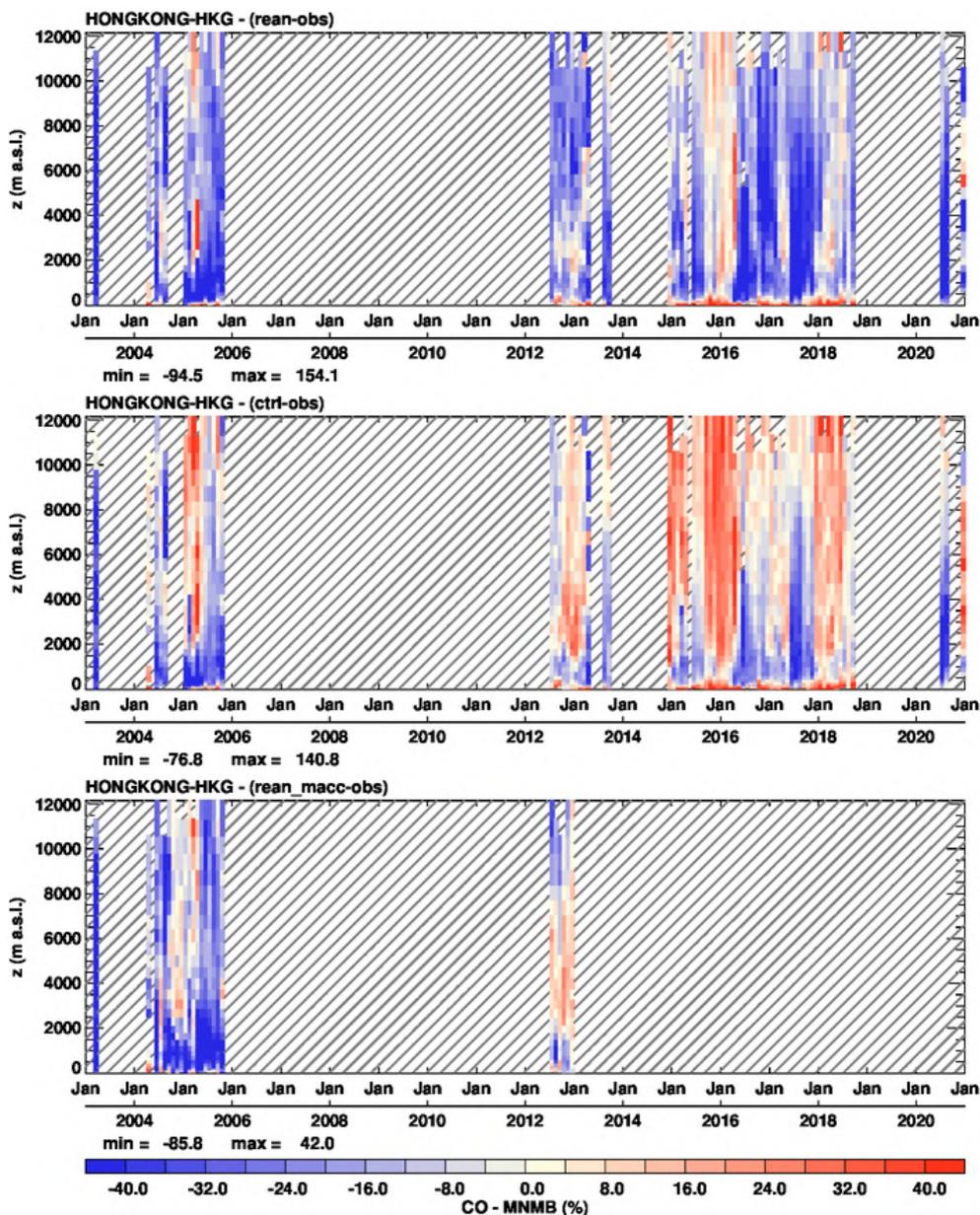


Figure 5.2.11: Time series of the monthly scores against IAGOS (MNMB) for CO at Hong Kong during the period 2003-2020. From top to bottom panel, the scores are for: CAMS reanalysis, control run and MACC reanalysis. Units: %.

The MACC results are available only at Tokyo (Fig. 5.2.10) and Osaka (not shown), and no notable difference is found with the results of the CAMS reanalysis (see also Fig. 5.2.13).

As found over India, the differences between the CAMS reanalysis and associated control run are found mainly in the mid-troposphere and UTLS, where in contrast to the CAMS reanalysis the control run mostly shows overestimations (Fig. 5.2.10, 5.2.11 and 5.2.13). Moreover, the control run shows a seasonal behaviour of the bias in the free troposphere which is not always present or less pronounced in the CAMS reanalysis (Fig. 5.2.10, 5.2.11 and 5.2.13). This behaviour of control run can be clearly seen for example at Taipei and Hong Kong (Fig. 5.2.10 and 5.2.11).

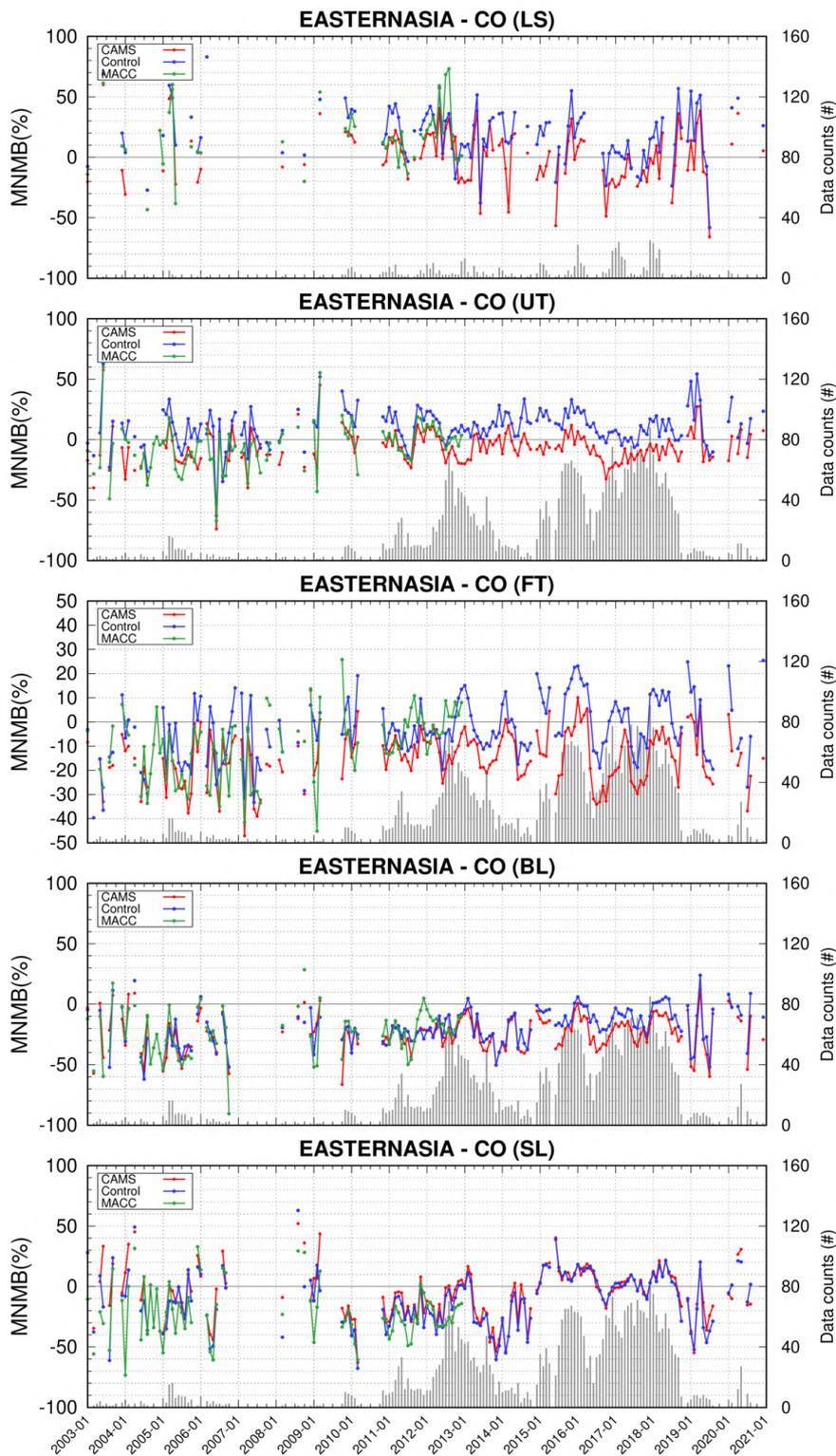


Figure 5.2.12: Monthly time series for ozone in different atmospheric layers over Eastern Asia during the period 2003-2020 (black: IAGOS, red: CAMS reanalysis, blue: Control run green: MACC reanalysis). From top to bottom for SL: Surface Layer, BL: Boundary Layer, FT: Free troposphere, UT: Upper troposphere, LS: Lower Stratosphere. Units: ppbv.

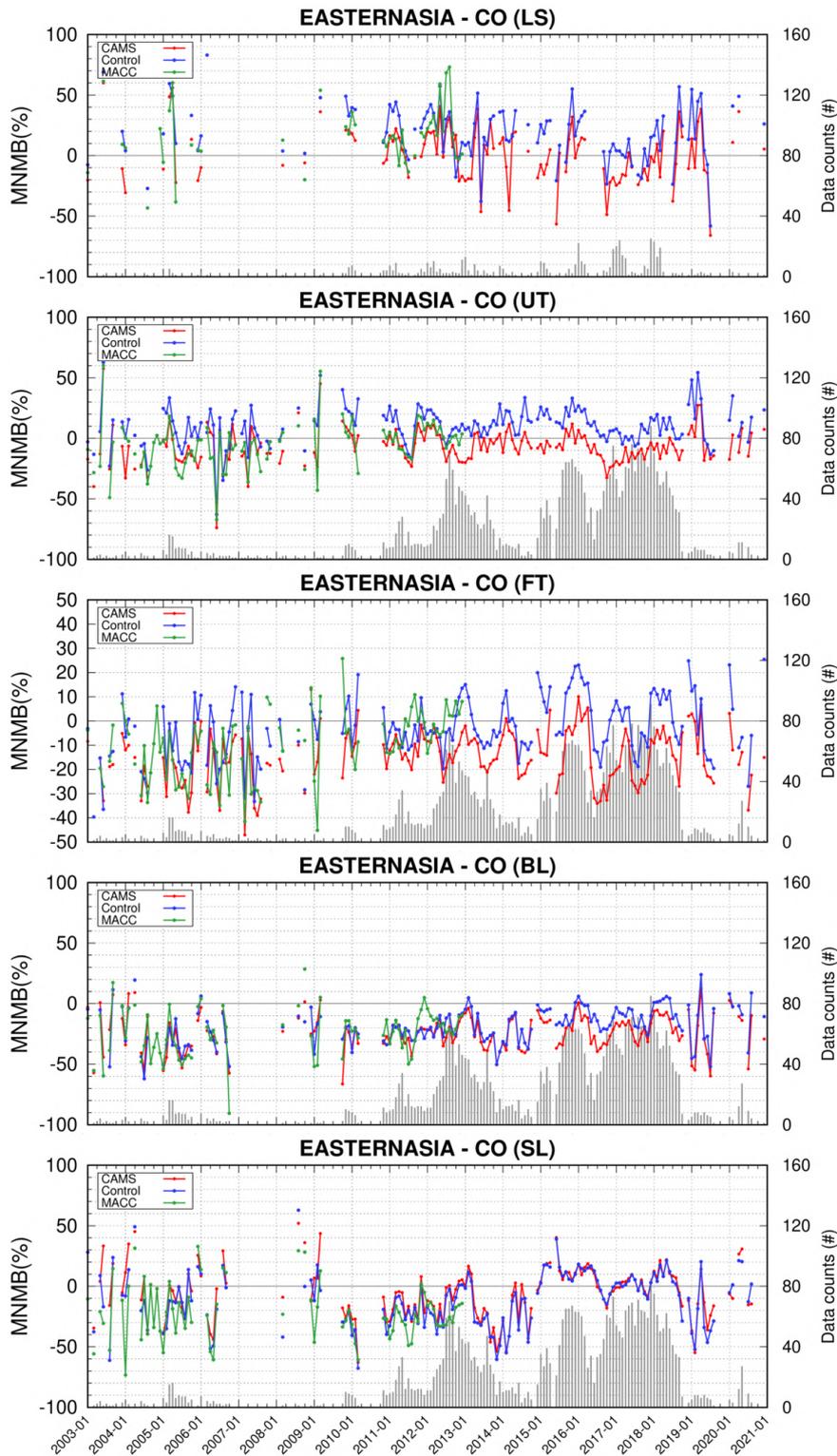


Figure 5.2.13: Time series of the monthly scores (MNMB) for ozone in different atmospheric layers over Eastern Asia during the period 2003-2020 (black: IAGOS, red: CAMS reanalysis, blue: Control run, green: MACC reanalysis). From top to bottom for SL: Surface Layer, BL: Boundary Layer, FT: Free troposphere, UT: Upper troposphere, LS: Lower Stratosphere. Units: %.

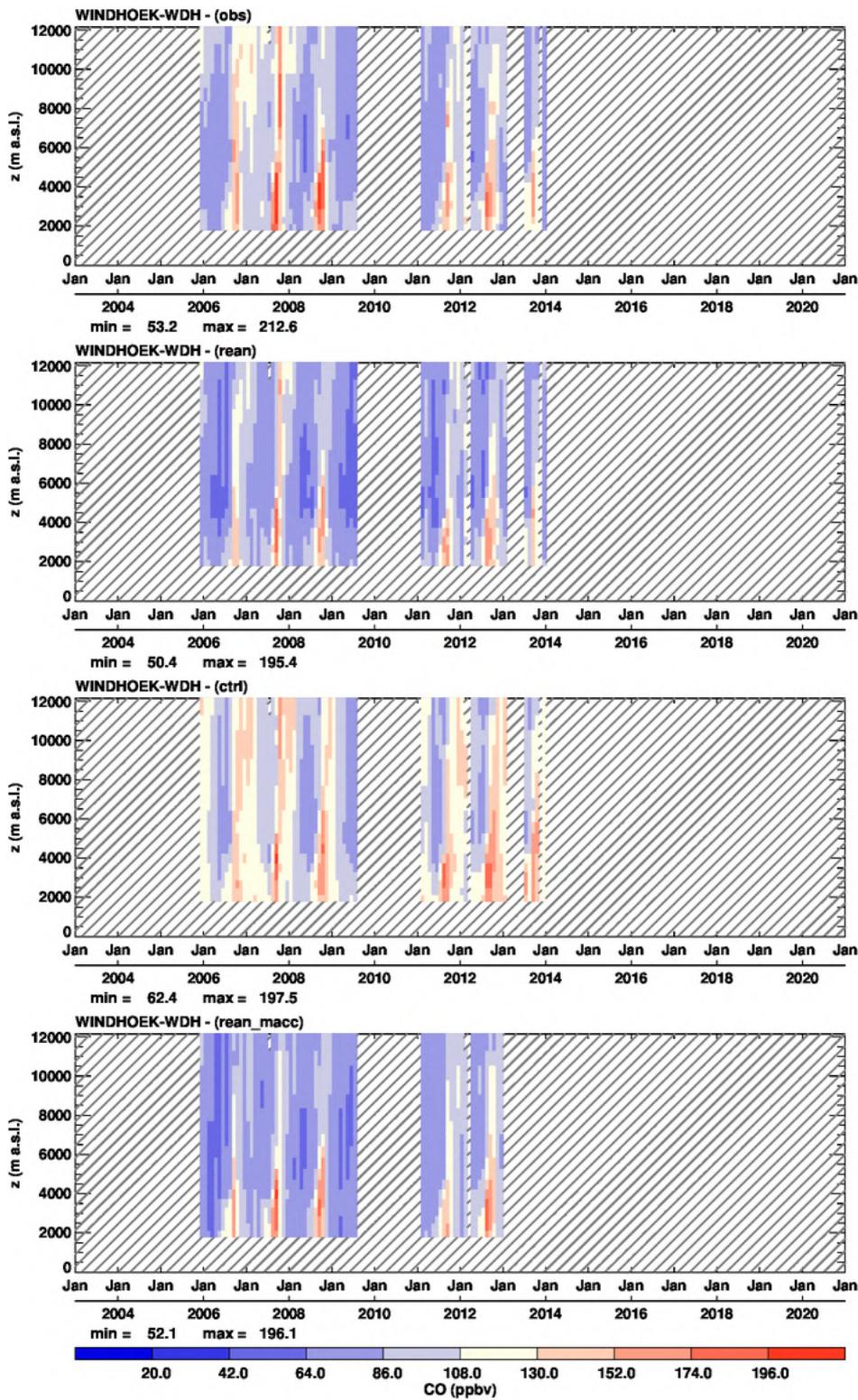


Figure 5.2.14: Time series of the monthly profiles for CO at Windhoek during the period 2003-2020. From top to bottom panel the time series correspond to: IAGOS (MOZAIC), CAMS reanalysis, control run and MACC reanalysis. Units: ppbv.

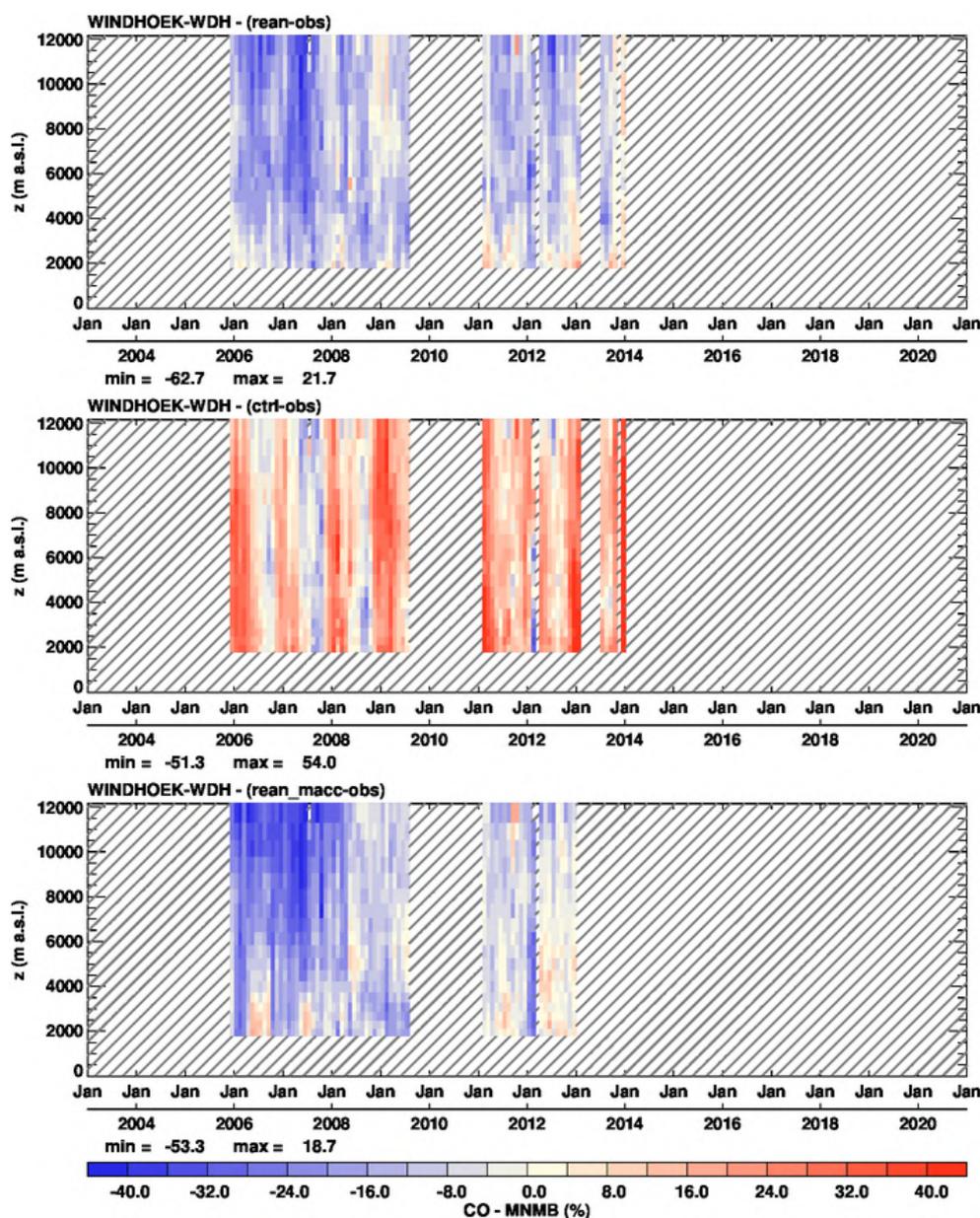


Figure 5.2.15: Time series of the monthly scores against IAGOS (MNMB) for CO at Windhoek during the period 2003-2020. From top to bottom panel, the scores are for: CAMS reanalysis, control run and MACC reanalysis. Units: %.

### Southern and Western Africa

The seasonality of CO at Windhoek (1700 m.a.s.l) is characterized by low CO values (< 90 ppbv) over the whole atmospheric column during the wet season (November-April), and high CO values (> 100 ppbv) with a maximum in the lowest layers during the dry season (May-October), as shown in Fig. 5.2.14. The opposite phase applies over the coastal area of Lagos with the 6 months shift of season between the northern and southern hemisphere with very large values in the lowest layers (not shown).



The seasonality is well reproduced by both reanalyses at these airports. As shown on Fig. 5.2.15, the behaviour of the bias is in general similar for the two reanalyses. Both CAMS and MACC mostly underestimates CO in all layers and for all seasons, with the exception of a small overestimation in the free troposphere during the dry season for MACC.

The largest underestimations (MNMB reaching more than 40%) are found in the mid-troposphere and UTLS at Windhoek (Fig. 5.2.15) and are much larger near the surface at the coastal airport of Lagos (not shown). These underestimations at Windhoek are particularly large before the end of 2008 with a worse performance for the MACC reanalysis. From 2009 onwards, both CAMS and MACC results at Windhoek seem to improve in the mid troposphere, especially during the dry season when fires are common. In this second period, the performance of MACC is also improving in the UTLS, while the performance remains stable for CAMS in the upper layers.

In the mid-troposphere to the UTLS, the control run behaves differently than the CAMS reanalysis at both Northern and Southern Hemisphere African airports, where the control run mostly shows overestimations. Moreover, at Windhoek the bias shows a seasonal feature in the mid-troposphere and UTLS with large overestimates in the wet season often beyond 32%, and a better agreement (small underestimation) in the dry season.

#### Comparison with cruise level data

Fig. 5.2.16 shows the gridded average maps for the comparisons with IAGOS observations of ozone at flight level. The spatial variability of CO is well represented by both CAMS configurations. Over the northern mid-latitudes, CO is mostly underestimated with biases below 30% in absolute value. The larger negative biases can be found over the boreal regions of North America and in particular over the boreal region of Siberia with absolute values beyond 50%. Over Eastern Asia the bias is rather small with a positive or negative sign, while over western Asia as well as over India positive values roughly between 10 to 20% are found. In Western Africa (Gulf of Guinea) CO is underestimated, but CO is overestimated in the most southern parts of Africa providing observations.

Control run results differ strongly from those of the CAMS reanalysis. CO is mostly overestimated over all regions of the world with the exception of the boreal region of Siberia which clearly present negative bias as well as some parts of the boreal region of North America although less clearly. The largest positive biases are found in southern Africa with values greater than 50%. Large biases are also found over India, the Middle East and the Caribbean region.

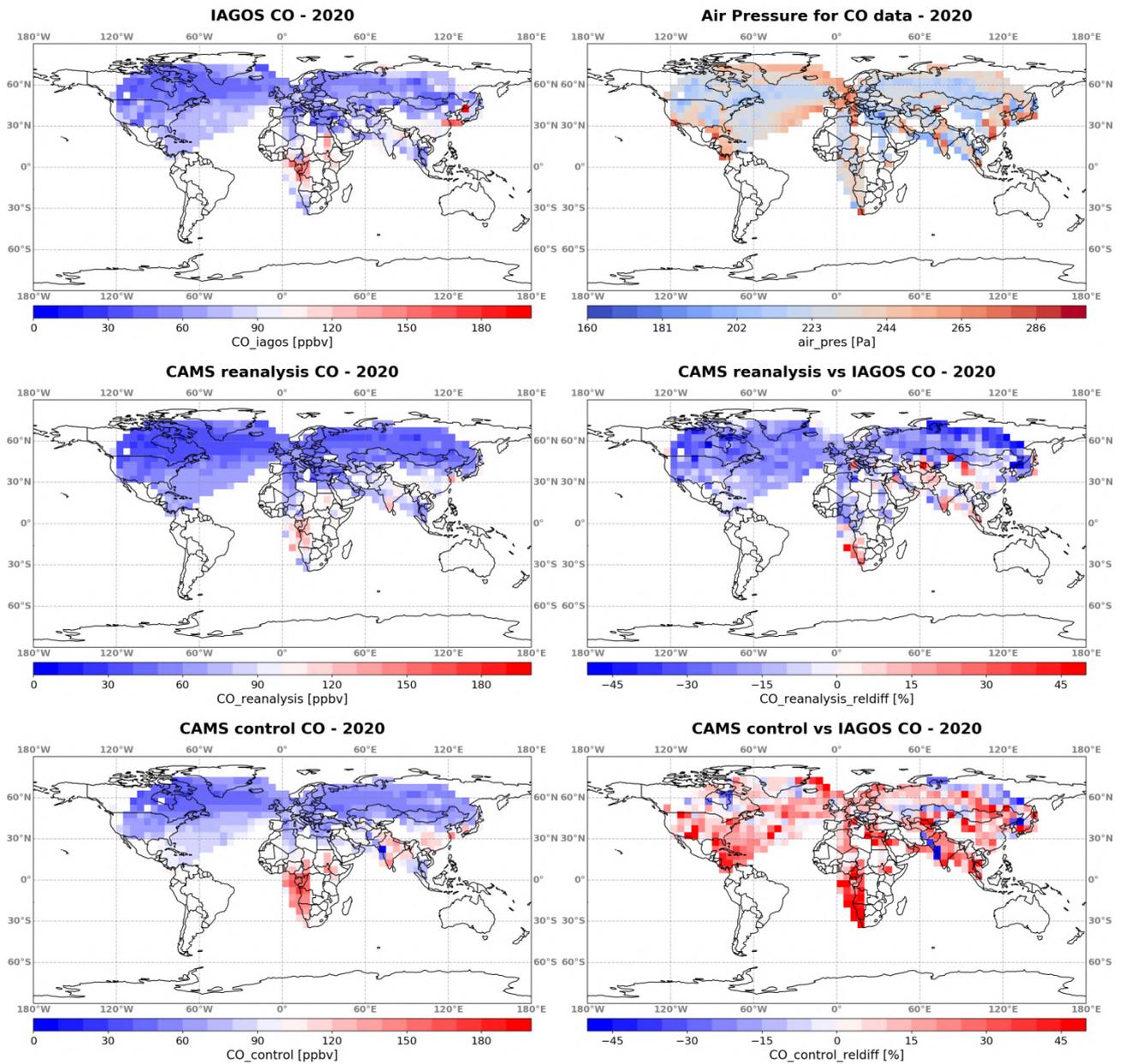


Figure 5.2.16: Global maps of gridded averages (5°x5°) for CO comparison with IAGOS cruise data for the year 2020. From left to right, top panel: IAGOS CO (in ppbv) and air pressure (in hPa). Bottom panel: CO from the CAMS reanalysis (in ppbv) and associated relative differences (in %) with respect to IAGOS.

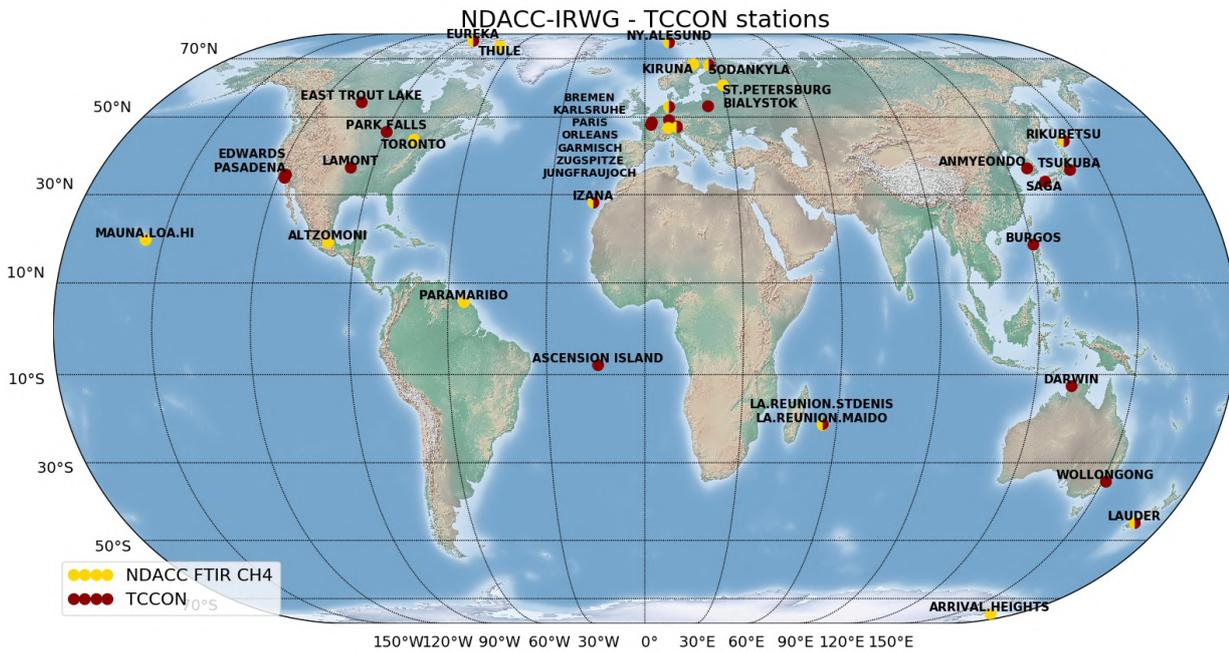


Figure 5.3.1: Overview of the geographical distribution of NDACC and TCCON FTIR instruments.

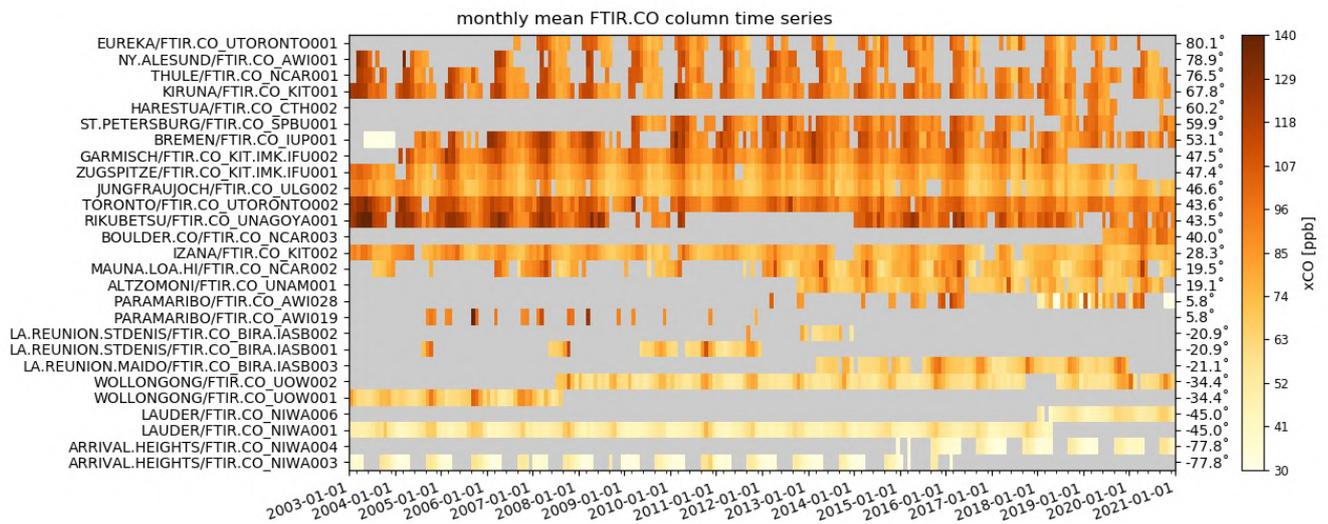


Figure 5.3.2: Overview of CO data product time series available on NDACC. Most FTIR data products are used in this study. Sites are sorted by latitude. The plot shows dry air column averaged CO.

### 5.3 Validation against FTIR CO observations from the NDACC network

In this section, we compare the CO profiles of the CAMS reanalysis model with FTIR measurements at selected NDACC sites (Fig. 5.3.1). 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. The FTIR partial columns (surface to 20km) are compared to the corresponding model’s partial column. In the co-location criteria, the line of sight of the FTIR measurement is taken into account and the FTIR averaging kernel is applied to the co-located model data. A more detailed

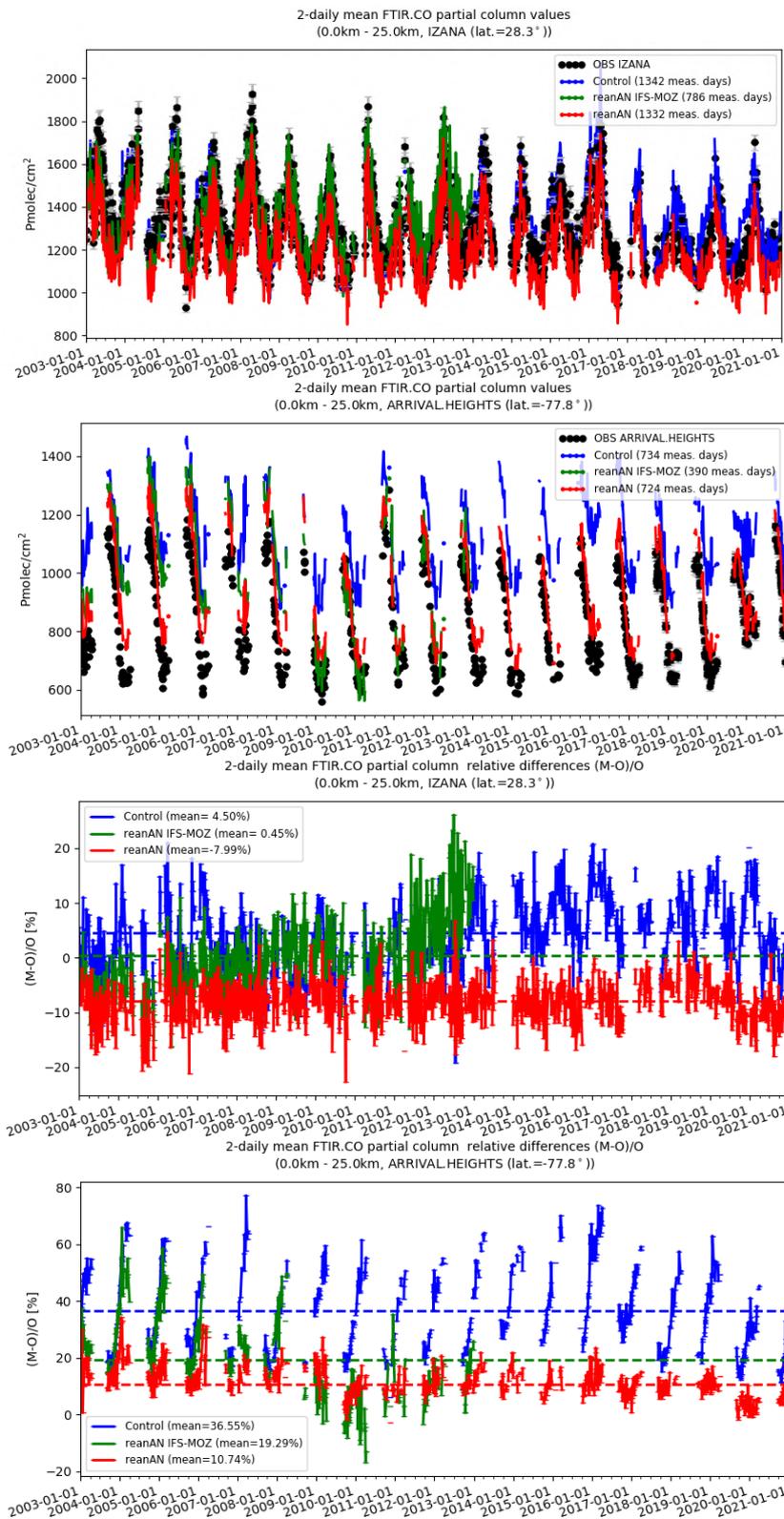


Figure 5.3.3: Upper two panels show the time series plots at the two NDACC stations Izana (tropical station) and Arrival Heights (Antarctic station). The lower two panels show the daily mean differences. For the CAMS reanalysis the time series do not show strong differences within the entire time period (2013-2020), which was the case for the previous the MACC reanalysis (e.g., the increased bias in 2011-2012 at Izana).



description of the instruments and applied methodologies can be found at <http://nors.aeronomie.be>. The typical total uncertainty of the FTIR CO column is approximately 3%.

Fig. 5.3.2 shows the CO data time series available on NDACC. Most stations are used in this study except Paramaribo (too sparsely represented), Wollongong (not conform NDACC template) and Bremen (not conform NDACC template for 2003-3005). It may happen that within the validation time period an instrument was replaced (St Denis and Arrival Heights). For these sites the data products of both instruments are concatenated to one single time series in the validation results below.

The MACC reanalysis (rean IFS-MOZART) is indicated in green and the new CAMS reanalysis in red (reanAN). From Fig. 5.3.3 we can deduce that the CAMS model captures well the day-to-day variations of the tropospheric columns of CO and seems to be more stable than the previous reanalysis: for the CAMS reanalysis the time series do not show strong differences from one year to the other, which was the case for MACC (for example the increased bias in 2011-2012 at Izana).

The following list summarizes the conclusions drawn from the figures presented here.

- Concerning the stability of relative differences throughout the entire time series (see Fig. 5.3.3 and Fig. 5.3.5) the CAMS reanalysis performs better than the MACC reanalysis. The relative bias of the control run shows a strong latitudinal dependence in the southern hemisphere and a positive trend in the relative bias for the northern hemisphere (Fig. 5.3.5). A small discontinuity in the evolution of the CAMS reanalysis biases can be seen at end 2016 and mid 2019 (see Fig. 5.3.5, most clearly at the tropical sites). This coincides with changes in the assimilated MOPITT CO product version (v6 to v7 end of 2016 and v7 to v8 mid 2019).
- From the Taylor diagrams in Fig. 5.3.4, the correlation coefficients of the CAMS reanalysis (most stations are above 0.9) are higher compared to the MACC reanalysis and the standard deviation of the difference between the CAMS reanalysis and the NDACC time series is below 0.6 for all stations (compared to 0.8 for the MACC reanalysis)
- The CAMS reanalysis shows an increased bias in the tropics in the northern hemisphere, the overall bias at these sites (Izana, Mauna Loa and Alzomoni) is well above the reported measurement uncertainty. This increased bias was not observed in the MACC run. However, some care must be taken because the three sites in question are all mountain stations and co-location issues might play, and also because the bias is reduced with the assimilation of MOPITT v6 in 2017-mid 2019 in the CAMS reanalysis.
- The overall bias changes sign between the NH and the SH. Although the CAMS reanalysis performs better than MACC at the Antarctic station, the overall bias is still well above the reported measurement uncertainty (Fig. 5.3.4).
- The relative differences for the stations in the southern hemisphere show a significant positive trend in the relative biases: Maitopo +0.26% per year, Lauder +0.13% per year and Wollongong +0.95% per year. At the Antarctic station the trend in the relative bias is negative (see Fig. 5.3.6). The high latitude stations (both Arctic and Antarctic) show a seasonal dependence in the relative differences with a reduced bias during local spring/summer).

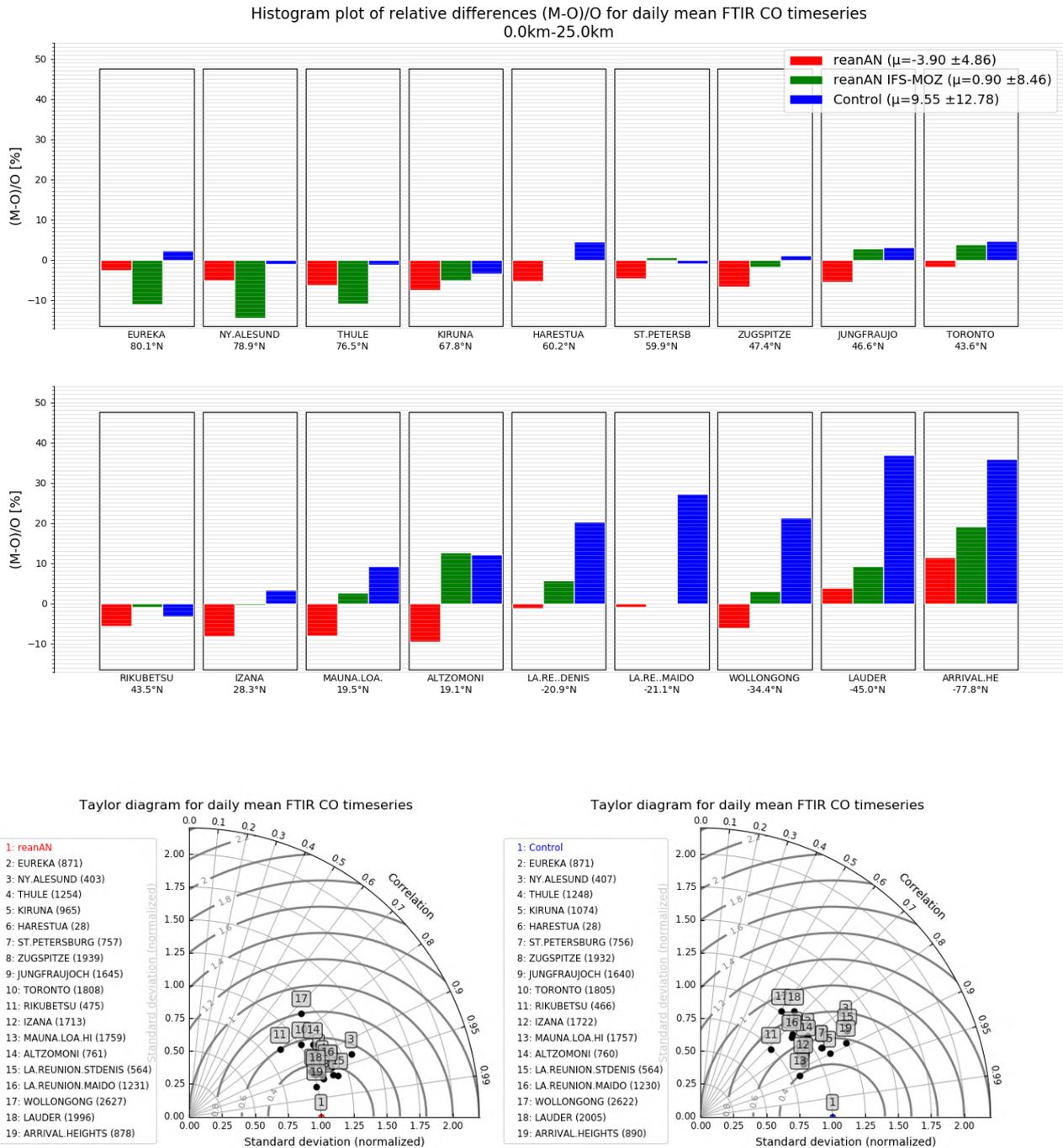


Figure 5.3.4: Averaged daily bias for each site (top, sorted by decreasing latitude) for the entire time period 2003-2020. Bottom plots: the Taylor diagrams for the period 2003-2020 comparing the CAMS reanalysis (left) and Control run (right). The CAMS reanalysis performs better at the higher latitude stations. In the northern hemisphere the CAMS reanalysis underestimates the CO abundance by 5% which increases towards the tropics (Izana) where the model underestimates the CO abundance by values up to 10%. In the southern hemisphere, the negative bias of the CAMS reanalysis at Reunion is nearly vanishing. The relative bias of the control run shows a strong latitudinal dependence in the southern hemisphere.

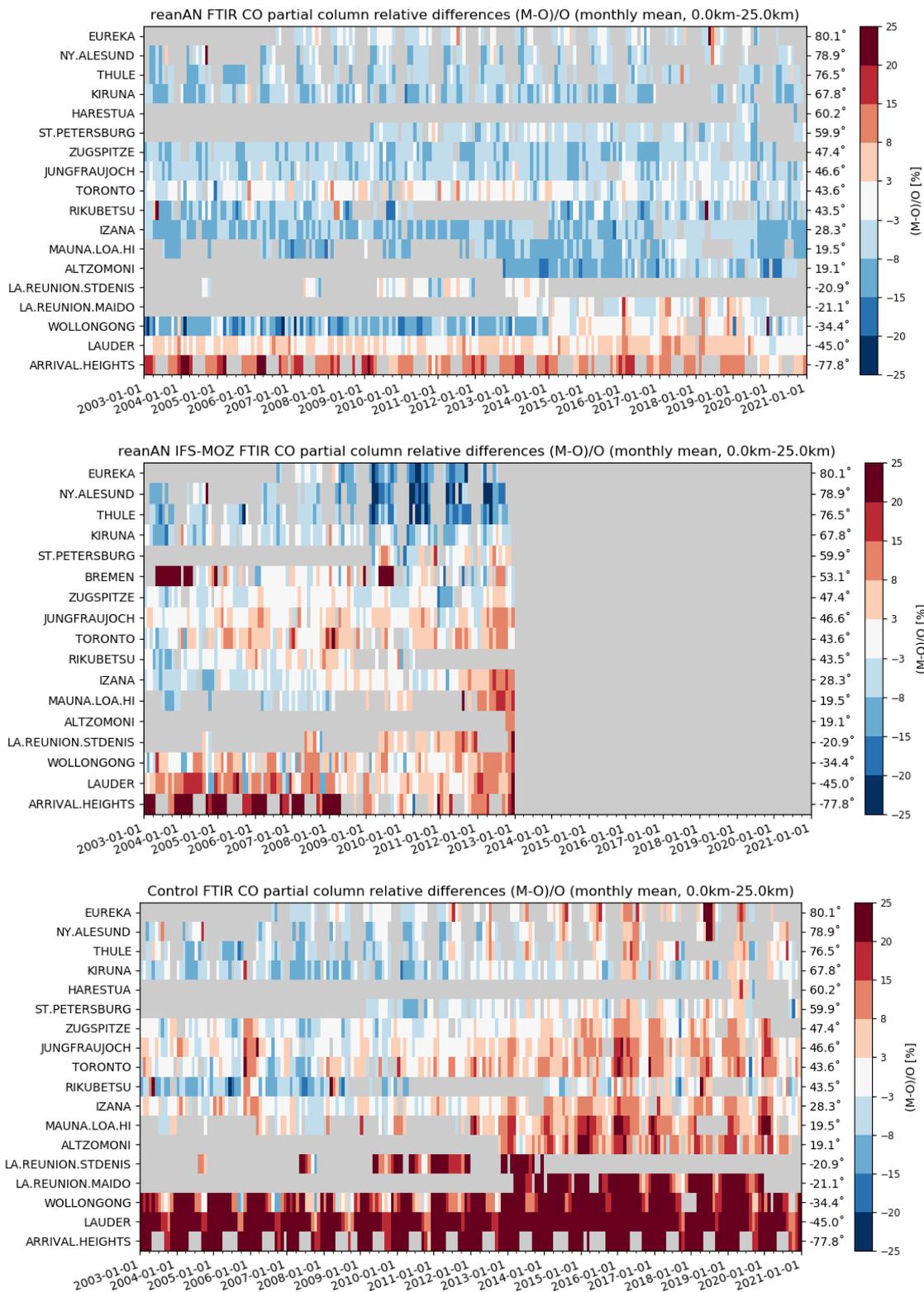


Figure 5.3.5: Mosaic plot of the CO seasonal relative biases at all NDACC-FTIR sites for the period 2003-2020. The CAMS reanalysis (top panel) performs more stable compared to the MACC reanalysis (middle panel). A negative bias is found in the Northern Hemisphere, a slightly positive bias in the Southern Hemisphere. At the tropical sites in the northern hemisphere the negative bias reach values up to 10% (Izana, Mauna Loa, Altzomoni). The control run (bottom panel) shows significant positive biases in the Southern Hemisphere.

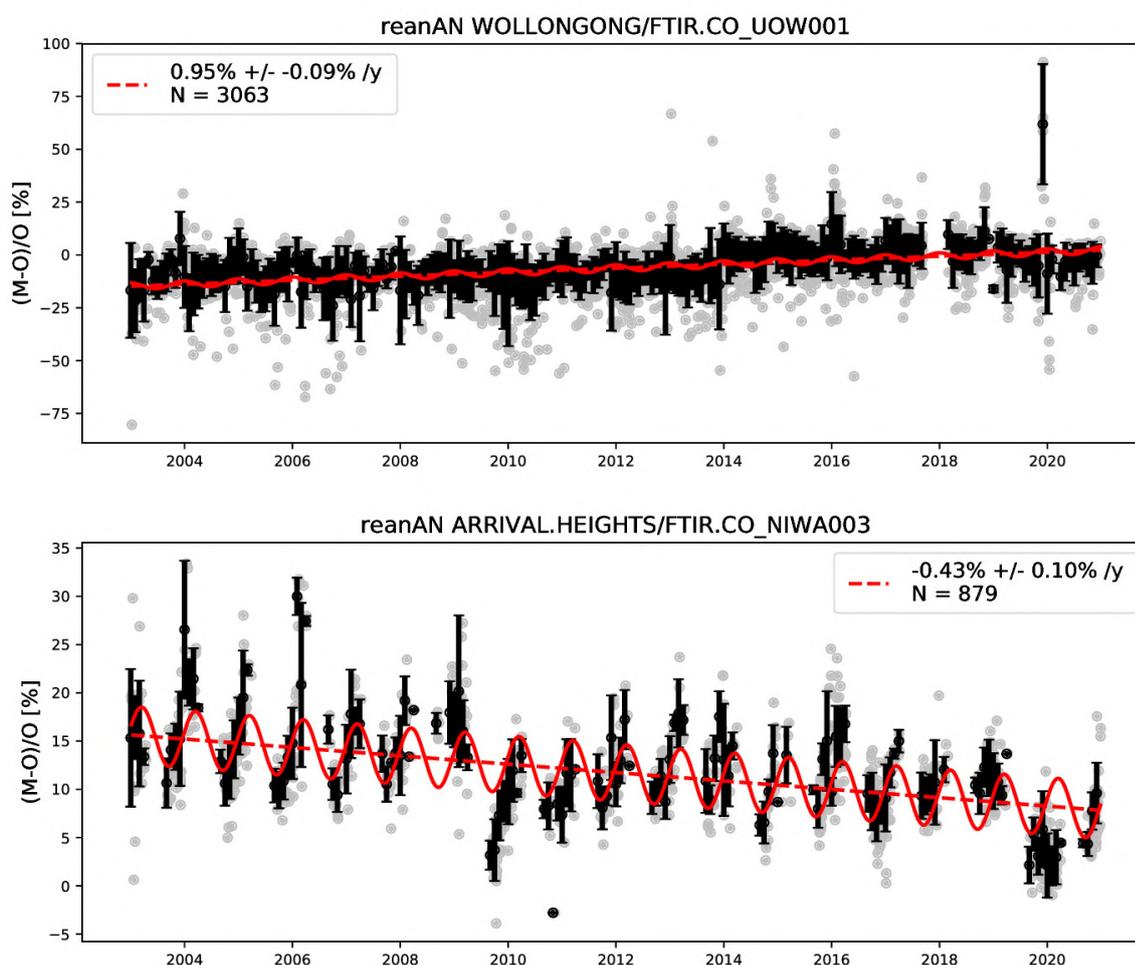


Figure 5.3.6: Trend estimates for the relative bias of the CAMS reanalysis for southern hemispheric NDACC-FTIR stations for the period 2003-2020. At Maïdo, Lauder and Wollongong a positive trend is observed, while at the Antarctic site Arrival Heights a negative trend is seen with a significant seasonal change in the relative differences.

#### 5.4 Validation against FTIR CO observations from the TCCON network

In this section, we compare column averaged mole fractions of CO (XCO) of the CAMS reanalysis model (reanAN) and the control model with TCCON retrievals. Column averaged mole fractions provide different information than the in-situ measurements and are therefore complementary to the in situ data. 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

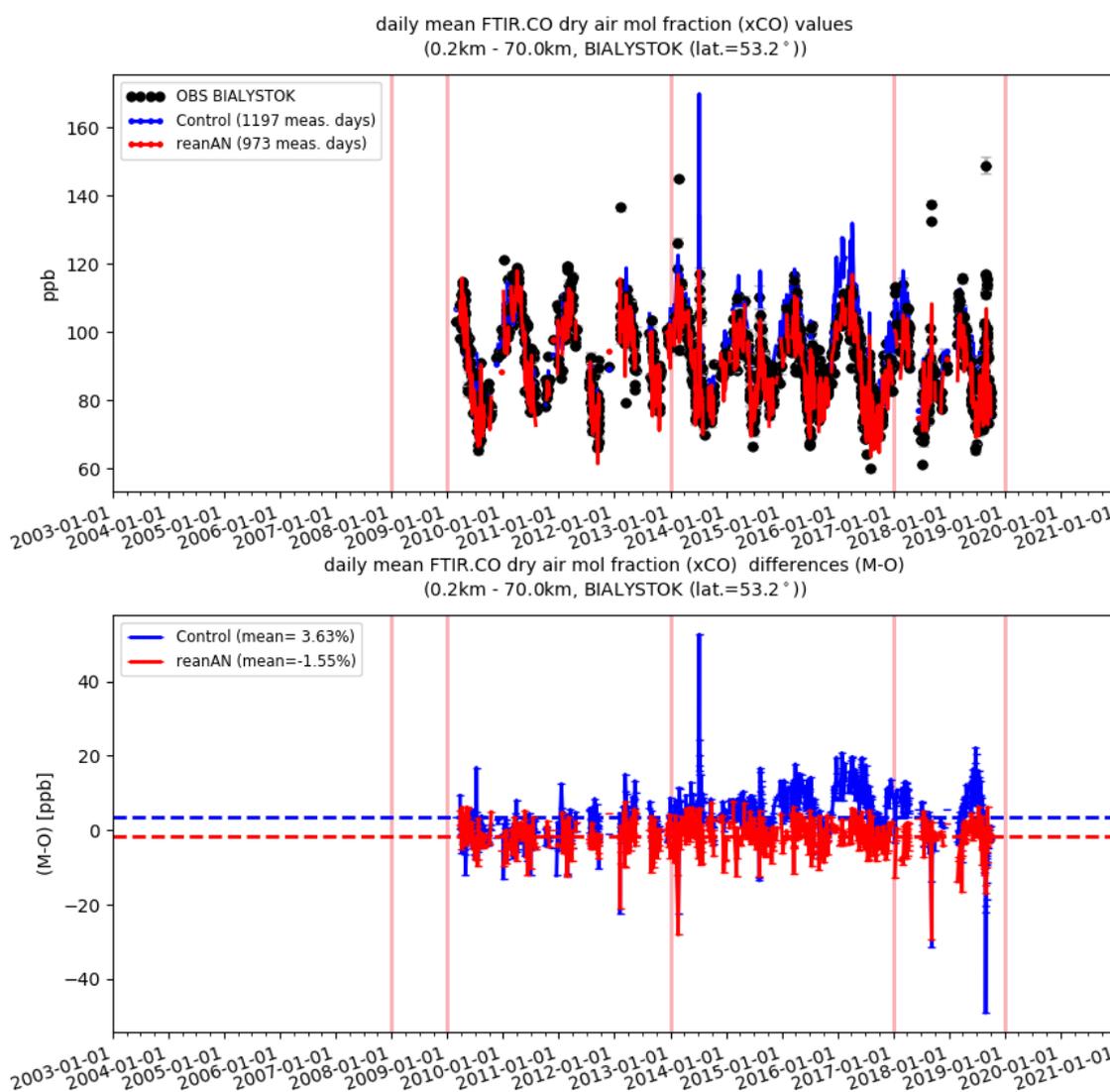


Figure 5.4.1.a: The upper panel shows the time series plots of CO (in ppb) at the TCCON station Bialystok for the period 2003-2020. The lower panel shows the daily difference in CO (in ppb) between the models and the observations.

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). The 2014 data version is described in Wunch et al., 2015. The references for the involved TCCON sites are given in Annex 1.

The TCCON ground-based, remote-sensing measurements are performed in the near IR spectral region and have the sensitivity is highest close to the Earth’s surface (Wunch et al., 2011). The standard TCCON data products are column averaged mole fractions. TCCON obtains the column averaged mole fractions from a ratio of the gas of interest and O<sub>2</sub>. This ratio is then multiplied by the mole fraction of O<sub>2</sub> yielding the dry air mole fraction. Since the variations of the O<sub>2</sub> mole fraction in the atmosphere is several magnitudes smaller than the uncertainty of the TCCON data product it can be regarded as constant. The advantage of this approach is that systematic errors

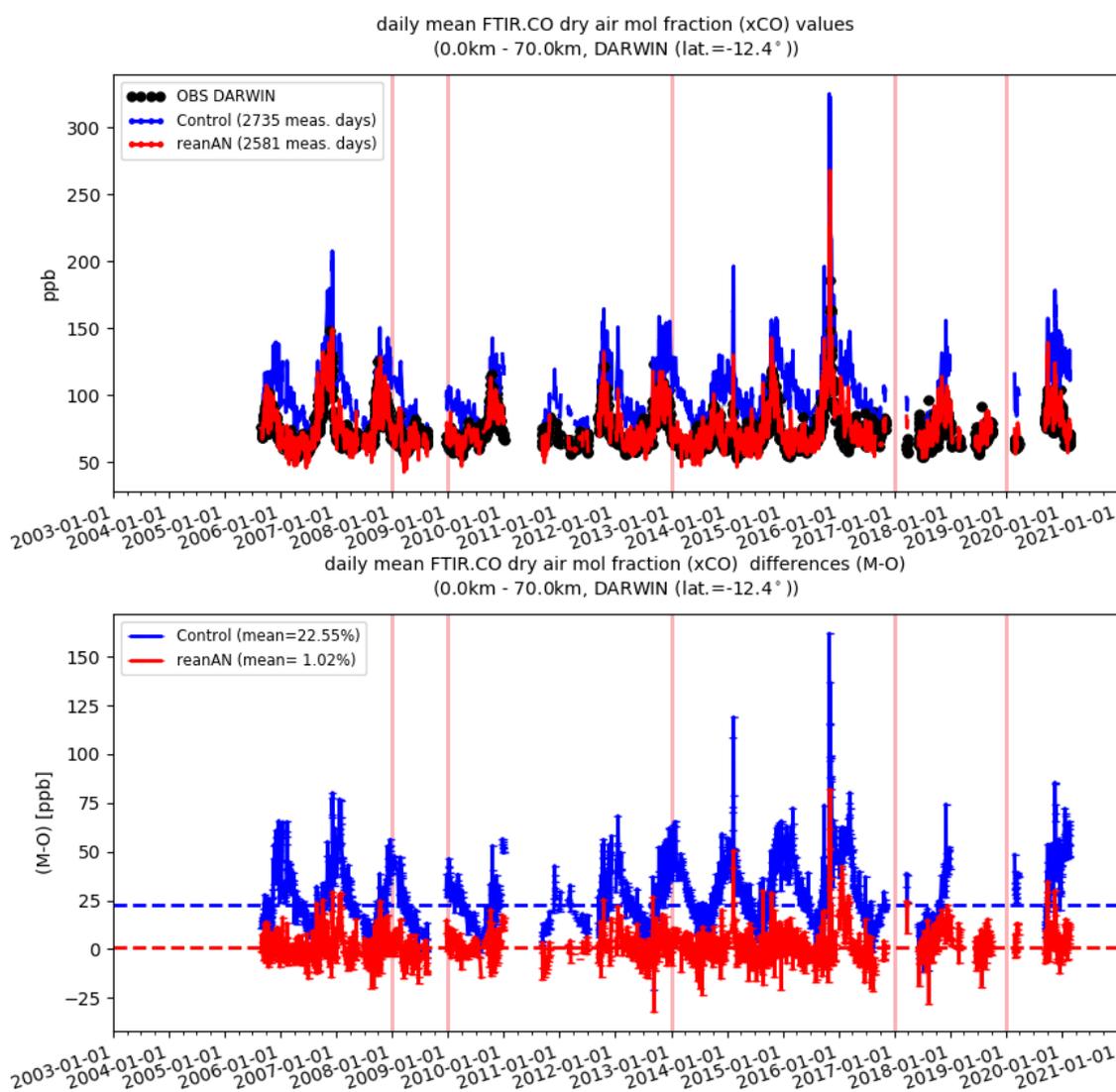


Figure 5.4.1.b: The upper panel shows the time series plots of CO (in ppb) at the TCCON station Darwin for the period 2003-2020. The lower panel shows the daily difference in CO (in ppb) between the models and the observations.

common to the columns of the gas and  $O_2$  partially cancel out. The column averaged mole fractions could also be calculated via the atmospheric pressure corrected by the atmospheric water content.

In the co-location criteria, the line of sight of the FTIR measurement is taken into account and the FTIR averaging kernel is applied to the co-located model data. The validation routines used for TCCON data are the same as used for the NDACC network. The routines have been adapted to use the TCCON data format. A description of the applied methodologies for the comparison is given in Langerock et al. (2015).

An example of such a comparison is shown for the TCCON sites Bialystok, Darwin, Orleans and Lauder in Figure 5.4.1.a-d respectively. While the CAMS reanalysis captures the seasonality quite well, the control model does not.

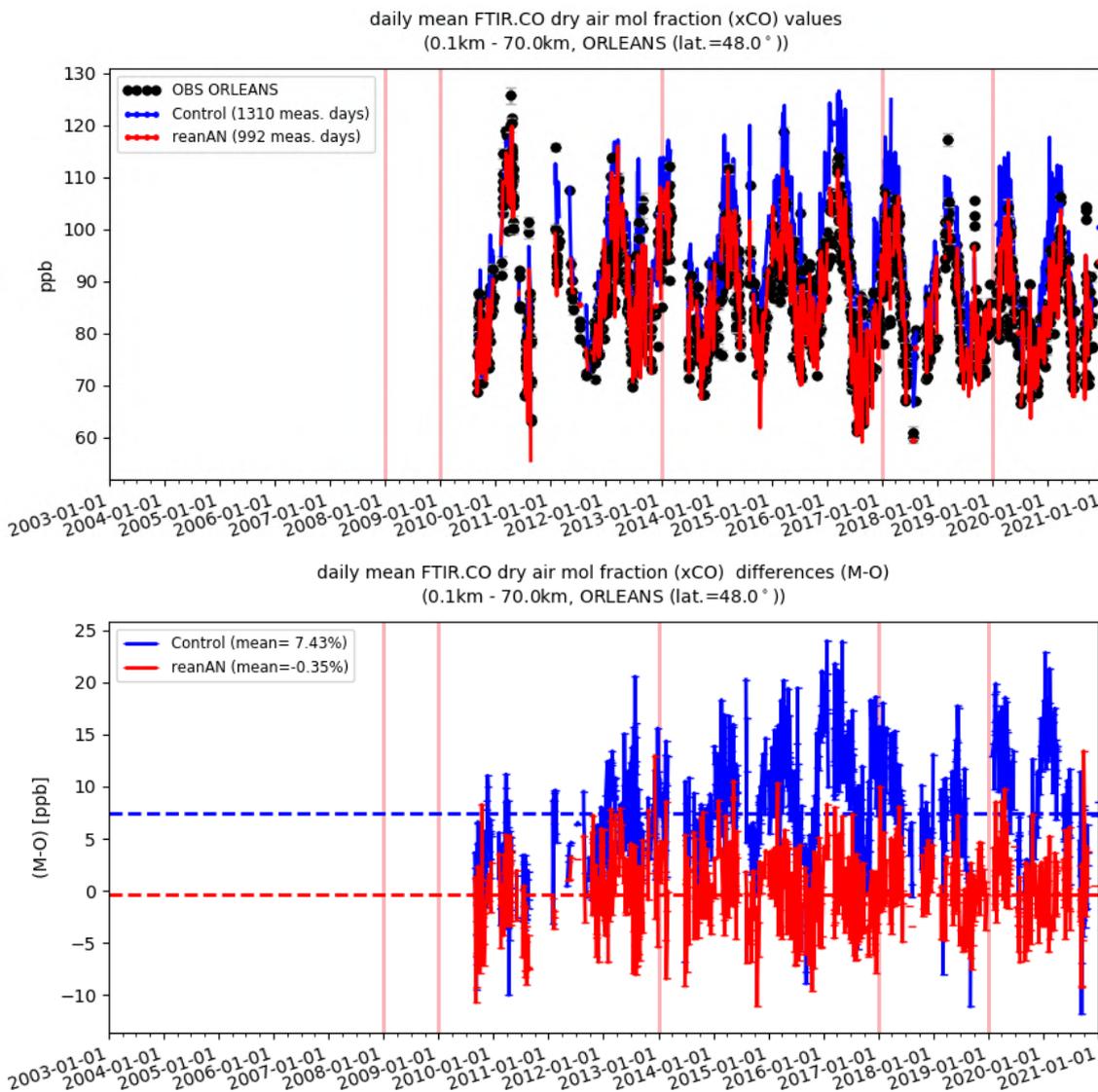


Figure 5.4.1.c: The upper panel shows the time series plots of CO (in ppb) at the TCCON station Orleans for the period 2003-2020. The lower panel shows the daily difference in CO (in ppb) between the models and the observations.

The following list summarizes the conclusions drawn from the figures presented here.

- Both experiments have a similar correlation performance with values ranging between 0.6 and 0.9 for most sites (Fig. 5.4.2).
- The CAMS reanalysis captures well the seasonality for all sites and the whole time period of the comparison: The difference between the CAMS reanalysis and TCCON is for all sites and most of the time series roughly within +/- 5 ppb (Figure 5.4.1.a-d, 5.4.2 and 5.4.3). The control model does not capture the seasonality well and shows an overestimation of the XCO<sub>2</sub> for many sites/times up to 40 ppb (Figure 5.4.1.a-d, 5.4.2 and 5.4.3).

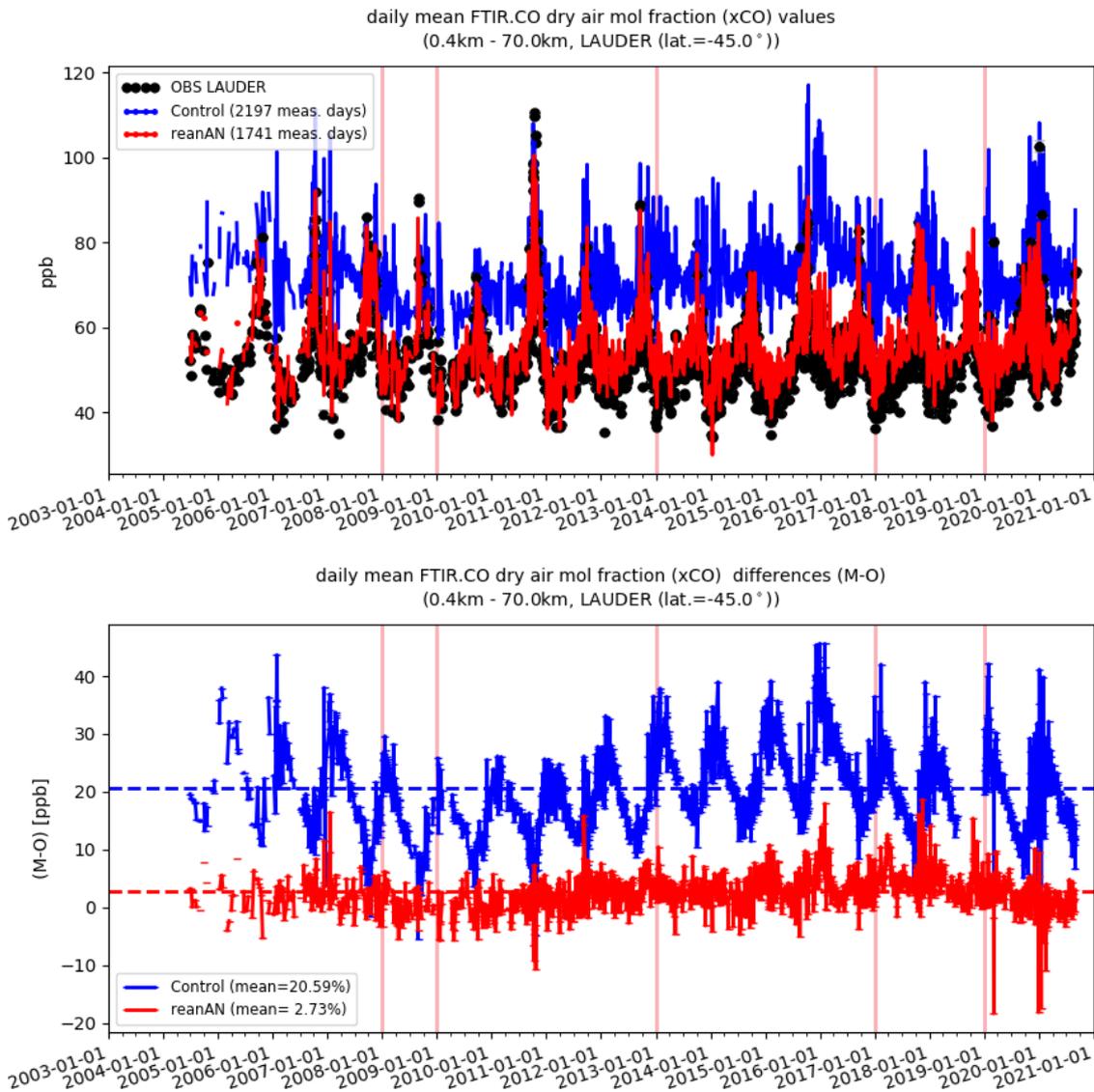


Figure 5.4.1.d: The upper panel shows the time series plots of CO (in ppb) at the TCCON station Lauder for the period 2003-2020. The lower panel shows the daily difference in CO (in ppb) between the models and the observations.

- The long-term (> annual) variation in the differences is small for the CAMS reanalysis (< 5ppb). For control run long-term (> annual) variations in the difference are > 10ppb for many sites. (Figure 5.4.1.a-d and 5.4.3).
- While the CAMS reanalysis performs well in both hemispheres. The control run shows a higher overestimation in the Southern Hemisphere compared to the Northern Hemisphere (Fig. 5.4.2 and 5.4.3).

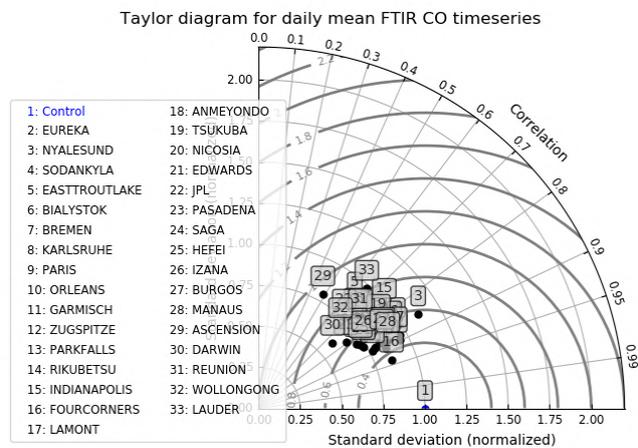
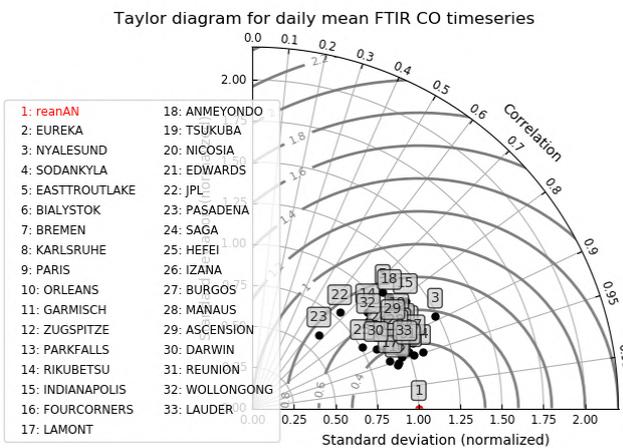
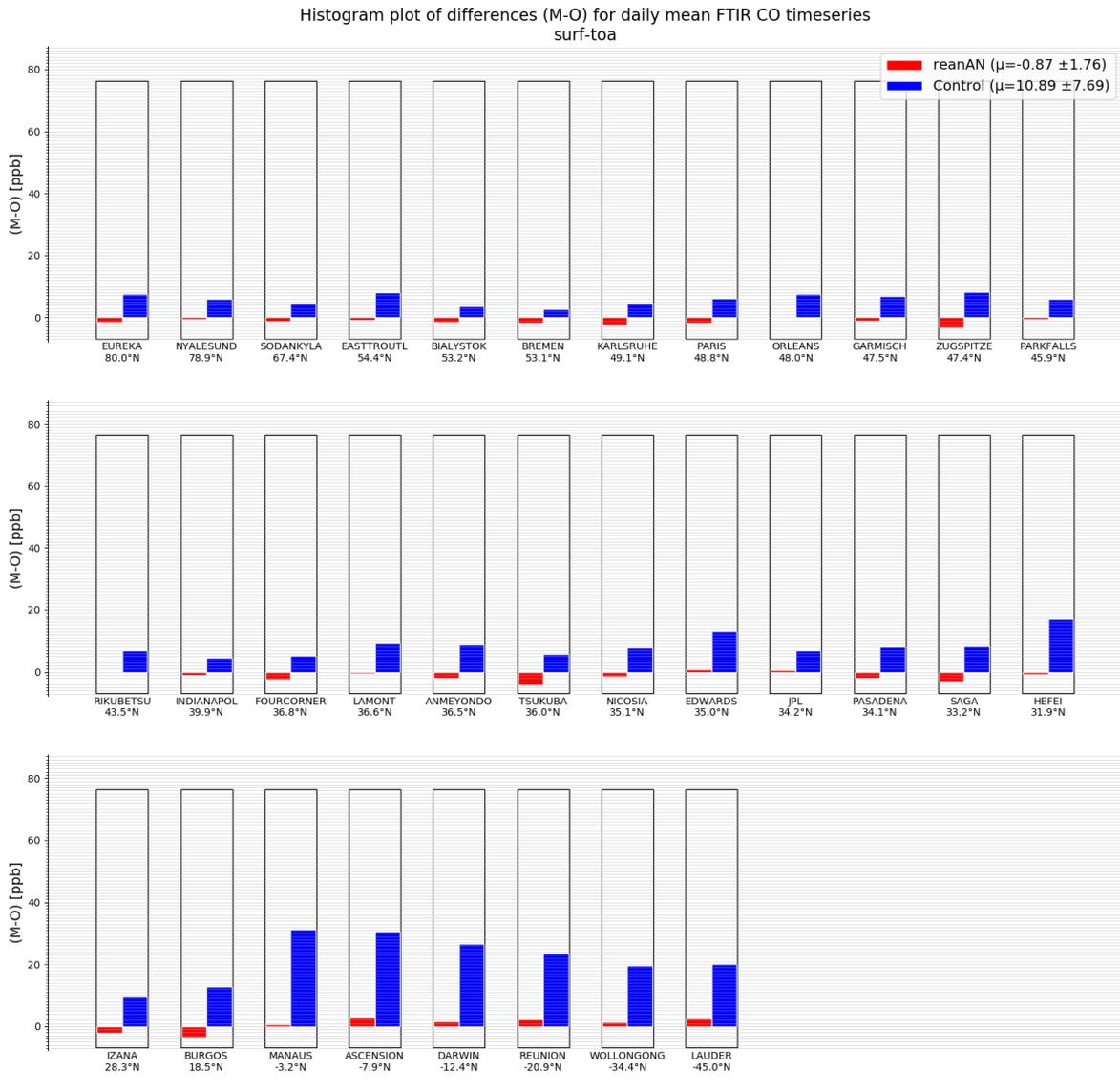


Figure 5.4.2: Histogram plot (top, sorted by decreasing latitude) of averaged daily XCO differences against FTIR for CAMS reanalysis (red) and Control run (blue) and Taylor diagrams (bottom-left: reanalysis; bottom-right: control run) for the period 2003-2020.

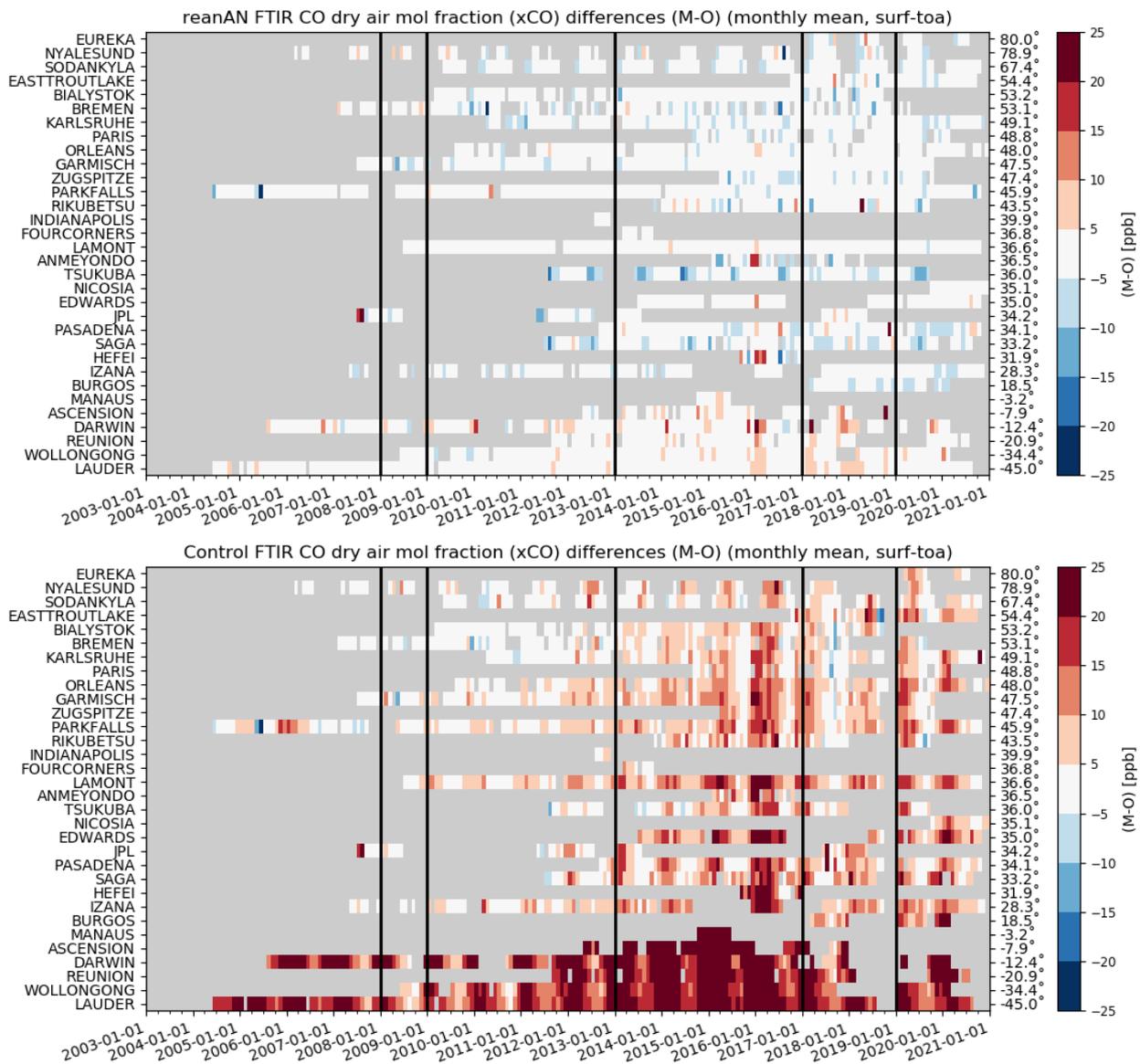


Figure 5.4.3: Mosaic plot of the XCO monthly differences at the TCCON sites for the period 2003-2020. The CAMS reanalysis (top panel) does not show the strong latitudinal bias seen in the control run (bottom panel). The CAMS reanalysis shows a small negative bias in the Northern Hemisphere and a small positive bias in the Southern Hemisphere.



## 5.5 Comparison of the CAMS reanalysis with MOPITT CO

CAMS reanalysis CO total column data are compared with IASI Metop-B satellite observations (Clerbaux et al., 2009) and MOPITT version 8 (thermal infrared radiances) (Emmons et al., 2009, Deeter et al., 2010). For the comparison with the satellite data, the modelled CO concentrations were transformed using averaging kernels from IASI and MOPITT respectively (Deeter, 2004).

The CAMS reanalysis CO total column is in good agreement with the MOPITT observations and have a slight negative bias within 10 % with some exceptions where the underestimation reach 20 % (Fig.5.5.1). From the year 2012 to 2015 the negative bias increased over the Northern Hemisphere low latitudes (up to 20%). End of 2015 was marked by a strong El Nino event which intensified the dry season over the large regions of Indonesia and as a consequence an increase of fire emissions in this region. The reanalysis overestimates the satellite observations over this area up to 20 %. Starting from autumn 2016, the reanalysis data shows reduced negative bias (within 5%) over the globe and improved agreement with the satellite observations. In the first half of year 2019 the CAMS reanalysis data show slightly higher positive bias over the Southern Hemisphere compared to year 2018. Second half of 2019 showing improvements compare to the previous year reflecting in the weaker bias over the Southern Hemisphere. Year 2020 shows a change in the sign of the bias over the Southern Hemisphere from positive to negative, but within 5%. We note that MOPITT V7 NRT data were assimilated in the CAMS reanalysis between 20170101 and 20190630, while before that time it was V6 and afterwards it is NRT V8. In contrast, a slight positive bias (within 5%) appears over the Southern Hemisphere in spring/summer seasons.

The control run overestimates CO values over the Southern Hemisphere and low latitudes in the Northern Hemisphere. Overestimation reaches 50% over the Southern Hemisphere tropics/subtropics in winter/spring seasons. Mid-latitudes in the Northern Hemisphere show negative bias up to 20%.

In comparison with IASI observations, the relative biases are larger (note that CO from IASI is not assimilated). The reanalysis data show a negative bias over the low latitudes within 20 % and over the Northern Hemisphere mid- and high latitudes during the summer season. Over the Southern Hemisphere high latitudes, the reanalysis overestimates CO values up to 30% from year 2011 till 2014 and up to 20% onwards (this is pronounced weaker over the Northern Hemisphere). Note that IASI data versions were changed on 2014.10.01, on 2019.05.15 and on 2019.12.11, which can be reflected in the comparison. Overall, both e-suite and control run show improved performance in year 2020, reflecting in smaller biases compared to satellite observations.

The control run shows strong overestimation compared to IASI data over the Southern Hemisphere (up to 50%), especially during the winter and spring seasons and within 20% during the same seasons over the Northern Hemisphere. An underestimation within 20% during autumn season is observed over the Northern Hemisphere high- and midlatitudes. Year 2020 shows improvements reflected by the smaller biases over the high latitudes for the reanalysis and over the high latitudes and mid-southern latitudes during winter/spring for the control run.

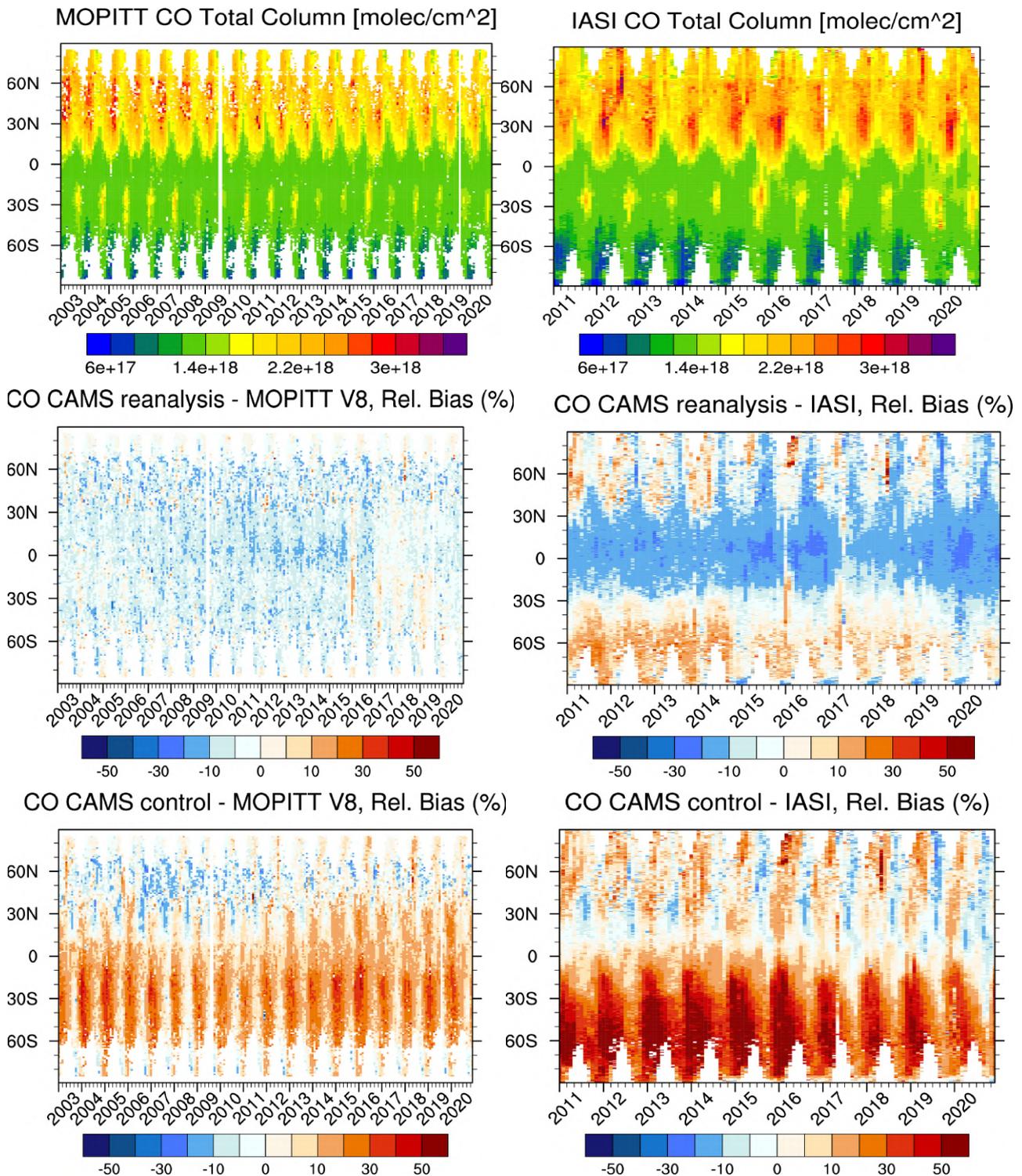


Figure 5.5.1: MOPITT V8 (top left) and IASI CO total column (top right) as function of latitude and time. Relative difference between CAMS reanalysis (middle left) and CAMS control run (bottom left) and MOPITT V8 and relative difference between CAMS reanalysis (middle right) and CAMS control run (bottom right) and IASI. Please, note the different time axis: from January 2003 to December 2020 for MOPITT and from January 2011 to December 2020 for IASI.

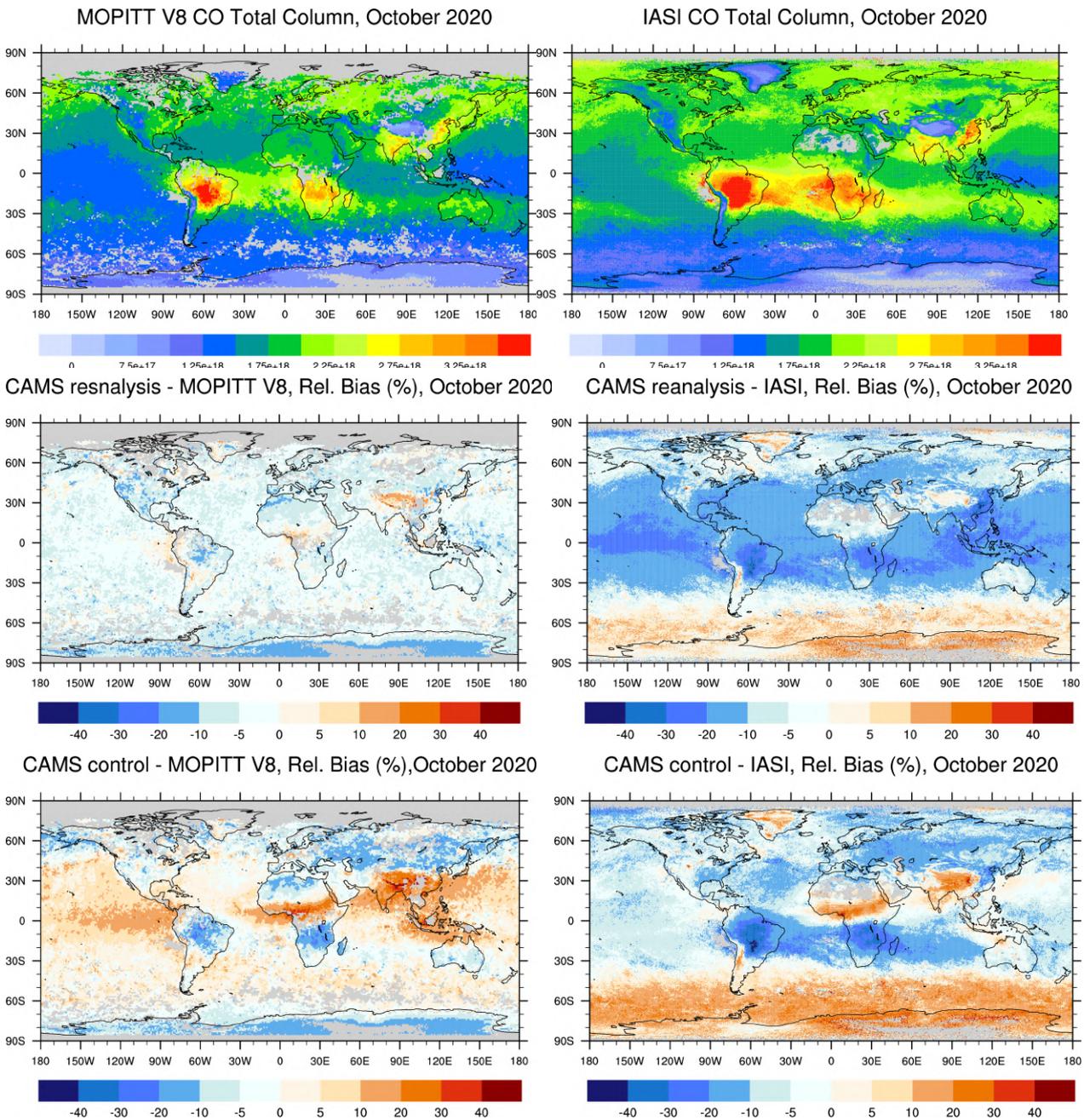


Figure 5.5.2: CO total column for MOPITT V8 (top left) and IASI (top right) satellite observations and relative difference between the reanalysis and MOPITT (middle left), reanalysis and IASI (middle right), control and MOPITT (bottom left), control and IASI (bottom right) for October 2020. Grey colour indicates missing values.

The global distributions of CO total columns from MOPITT and IASI and relative biases of the model runs, averaged for October 2020, are shown in Fig. 5.5.2. Both, MOPITT and IASI show high CO values over the biomass burning regions in South America and Africa and over east of China and India, with IASI values being higher than MOPITT.



The reanalysis data show good agreement with MOPITT observations with a bias within 10%. The control run shows overestimation over the equatorial band and over China, Indonesia and India up to 30% and underestimation over some areas in Eurasia and over the biomass burning areas in South America and Southern Africa, up to 20%.

In comparison with IASI data, the reanalysis shows a pronounced negative bias over the low latitudes within 20% and overestimation over southern high latitudes up to 30%. The control run has a negative bias over biomass burning regions in South America and southern Africa up to 30% and over Eurasia within 20%. An underestimation is found over the Sahel region, north of China and for high southern latitudes, with values up to 30%.

The time series of CO total column from MOPITT, IASI and the model data with applied averaging kernels from IASI, over the eight selected regions from 2011 to December 2020 are shown in Fig. 5.5.3. The two instruments show a relatively similar seasonality of CO total columns, but there is a systematic difference over most regions, indicating that CO values from MOPITT are higher than for IASI, especially during the winter and spring seasons (except for North and South Africa). The difference between the two observational data sets increased in 2011, 2012, 2013 and 2014, which is especially notable in Alaskan and Siberian fire regions where the difference reaches 40%. This is due to decreased IASI CO values, which was not observed by the MOPITT instrument.

The CAMS reanalysis and control run both capture the seasonal and interannual variability of CO total columns, see Fig. 5.5.3. The reanalysis data show underestimations over most of the regions, with exceptions for the Alaskan and Siberian fire regions during February-April where the models overestimate the observations. In Europe and US, after the year 2014, the reanalysis data are close to the observations during the winter season (bias within 5% for Europe and 10% for US) and show a maximum bias at the end of summer/beginning of autumn (bias up to 20%). The control run shows a maximum overestimation during winter (up to 15%) and underestimation around September (up to 15%). The reanalysis is in good agreement with the observations over the Siberian and Alaskan fire regions with generally a bias within 10%, with the exception of an underestimation in the summer season (up to 20%). Both reanalysis and control run are able to capture big fire events in Siberia, leading to enhanced CO values in summer 2012 (bias is within +5% for the control and within -10% for the reanalysis run). In the South Asian region, the reanalysis shows a good agreement, with a bias within 10% which increased in the year 2020 to up to 15%. The control run shows overestimation in winter/spring seasons up to 20%. The reanalysis data captured very well the large fires in South Asia at the end of 2015, while the control run overestimated the event by about 25%. Over the East Asian region, the reanalysis underestimated observations up to 20% during the low emission seasons (summer/autumn). The control run shows better agreement with a bias within 10% (except for winter season with a positive bias up to 15%). The reanalysis shows a negative bias of 20% over the North African region, up to 30% over the South African region. The control run, in contrast, shows a positive bias up to 20% over the North African region and up to 30% over the South African region.

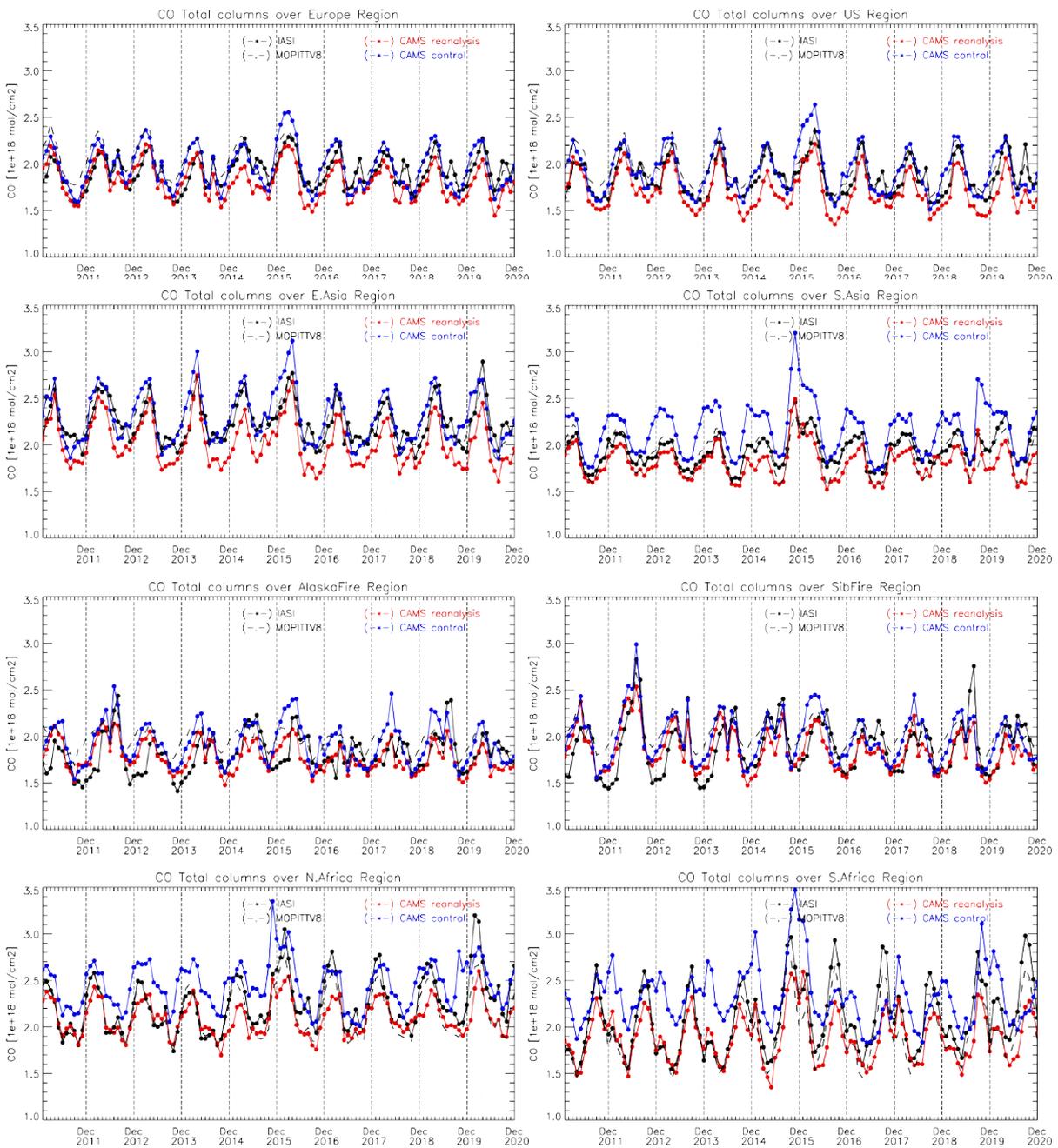


Figure 5.5.3: CO total columns for satellite retrievals MOPITT V8 and IASI (black) and reanalysis (red) and control (blue) data over selected regions from January 2011 to December 2020. The modelled CO concentrations were transformed using averaging kernels IASI.

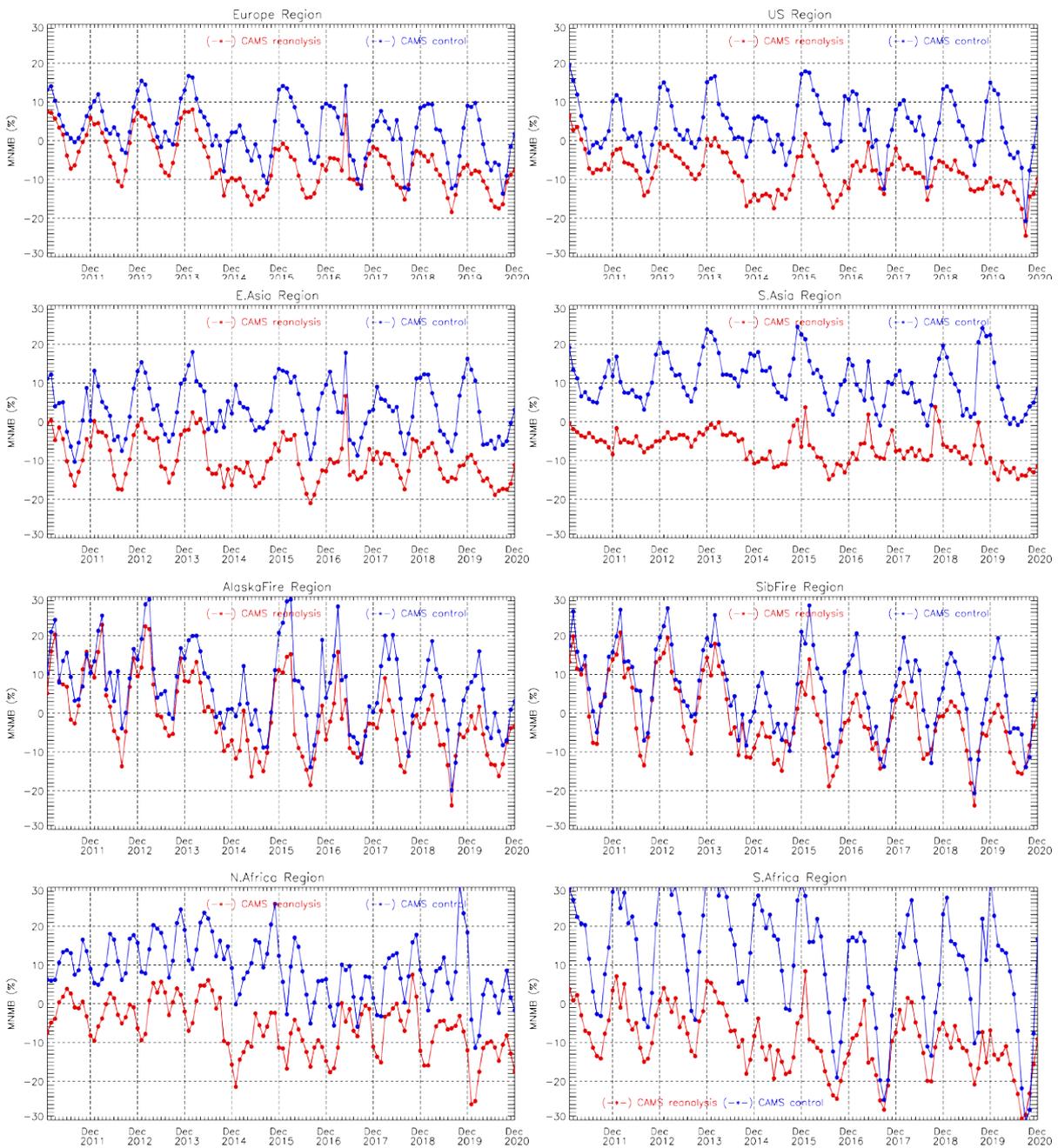


Figure 5.5.4: Time series of modified normalized mean bias (%) for CO total columns from the model simulations vs. IASI retrievals over selected regions from January 2011 to December 2020.

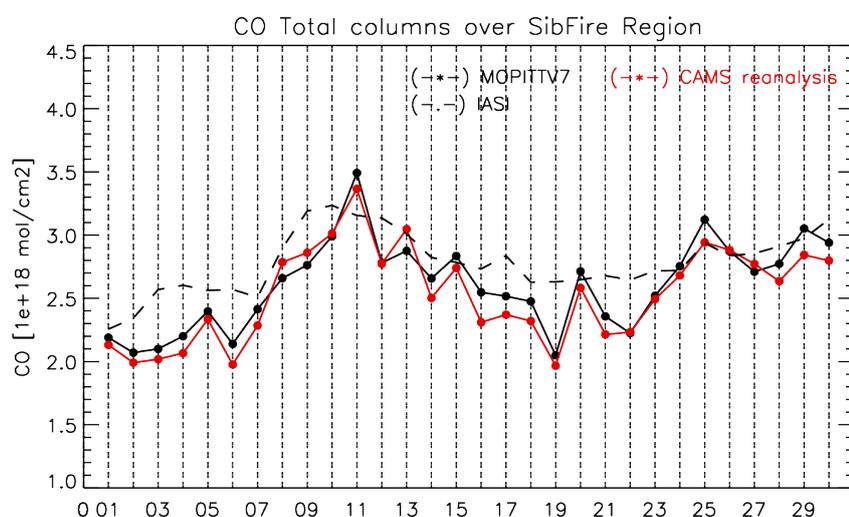


Figure 5.5.5: CO total column from satellite retrievals MOPITT V7 and IASI (black) and reanalysis (red) over Siberian region during July 2012.

### Case study: Large fire events in Siberia in summer 2012

The Summer 2012 was characterized by a large number of fire events in Siberia and high amount of fire emissions transported across Pacific Ocean towards the west coast of North America. Satellite observations show a clear increase of CO values over Siberian region during July 2012 (Fig. 5.5.5.). IASI data are generally higher than MOPITT by about 15 %. The reanalysis reproduces the enhanced CO values. The CAMS reanalysis data show good agreement with the MOPITT observations with difference of less than 10 %.

The geographical distribution of IASI CO data indicates a large emission plume transported from an active fire region towards the west coast of North America over the Pacific Ocean (Fig. 5.5.6.). Both runs capture location of the plume and transpacific transport of CO emissions. The reanalysis overestimates CO total columns over the fire regions. In the transportation pathway over the ocean, the reanalysis data show underestimation.

### Case study: Fire events in Indonesia in autumn 2015

The year 2015 was marked by a strong El Nino event which intensified the dry season over large parts of Indonesia. During September and October 2015, the largest fire emissions were recorded in Indonesia since 1997, based on GFAS and GFED emissions time series (Huijnen et al., 2016). Daily CO total columns from the CAMS reanalysis over a region covering Indonesia (70E-150E and 11S-6N) for September 2015 were compared to IASI and MOPITT data. Observations show an area-average increase of CO total column from  $\sim 1.8 \times 10^{18}$  molec/cm<sup>2</sup> at the beginning of September to  $\sim 2.6 \times 10^{18}$  molec/cm<sup>2</sup> in MOPITT and  $3.7 \times 10^{18}$  molec/cm<sup>2</sup> in IASI by the end of September, indicating IASI data up to 40 % higher compared to MOPITT (Fig. 5.5.7). Evaluation of the model results shows that the CAMS reanalysis CO total columns are in good agreement with the MOPITT observations with a positive bias of less than 15 %.

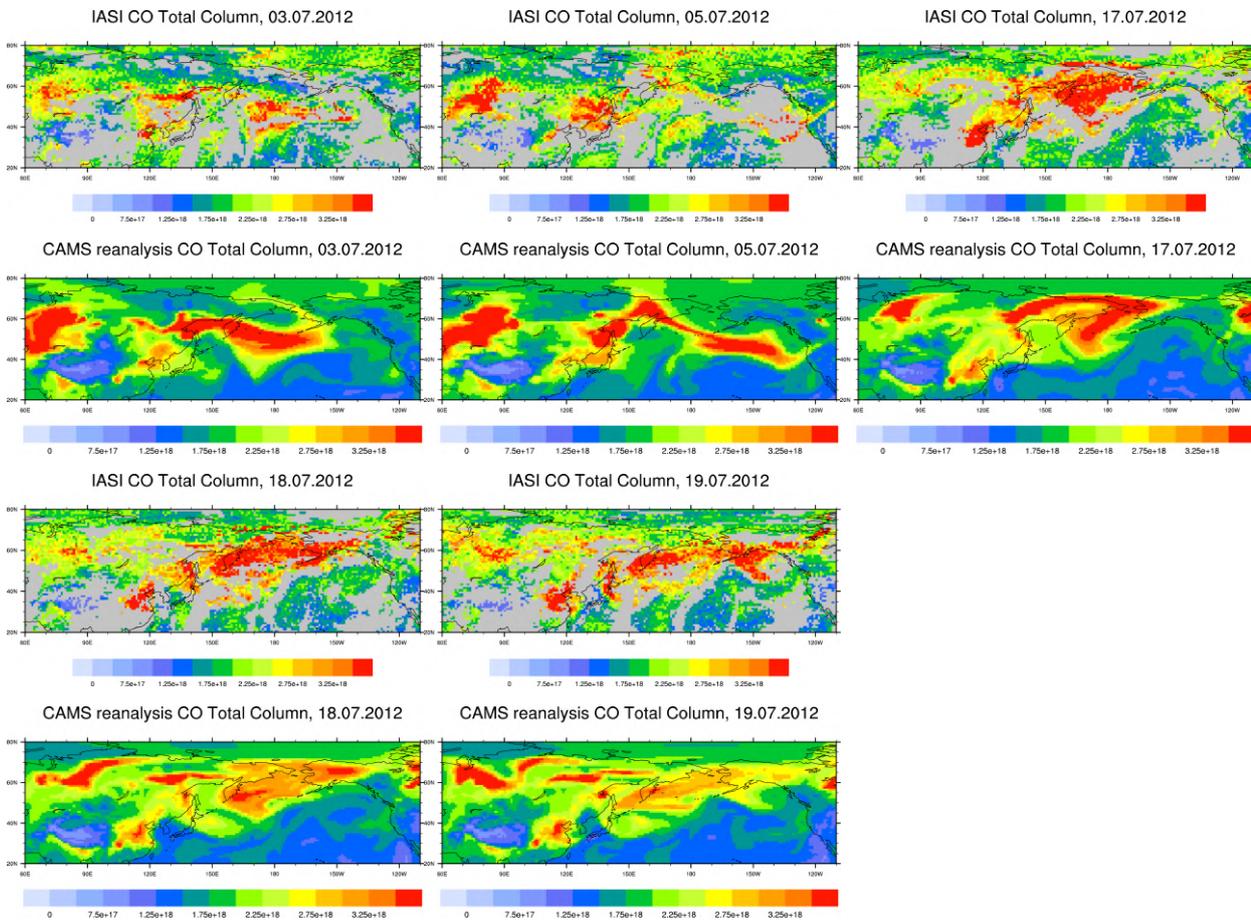


Figure 5.5.6: CO total column over the Siberian region and the North Pacific for 03, 05, 17, 18 and 19 July 2012 from IASI (first and third row) and the CAMS reanalysis (second and fourth row). Unit: molecules / cm<sup>2</sup>.

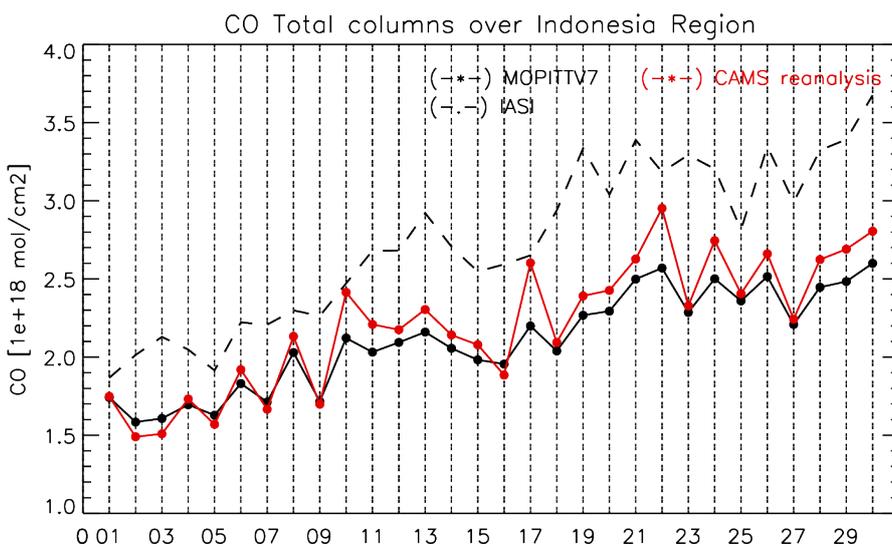


Figure 5.5.7: CO total column from satellite observations MOPITT V7 and IASI (black) and reanalysis (red) over Indonesia (70E-150E and 11S-6N) during September 2015.

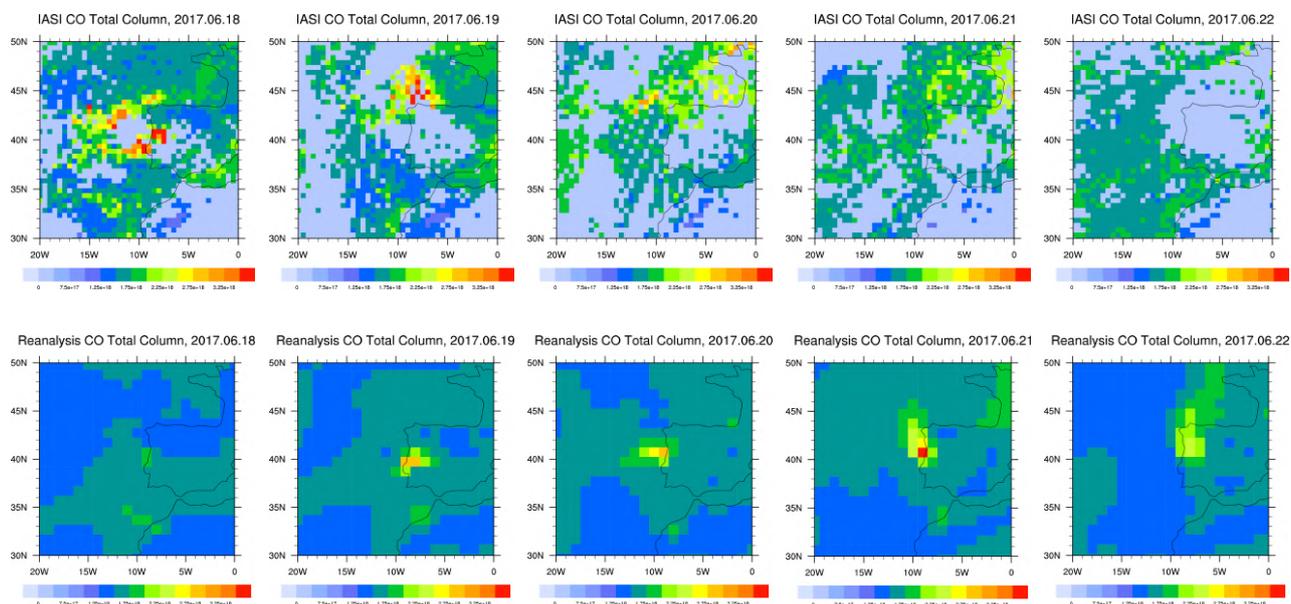


Figure 5.5.8: Map of CO total columns for IASI satellite observations (top panels) and CAMS reanalysis data (bottom panels) over Portugal between 18 and 22 June 2017.

### Case study: Fire event in Portugal, June 2017

On June 17, 2017, lightning reportedly ignited a wildfire that spread across the mountainous areas of Pedrogao Grande – a municipality in central Portugal located about 160 kilometers northeast of Lisbon. IASI data show rapidly increasing CO over this region starting from 18<sup>th</sup> June (Fig. 5.5.8). CAMS reanalysis data show the increase of CO the next day, on June 19<sup>th</sup> and maximum values on June 21<sup>st</sup>. The emission patterns are less extended compared to IASI. **Case study: Fire events in the western United States, August-September 2020**

The western United States experienced many severe wildfires in the year 2020, and, at times, the intensities were ten to hundred times higher than the 2003-2019 average (<https://atmosphere.copernicus.eu/cams-monitors-smoke-release-devastating-us-wildfires>). The fire activity started in California and Colorado during the second half of August and seriously affected also Oregon and Washington since the beginning of September.

Daily CO total columns from the CAMS reanalysis and control runs over the US region (65° - 120°W and 30° - 45°N) during 15 August – 20 September 2020 were compared to IASI and MOPITT data. The satellite observations show the increase of carbon monoxide during 19 – 29 August and strong enhancements during the first half of September, with maximal emissions on 15-16 September (Fig. 5.5.9). The reanalysis well reproduces the timing and transport of the pollution plumes but shows an underestimation of up to 20% in August and up to 40% during the peak of emissions in September. The control run performs better, with a smaller negative bias. Note that for this comparison the modelled data were convoluted with the IASI averaging kernels. The map of the geographical distribution of the plume location and the transport (Fig. 5.4.8) shows that the locations of the fireplaces agree with the satellite data. The strengths of the plume and the magnitude of emissions on the transportation pathways are underestimated. It is observed that the further away from the fire, the more the model underestimates the CO concentrations.

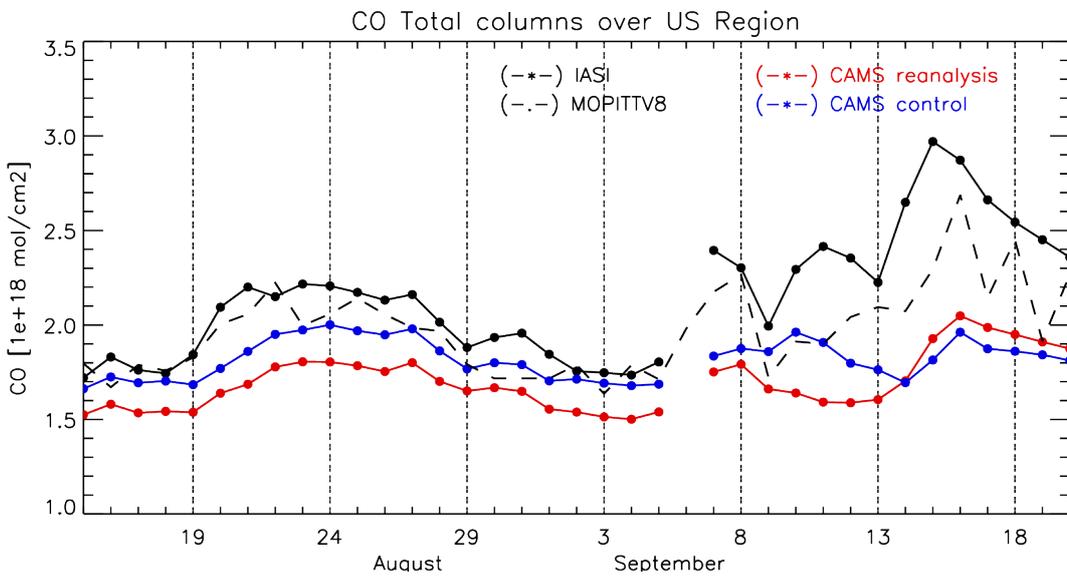


Figure 5.5.9: CO total column from satellite observations IASI and MOPITT V8 (black) and reanalysis (red) over the US region (65° – 120°W and 30° – 45°N) during 15 August – 20 September 2020.

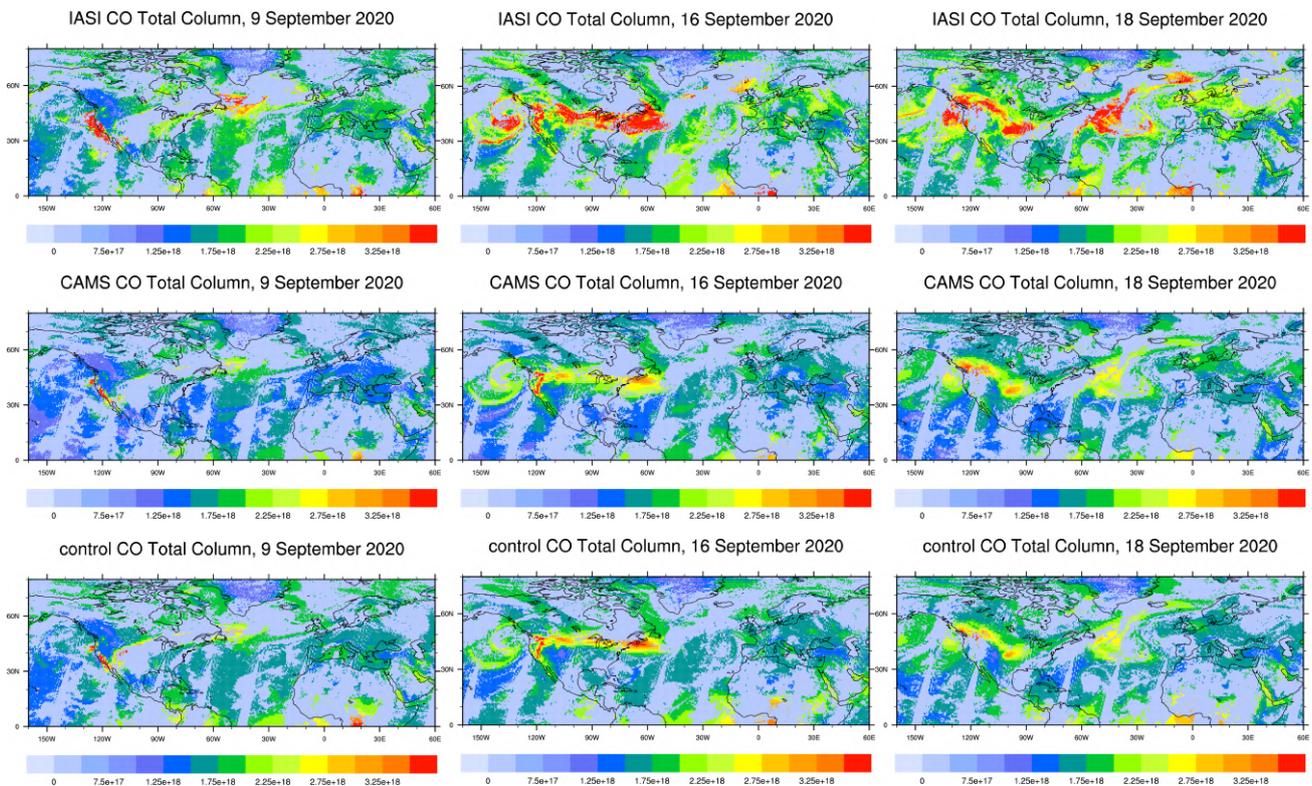


Figure 5.4.8: CO total columns over the North America and Atlantic region (160°W – 60°E; 0° – 80°N) on 9, 16 and 18 September from IASI (top), the CAMS reanalysis (middle) and control runs (bottom). Unit: molecules / cm<sup>2</sup>. Note that region differs from figure 5.4.7.



## 6. Tropospheric nitrogen dioxide and formaldehyde

### 6.1 Evaluation against SCIAMACHY and GOME-2 NO<sub>2</sub> retrievals

In this section, model columns of tropospheric NO<sub>2</sub> are compared to SCIAMACHY/Envisat NO<sub>2</sub> satellite retrievals (IUP-UB v0.7) [Richter et al., 2005] for model data before April 2012, and to GOME-2/MetOp-A NO<sub>2</sub> satellite retrievals (IUP-UB v1.0) [Richter et al., 2011] for more recent simulations. This satellite data provides excellent coverage in space and time and very good statistics. However, only integrated tropospheric columns are available, and the satellite data is always taken at the same local time, roughly 10:00 LT for SCIAMACHY and 09:30 LT for GOME-2, and at clear sky only. Therefore, model data are vertically integrated, interpolated in time and then sampled to match the satellite data. GOME-2 data were gridded to model resolution (i.e. 0.75° x 0.75°). Model data were treated with the same reference sector subtraction approach as the satellite data. Uncertainties in NO<sub>2</sub> satellite retrievals are large and depend on the region and season. Winter values in mid and high latitudes are usually associated with larger error margins. As a rough estimate, systematic uncertainties in regions with significant pollution are on the order of 20% – 30%. Note that SCIAMACHY and GOME-2 NO<sub>2</sub> observations are assimilated in the CAMS reanalysis but using a different retrieval product.

Time series of tropospheric NO<sub>2</sub> for selected regions are shown in Fig. 6.1.1. Apart from East Asia for the year 2003 and apart from an underestimation of the seasonal cycle due to an underestimation of wintertime values over East Asia and also in part for Europe probably related to uncertainties in anthropogenic emissions or photochemistry, the seasonality and magnitude of satellite values is in general reasonably represented by the CAMS reanalysis for the regions investigated. Over East Asia in 2003, the CAMS reanalysis shows a strong variation of values from one month to another and completely fails to reproduce the observed seasonality, due to the assimilation of NO<sub>2</sub> satellite data of degraded data quality for this year. The increase in wintertime values over East-Asia up to 2014 and the following decrease for recent years observed by satellite is reproduced by the CAMS reanalysis. This is in contrast to the control, MACC reanalysis as well as o-suite model results shown in the last CAMS validation report. This could be the result of assimilating NO<sub>2</sub> from more instruments (with different overpass times) in the CAMS reanalysis, i.e., SCIAMACHY, OMI and GOME-2A. However, the magnitude of the positive trend up to 2014 and following negative trend are underestimated by the CAMS reanalysis compared to the satellite observations, so that satellite and model results agree better for the beginning and last years compared to the rest of the time series. For 2017-2020 however, a strong overestimation of summertime values is apparent. The unusually high satellite values for December 2020 over East-Asia and Europe are possibly caused by a problem with the satellite retrievals for this month and are still under investigation.

The CAMS reanalysis tends to underestimate values around December to April over South Africa (here the MACC reanalysis performs better) , while December/January values are overestimated for North-Africa probably related to an overestimation of fire emissions in this region (the CAMS reanalysis performs better here). Satellite observations in winter are associated with larger uncertainties, but inaccuracies in winter NO<sub>x</sub> emissions and NO<sub>2</sub> lifetime in highly polluted regions such as in China could lead to larger model uncertainties during winter.

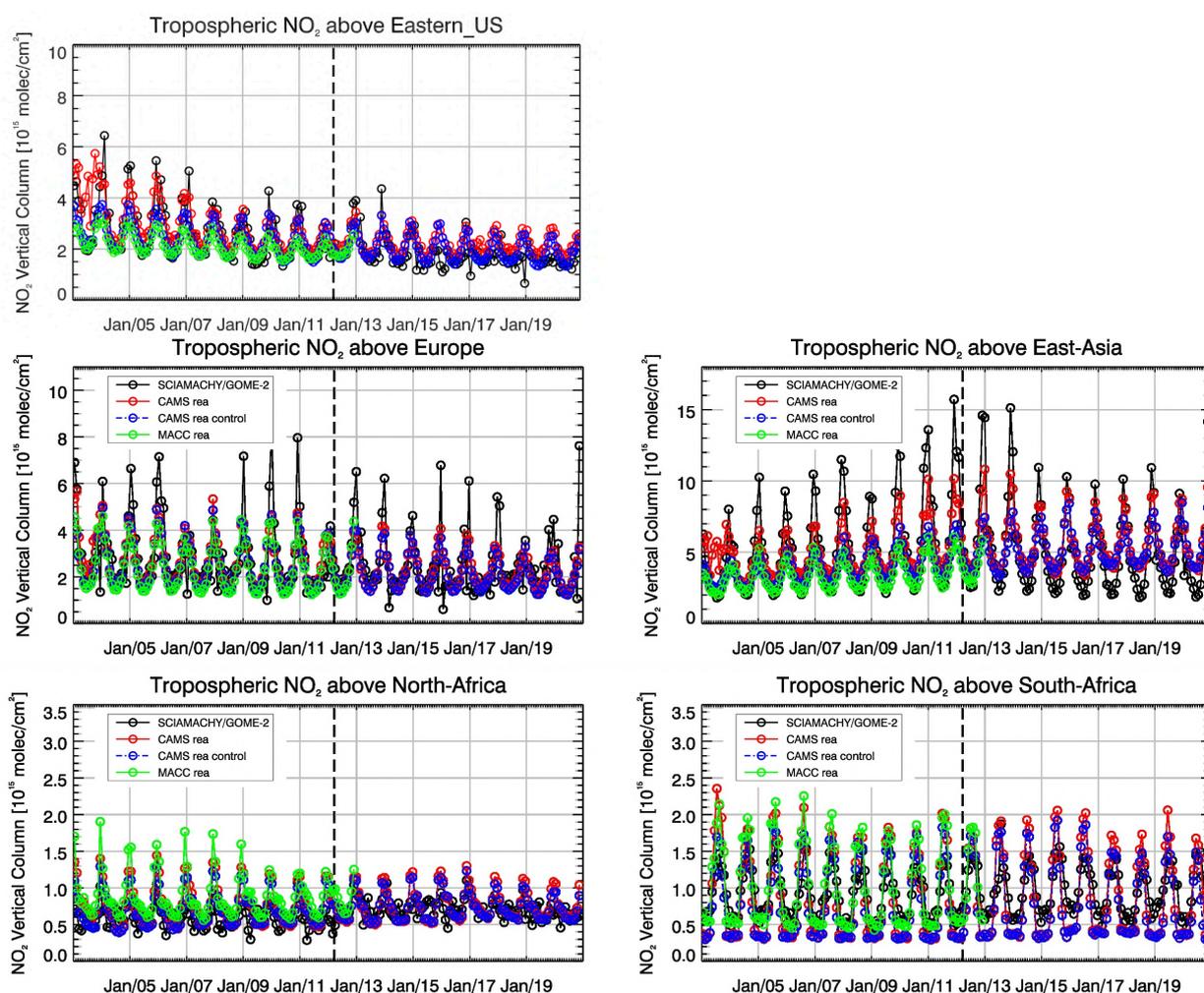


Figure 6.1.1: Comparison of time series of tropospheric  $\text{NO}_2$  columns from SCIAMACHY (up to April 2012) and GOME-2A (from April 2012 onwards) to model results over selected regions for the years 2003-2020. The switch from SCIAMACHY to GOME-2A is indicated by the vertical black dashed lines. Upper panels represent regions dominated by anthropogenic emissions; lower panels represent those dominated by biomass burning. The unusually high satellite values for December 2020 over East-Asia and Europe are possibly caused by a problem with the satellite retrievals for this month and are still under investigation.

Global monthly mean map comparisons (see Fig. 6.1.2 for an example of seasonal averages for 2020) show that the overall spatial distribution and magnitude of tropospheric  $\text{NO}_2$  is well reproduced by the CAMS reanalysis, indicating that emission patterns and  $\text{NO}_x$  photochemistry are reasonably represented. Some differences are apparent between observations and simulations, with generally larger shipping signals simulated by the models. Boreal forest fire emissions are regularly overestimated for example over Siberia in August 2020 (which can be seen in the JJA mean shown in Fig. 6.1.2). This overestimation also shows up for tropospheric  $\text{HCHO}$ , but only over parts and not all of the regions with overestimations of tropospheric  $\text{NO}_2$  columns (see section 6.2). This may point to uncertainties regarding fire emission factors, with the results in general indicating a different performance depending on the trace gas, region and season. The CAMS reanalysis overestimates values over the Persian Gulf and the Red Sea from June to October. Values over

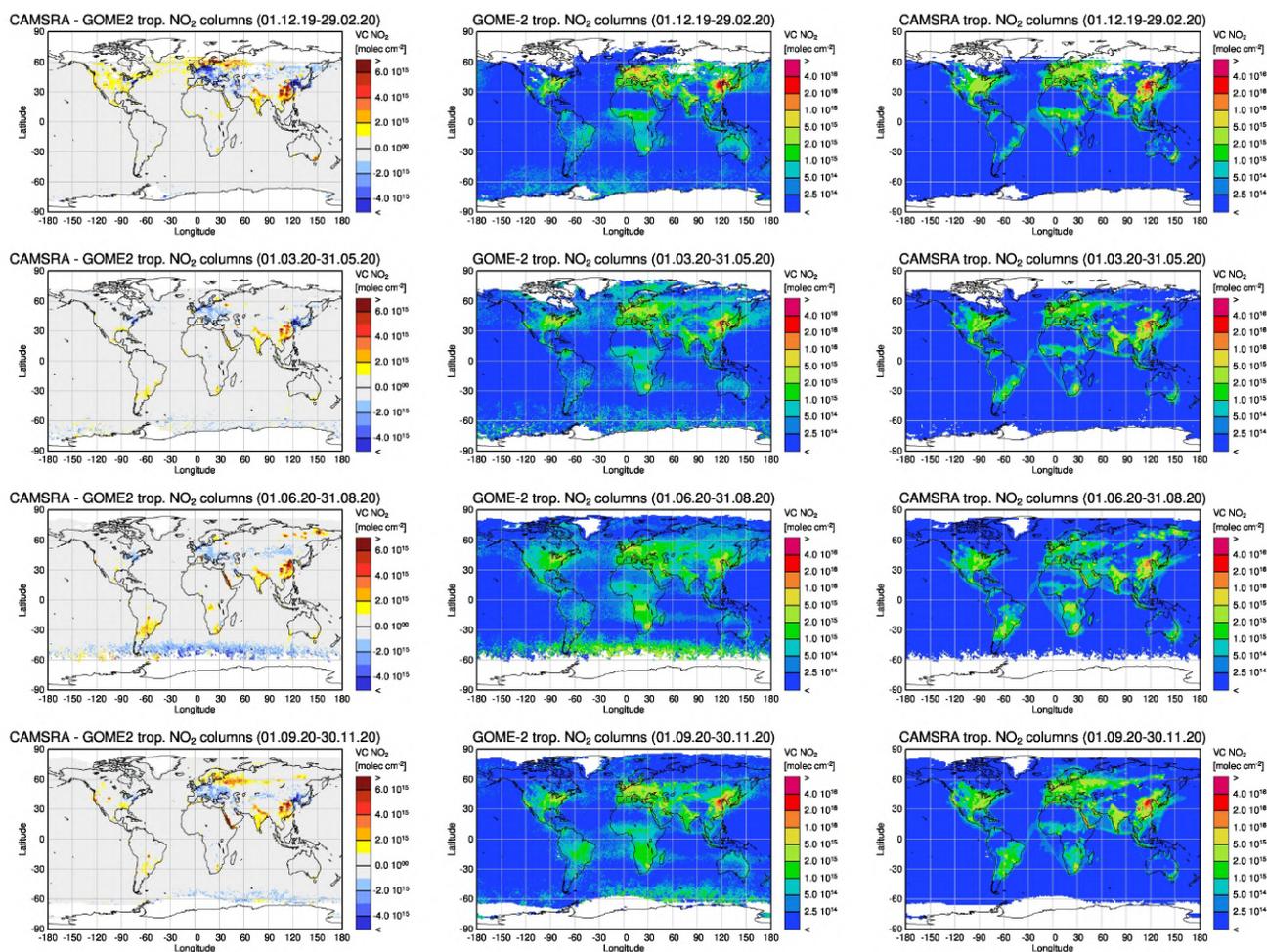


Figure 6.1.2: Global map comparisons of satellite-retrieved and model-simulated seasonally averaged tropospheric  $\text{NO}_2$  columns [ $\text{molec. cm}^{-2}$ ] for (from top to bottom): DJF 2019/2020, MAM 2020, JJA 2020 and SON 2020. Shown are mean tropospheric  $\text{NO}_2$  columns for (from left to right): GOME-2A, CAMS reanalysis and the difference between CAMS reanalysis and GOME-2A. GOME-2A data were gridded to model resolution (i.e.,  $0.75^\circ \times 0.75^\circ$ ). Model data were treated with the same reference sector subtraction approach as the satellite data.

anthropogenic pollution hotspots broadly around the Benelux countries and the German Ruhr area tend to be underestimated (this underestimation is most pronounced in winter), while others are overestimated (e.g., Moscow and Helsinki).

## 6.2 Validation against SCIAMACHY HCHO satellite data

In this section, simulations of tropospheric formaldehyde are compared to SCIAMACHY/Envisat HCHO satellite retrievals (IUP-UB v1.0) [Wittrock et al., 2006] for model data before April 2012 and to GOME-2/MetOp-A HCHO data (IUP-UB v1.0) [Vrekoussis et al., 2010] afterwards. As the retrieval is performed in the UV part of the spectrum where less light is available and the HCHO absorption signal is smaller than that of  $\text{NO}_2$ , 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.

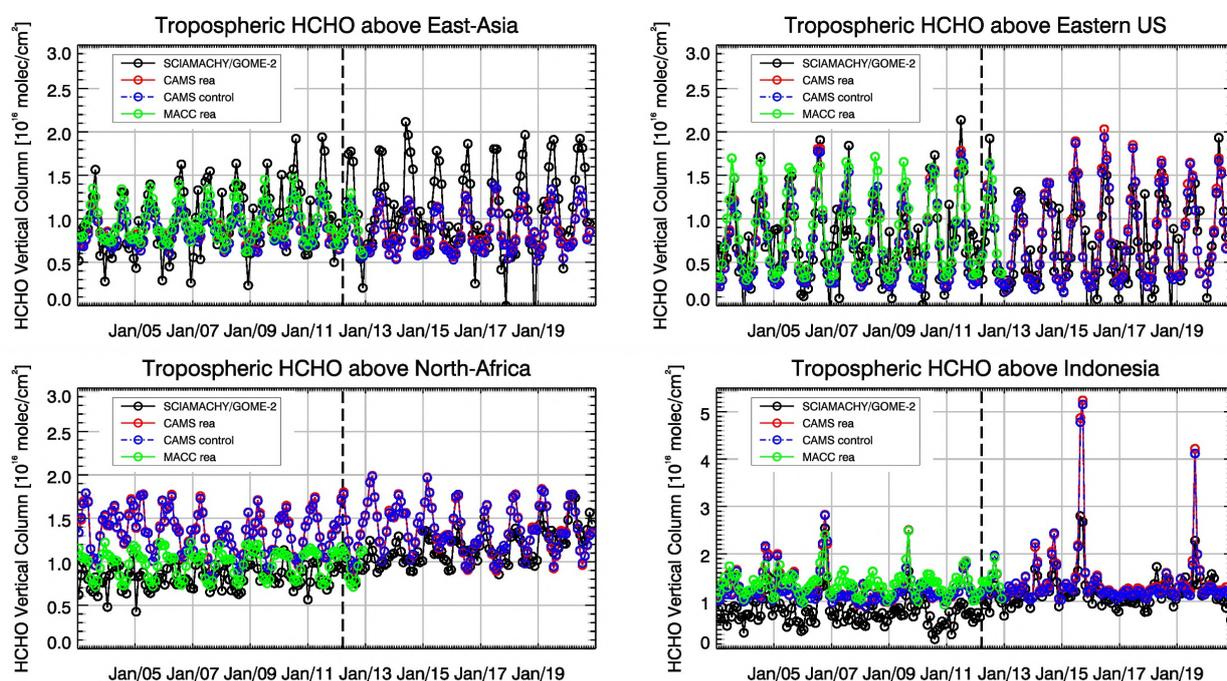


Figure 6.2.1: Comparison of time series of tropospheric HCHO columns from SCIAMACHY (up to April 2012) and GOME-2A (from April 2012 onwards) to model results over selected regions. The switch from SCIAMACHY to GOME-2A is indicated by the vertical black dashed lines. The regions differ from those used for NO<sub>2</sub> to better focus on HCHO hotspots: East Asia (25–40°N, 110–125°E), Eastern US (30–40°N, 75–90°W), Northern Africa (0–15°N, 15°W–25°E) and Indonesia (5°S–5°N, 100–120°E). Negative satellite retrieved values over Eastern US and East-Asia are due to the limited amount of data during Northern Hemisphere winter months for these regions.

The time series in Fig. 6.2.1 show different cases for HCHO: regions dominated by biogenic emissions with some anthropogenic input (East Asia, Eastern US) and regions with both biogenic and pyrogenic sources (North-Africa and Indonesia). The CAMS reanalysis reproduces satellite observations for Eastern US with respect to absolute values and seasonality. The seasonality over East Asia is generally underestimated by the models (the MACC reanalysis performed better than the CAMS reanalysis here). The CAMS reanalysis shows a positive offset compared to satellite retrievals for North Africa (where the MACC reanalysis is closer to the retrievals) and Indonesia. The seasonality is overestimated for North Africa, but in agreement with the retrievals for Indonesia. The slight increase in satellite retrieved values for North Africa is not present in the model data so that simulations and observations are in better agreement for more recent years. For September and October 2015 over Indonesia, satellite retrievals and simulations show a pronounced maximum, which is however strongly overestimated by the models. SON 2015 were strong El Niño months (e.g. ) which caused droughts and higher fire activity in Indonesia. Fire emissions used by CIFS-rea seem to be largely overestimated for this El Niño year, resulting in an overestimation of about a factor of 1.8 compared to the observations. This overestimation was also reported for CAMS o-suite NRT data in the validation reports and it was also shown that it is not due to cloud flagging applied to the satellite and model data. It should be noted that weaker El Niño also occurred during SON 2006, 2009 and 2019 which resulted in maxima in the satellite retrieved values which are reproduced for 2006 but again overestimated for 2009 and especially for 2019.

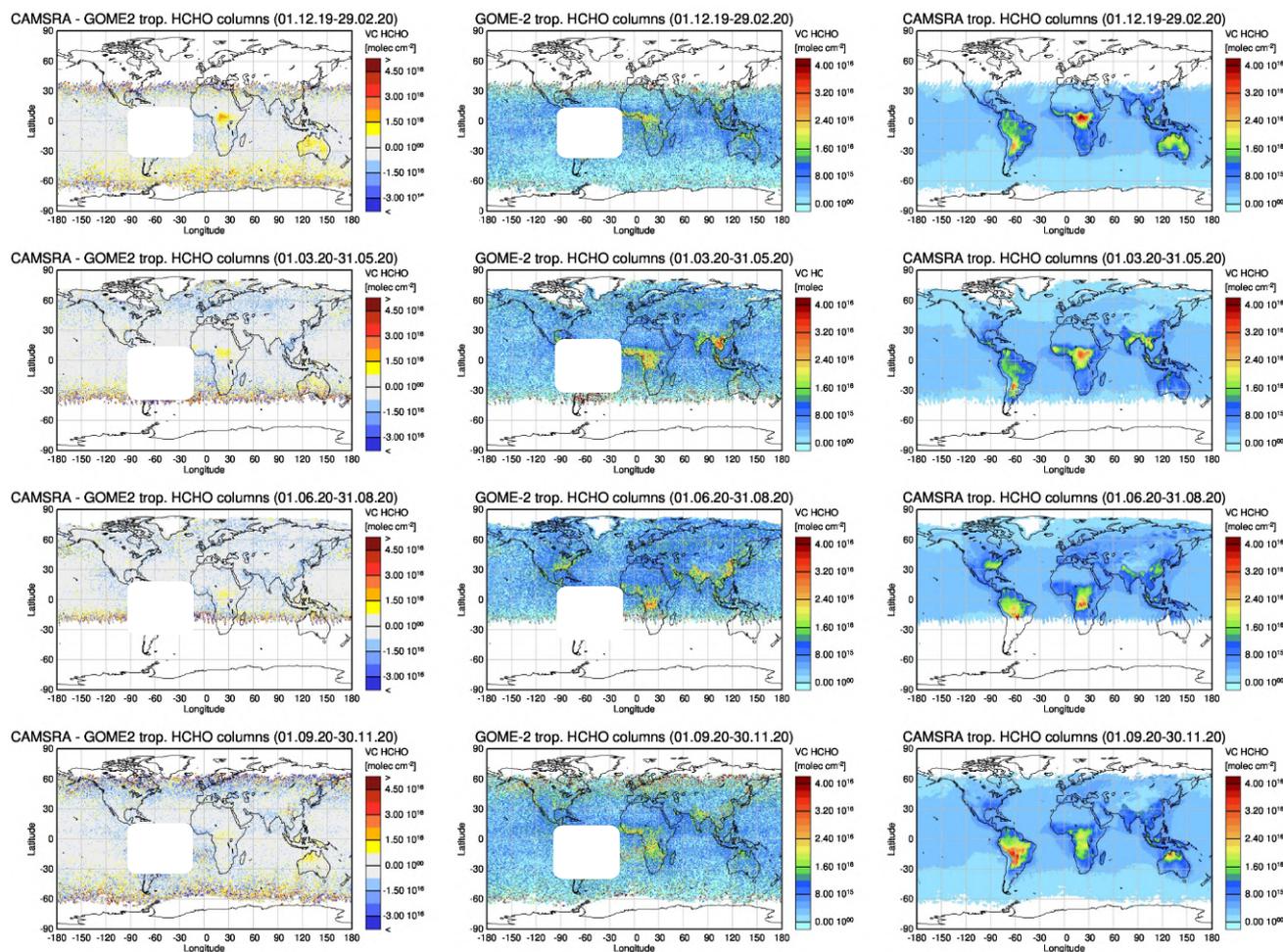


Figure 6.2.2: Global map comparisons of satellite retrieved and model simulated tropospheric HCHO columns [molec. cm<sup>-2</sup>] for (from top to bottom): DJF 2019/2020, MAM 2020, JJA 2020 and SON 2020. Shown are monthly mean tropospheric HCHO columns for (from left to right): GOME-2A, CAMS reanalysis and the difference between CAMS reanalysis and GOME-2A. GOME-2A data were gridded to model resolution (i.e. 0.75° x 0.75°). Model data were treated with the same reference sector subtraction approach as the satellite data. Satellite retrieved values in the region of the South Atlantic anomaly are not valid and therefore masked out (white boxes in all satellite images). Noise at higher latitudes is an artefact of the satellite retrievals.

Global maps of monthly means (see Fig. 6.2.2 for an example of seasonal averages for 2020) show that the magnitude of oceanic and continental background values and the overall spatial distribution are well represented by the CAMS reanalysis. Compared to SCIAMACHY and GOME-2A satellite retrievals, there is an overestimation of values for Central Africa during the whole season as well as Northern Australia which is most pronounced during autumn and winter. Values over

Europe are underestimated during spring, and there is an overestimation of values for boreal forest fires e.g., over Siberia in August 2020 (not shown). In general, tropospheric NO<sub>2</sub> is also overestimated over the same regions, but larger as well as additional areas of overestimation tend to occur compared to HCHO (see section 6.1, where in contrast to HCHO the overestimation is also visible in the summer seasonal average shown in Fig. 6.1.2).



## 7. Comparisons with aircraft campaigns

### 7.1 Campaign description

Measurements from several aircraft campaigns have been used to validate the CAMS reanalysis and control runs from 2003 to 2018. The campaign information is listed in Tab. 7.1.1, Tab. 7.1.2 (HIPPO) and 7.1.3 (ATOM). The flight tracks associated to campaigns described in Tab. 7.1.1 and 7.1.2 are shown in Fig. 7.1.1.

For comparing the reanalysis and control with the campaign data, the model grids nearest the measurement locations (latitude, longitude, and pressure layer) and times are extracted from the 4D field. For the spatial evaluation, all the airplane measurements and the extracted model data points are combined regardless of the measurement time. The HIPPO campaigns that flew over Pacific Ocean include five missions at different times of the year and can be used to evaluate the reanalysis as a function of latitudes and season. The HIPPO campaigns are used to evaluate O<sub>3</sub> and CO only (see Tab. 7.1.2). The ATOM campaign is used to evaluate several chemical species as indicated in Tab.7.1.3.

It should be noted that in this report the comparisons for control run were carried out only for ozone and CO for the campaigns listed in Table 7.1.1 and for all other available species in the case of the ATOM campaigns. Hereafter, the comparisons with the campaigns listed in Tab. 7.1.1 and 1.1.2 are presented in section 7.2 and the comparisons with the ATOM campaign in section 7.3.



Table 7.1.1 List of used aircraft campaigns

Campaign	Date	Location	Species used
INTEX-NA	2004.07-08	East America	O <sub>3</sub> , CO, NO, NO <sub>2</sub> , OH, HO <sub>2</sub> , HCHO
<a href="https://www-air.larc.nasa.gov/missions/intexna/intexna.htm">https://www-air.larc.nasa.gov/missions/intexna/intexna.htm</a>			
NEAQS-ITCT	2004.07-08	East America	O <sub>3</sub> , CO, NO, NO <sub>2</sub> , OH
<a href="https://www.esrl.noaa.gov/csd/projects/2004/">https://www.esrl.noaa.gov/csd/projects/2004/</a>			
ITOP	2004.07-08	North Atlantic	O <sub>3</sub> , CO
<a href="http://artefacts.ceda.ac.uk/badc_datadocs/itop/itop.html">http://artefacts.ceda.ac.uk/badc_datadocs/itop/itop.html</a>			
INTEX-B	2006.03-05	West America	O <sub>3</sub> , CO, NO, NO <sub>2</sub> , OH, HO <sub>2</sub> , HCHO
<a href="https://www-air.larc.nasa.gov/missions/intex-b/intexb.html">https://www-air.larc.nasa.gov/missions/intex-b/intexb.html</a>			
AMMA	2006.07-08	West Africa	O <sub>3</sub> , CO
<a href="http://artefacts.ceda.ac.uk/badc_datadocs/amma/amma.html">http://artefacts.ceda.ac.uk/badc_datadocs/amma/amma.html</a>			
ARCTAS	2008.04-07	North America to Arctic	O <sub>3</sub> , CO, NO, NO <sub>2</sub> , OH, HO <sub>2</sub> , HCHO
<a href="https://www-air.larc.nasa.gov/missions/arctas/arctas.html">https://www-air.larc.nasa.gov/missions/arctas/arctas.html</a>			
VOCALS	2008.10-11	Chile	O <sub>3</sub> , CO
<a href="http://data.eol.ucar.edu/master_list/?project=VOCALS">http://data.eol.ucar.edu/master_list/?project=VOCALS</a>			
YAK-AEROSIB	2006-2008, 2014	Russia	O <sub>3</sub> , CO
<a href="https://yak-aerosib.lsce.ipsl.fr/doku.php">https://yak-aerosib.lsce.ipsl.fr/doku.php</a>			
HIPPO	2009-2011	Pacific	O <sub>3</sub> , CO
<a href="https://hippo.ornl.gov/data_access">https://hippo.ornl.gov/data_access</a>			
KORUS-AQ	2016.04-06	Korea	O <sub>3</sub> , CO, NO, NO <sub>2</sub> , OH, HO <sub>2</sub> , HCHO
<a href="https://www-air.larc.nasa.gov/missions/korus-aq/index.html">https://www-air.larc.nasa.gov/missions/korus-aq/index.html</a>			
ATOM-1	2016.07-08	Pacific and Atlantic	O <sub>3</sub> , CO, NO, NO <sub>2</sub> , OH, HO <sub>2</sub> , HCHO
<a href="https://espo.nasa.gov/atom/archive/browse/atom/id3">https://espo.nasa.gov/atom/archive/browse/atom/id3</a>			

Table 7.1.2. List of HIPPO missions ([https://www.eol.ucar.edu/field\\_projects/hippo](https://www.eol.ucar.edu/field_projects/hippo)).

Mission	Date	latitudes
HIPPO-1	2009.01.09 – 2009.01.30	67°S – 80°N
HIPPO-2	2009.10.31 – 2009.11.22	66°S – 83°N
HIPPO-3	2010.03.24 – 2010.04.15	66.8°S – 85°N
HIPPO-4	2011.06.14 – 2011.07.10	58°S – 84°N
HIPPO-5	2011.08.09 – 2011.09.08	67°S – 87°N

Table 7.1.3 List of aircraft campaigns ATOM (<https://espo.nasa.gov/atom/archive/browse/atom/id3>).

Campaign	Date	Location	Species used
ATOM-1	29.07-23.08.2016	Pacific and Atlantic	O <sub>3</sub> , CO, NO, NO <sub>2</sub> , OH, HCHO
ATOM-2	26.01-21.02.2017	Pacific and Atlantic	O <sub>3</sub> , CO, NO, NO <sub>2</sub> , OH, HCHO
ATOM-3	28.09-27.10.2017	Pacific and Atlantic	O <sub>3</sub> , CO, NO, NO <sub>2</sub> , OH, HCHO
ATOM-4	24.04-21.05.2018	Pacific and Atlantic	O <sub>3</sub> , CO, NO, NO <sub>2</sub> , OH, HCHO

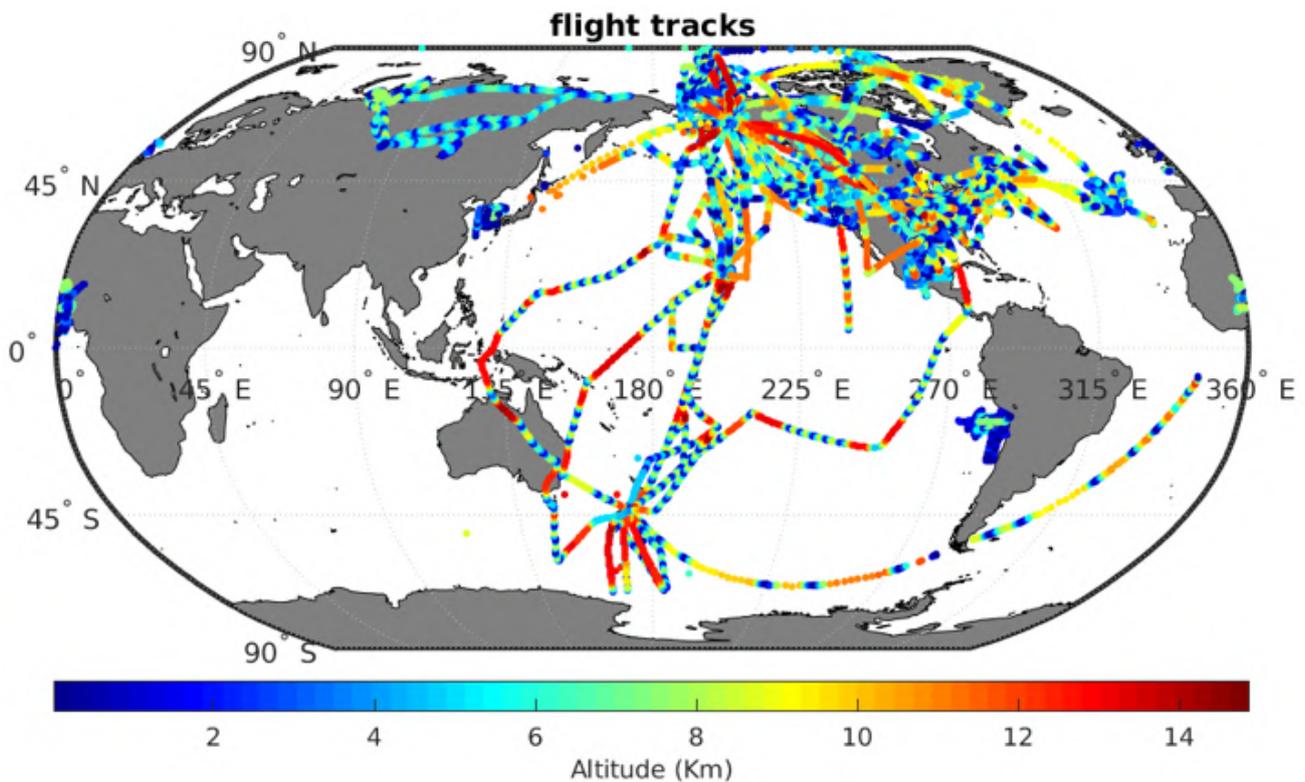


Figure 7.1.1: Flight tracks of the campaigns included in Table 7.1.1 and 7.1.2.

## 7.2 Comparison with several campaigns including HIPPO

### Ozone

The  $O_3$  comparison between observations and the CAMS reanalysis is generally good. The correlation coefficient squared of all the data points is 0.94, and the simulated slope is 0.94. The spatial distributions of the comparison at different altitude ranges are shown in Fig. 7.2.1. The reanalysis mostly overestimates  $O_3$  compared to the observations. In the surface layer (0-3Km), the biases of the reanalysis are mostly less than 15%, and the bigger differences occur in the tropics and Arctic, where the reanalysis overestimates the measurements by about 30%. In the free troposphere (3-9Km), the biases of the reanalysis are larger than in the lowest kilometers, especially over the tropical Pacific Ocean. Differences are small in the West Africa region. For the comparison above 9 km, the biases over the Pacific Ocean are even larger, up to 50%; however, the ozone concentrations are low in that region, so that the absolute biases are small. On the other hand, the absolute biases of the reanalysis in the polar stratosphere are large, but the relative differences are smaller because the  $O_3$  concentrations are high in that region.

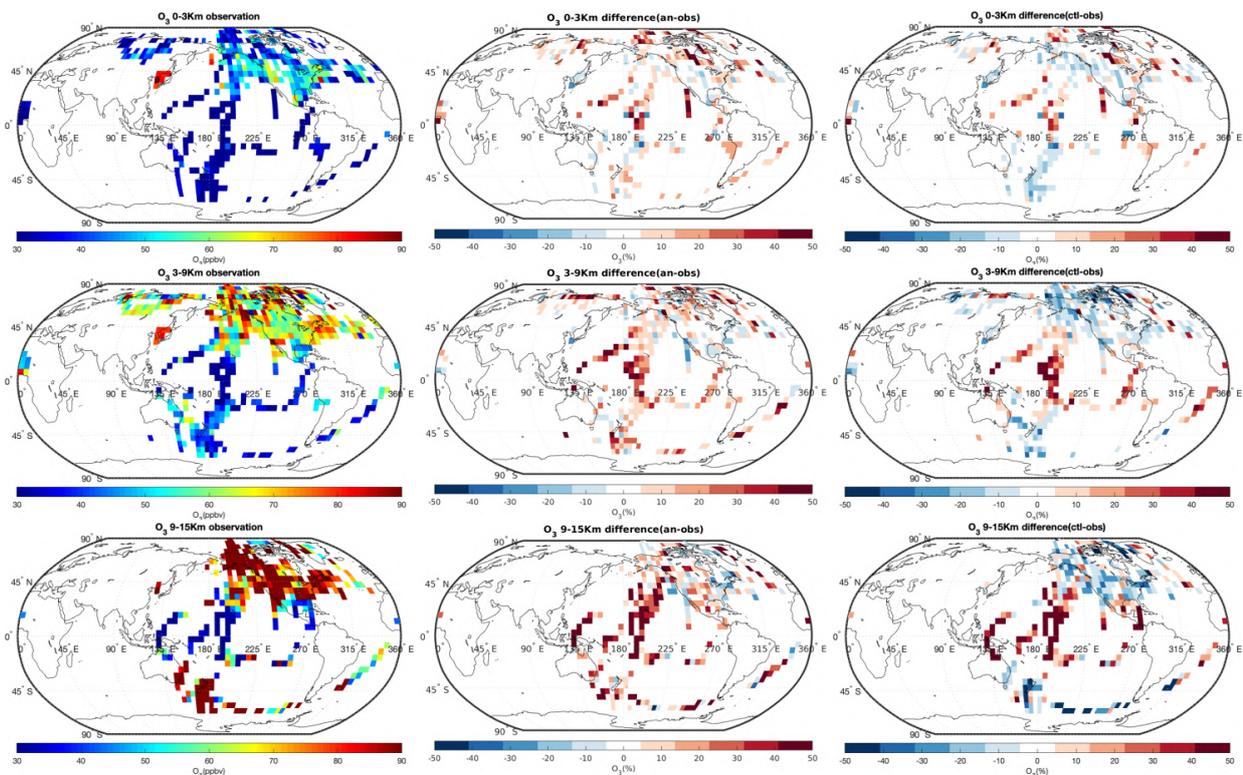


Figure 7.2.1: The campaign observations of  $O_3$  in ppbv (left, see table 7.1.1 and 7.1.2), the difference between the reanalysis and the observation (difference = reanalysis – observation; centre), and the relative difference in % between the control run and the observation (difference = control – observation; right). The data are averaged to  $5^\circ \times 5^\circ$  (latitude x longitude) and to three altitude bins: 0-3 km (top), 3-9 km (middle), and 9-15 km (bottom).

The agreement between the control run and the measurements is generally good too, but the overall bias is slightly larger than for the reanalysis. The calculated squared correlation coefficient is 0.90. In the surface layer, the biases of the control run are mostly about 15% too, but the pattern is different from the reanalysis. The control run underestimates the measurements in the west coast of America and the southern hemisphere, where the reanalysis is higher than the observation. In the free troposphere, the control run is lower than the observation in the polar region with a bias of about 20%; while it overestimates the tropical region, similar to the reanalysis. In the upper layer, the biases are still negative over the Southern Ocean and increase to 30%. Different from the lower layers, the biases in America become negative too.

The comparison with five HIPPO campaigns is shown in Fig. 7.2.2. Over the tropical ocean and for the layers above 4 Km, the reanalysis overestimates the observation for all the missions in different seasons. In the lower layers of the tropical ocean, the reanalysis underestimates the measurements during the spring (HIPPO-3) and autumn (HIPPO-2 and HIPPO-5) campaigns. In the lower layers of the southern hemisphere, the reanalysis overestimates the  $O_3$  concentration during the HIPPO-5 campaign of about 20%, which is in late summer/early autumn, while it underestimates  $O_3$  for the other missions with a small bias of less than 10%. The reanalysis is higher than the measurements in the stratosphere of the southern high latitudes for all the missions, and the biases are biggest during the summer campaign (HIPPO-4). In the northern hemisphere, the reanalysis is higher than

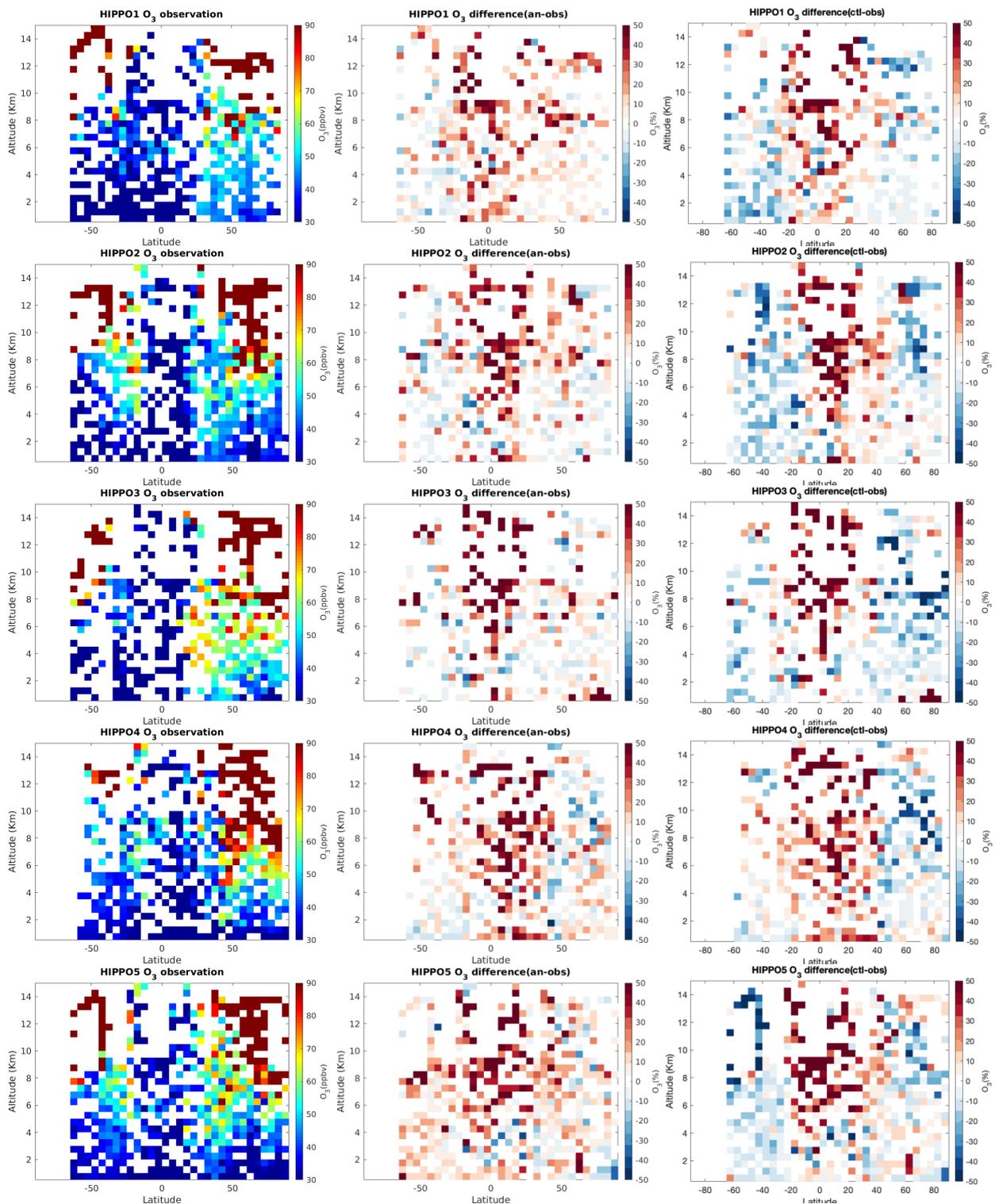


Figure 7.2.2: The HIPPO observation of O<sub>3</sub> in ppbv (left, see table 7.1.2), the difference between the reanalysis and the observation (difference = reanalysis – observation; centre), and the difference between the control run and the observation (right) against latitudes and altitudes. From top to bottom results from the 5 HIPPO campaigns (table 7.1.2) are shown. The data points are averaged to 5° (latitude) and 0.5 Km (altitude) bins.



the measurements during the winter campaign (HIPPO-1). During the autumn campaigns (HIPPO-2 and HIPPO-5), the biases are negative in the lower layers, but positive in the upper layers. The pattern is the opposite for the spring (HIPPO-3) and summer (HIPPO-4) campaign.

Over the tropical ocean, the control run is generally consistent with the reanalysis and the biases of the control run in that region (50%) are slightly larger than those of the reanalysis (40%). In the higher latitudes, the reanalysis and the control run behave differently. The control run underestimates the O<sub>3</sub> by about 20% in the southern hemisphere for all the missions except the mid-troposphere during the summer mission and some grids in the stratosphere. In the northern high latitudes, the control run underestimates the O<sub>3</sub> concentration in the free troposphere for all the missions, and it overestimates the measurements in the surface layer except some grids in HIPPO-5 and HIPPO-1. The control run overestimate the O<sub>3</sub> in the stratosphere from spring to early autumn (HIPPO-3, HIPPO-4, and HIPPO-5).

### **Carbon monoxide**

The agreement of CO between the measurements and the reanalysis is good with a squared correlation coefficient of 0.71 and a slope of 0.79. The comparison of all the campaigns at different layers is shown in Fig. 7.2.3. The biases of the most averaged grids are smaller than 15%. In the surface layer, the reanalysis is lower than the measurements over the north Pacific Ocean, while it is higher than the measurements in North America, north Atlantic Ocean, north Russia, and the southern hemisphere. In the free troposphere, the reanalysis underestimates CO over most of the northern hemisphere and the tropical ocean, and the positive biases are mainly over the Southern Ocean. In the upper layer, the biases are positive in the polar region, while negative in the lower latitudes.

The comparison of the control run is slightly worse than of the reanalysis. The calculated correlation coefficient squared is 0.66. The bias pattern of the control run in the surface layer is similar to the reanalysis, but the positive biases in the southern hemisphere are bigger than those of the reanalysis, which are about 30%. In the free troposphere, the control run underestimates the CO concentration at the latitudes north than 40°N, consistent with the reanalysis, but overestimates the CO in the rest of the globe. In the upper layer, the biases are positive in most of the region except in the east of America, and the biases are large in the polar stratosphere, which is about 50%.

The comparison with HIPPO (Fig. 7.2.4) shows that the biases of the reanalysis in the tropics are negative throughout the year; the biases are positive in the lower layers of the southern hemisphere; the biases are positive in the upper layers of the northern high latitudes. In the winter (HIPPO-1)/late autumn (HIPPO-2) time, the reanalysis is higher than the measurements in the upper layers of the southern high latitudes, while it is lower in the spring to early autumn time (HIPPO-3, HIPPO-4, and HIPPO-5). In the surface layer of the Arctic region, the reanalysis underestimates the CO concentration in spring-time when CO peaks (HIPPO-3), but overestimates CO in summer (HIPPO-4) and early autumn (HIPPO-5) when the CO concentrations are relatively low in that region.

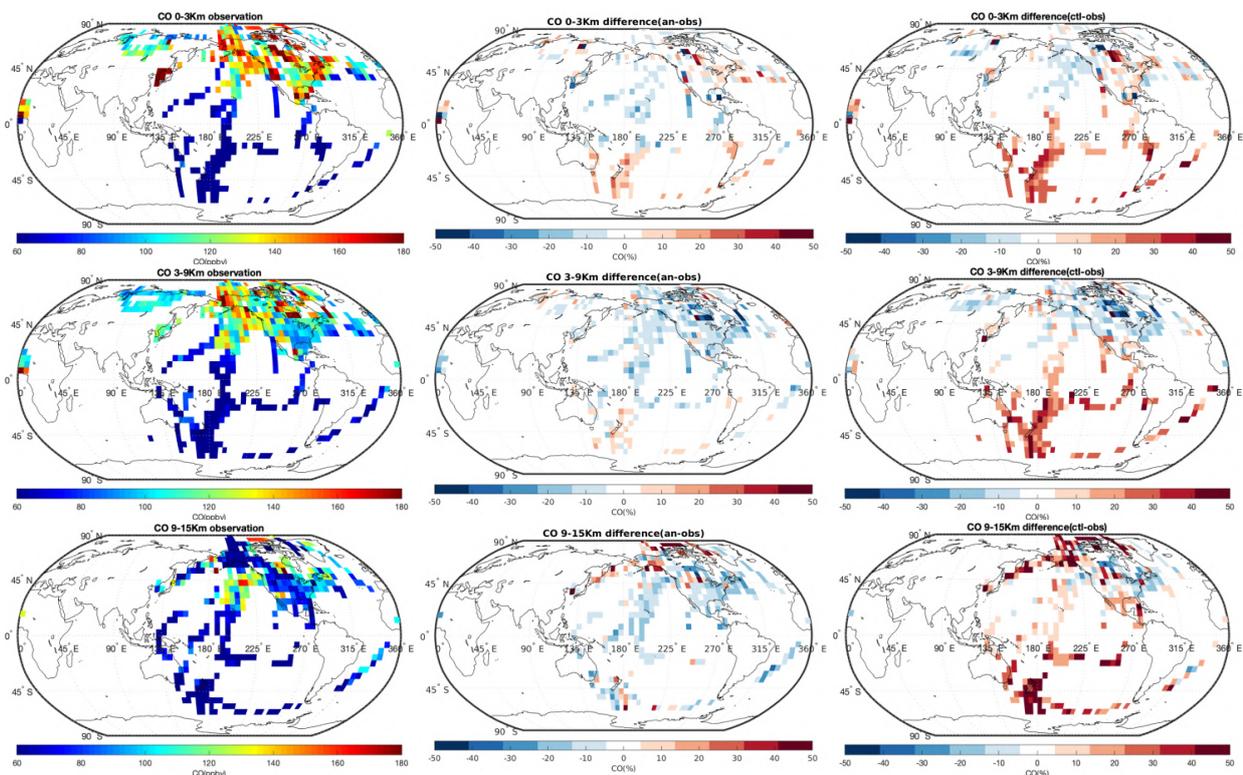


Figure 7.2.3: The campaign observation of CO in ppbv (left, see table 7.1.1 and 7.1.2), the difference between the reanalysis and the observation (difference = reanalysis – observation; centre), and the difference between the control run and the observation (difference = control – observation; right). The data are averaged to  $5^{\circ} \times 5^{\circ}$  (latitude x longitude) and to three altitude bins: 0-3 Km (top), 3-9 Km (middle), and 9-15Km (bottom).

The control run is higher than the measurements in the southern hemisphere all over the year, and the biases are larger in winter (HIPPO-1, 40%) and spring (HIPPO-3, 50%). The control run also shows overestimation in the high altitudes of the northern hemisphere. In the summer (HIPPO-4) to early autumn (HIPPO-5), the biases are positive in the northern low altitude, while the biases are negative in the other time of the year.

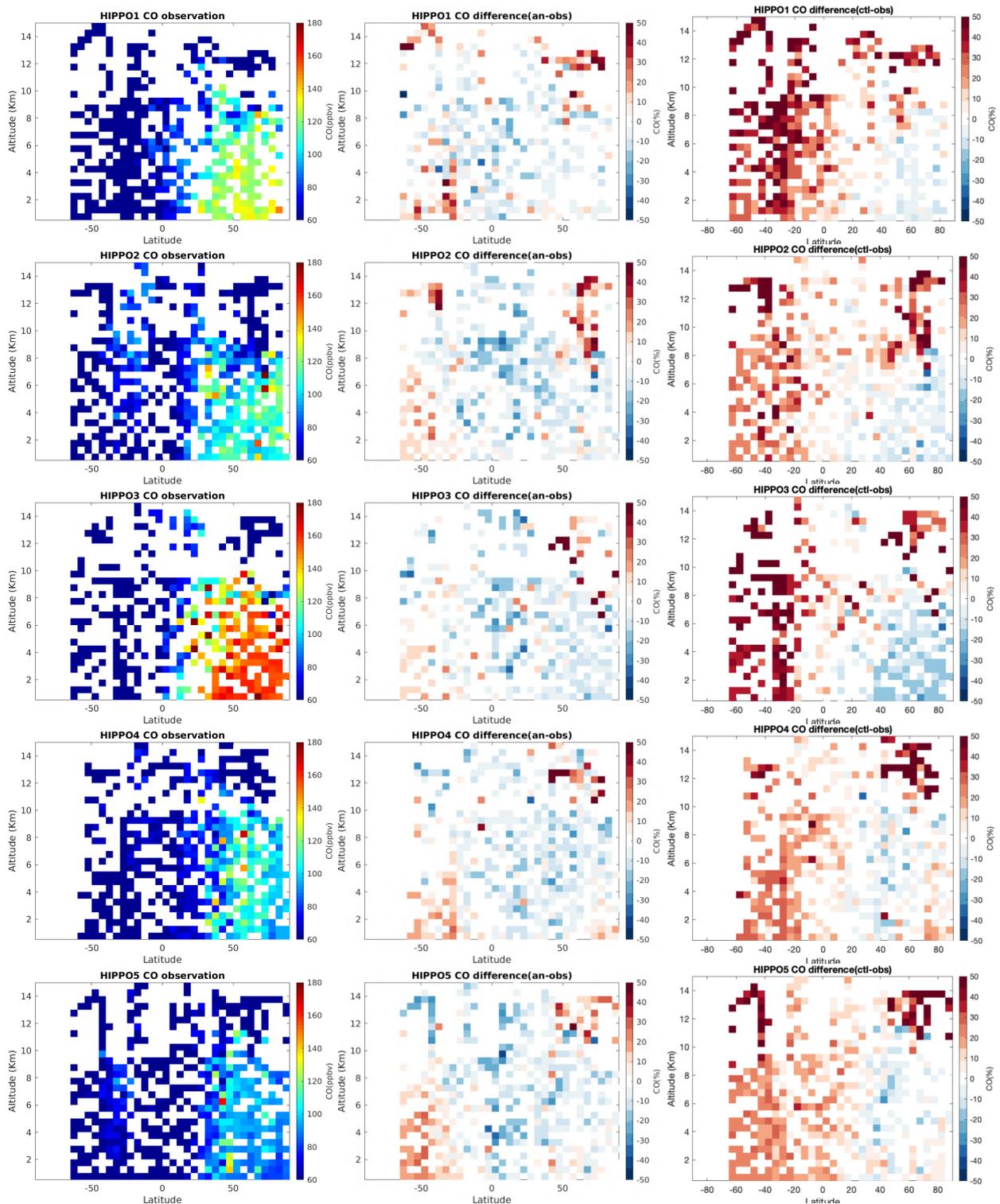


Figure 7.2.4: The HIPPO observation of CO in ppbv (left, see table 7.1.1 and 7.1.2), the difference between the reanalysis and the observation (difference = reanalysis – observation; centre), and the difference between the control run and the observation (difference = control – observation; right) against latitudes and altitudes. The data points are averaged to 5° (latitude) and 0.5 Km (altitude) bins.

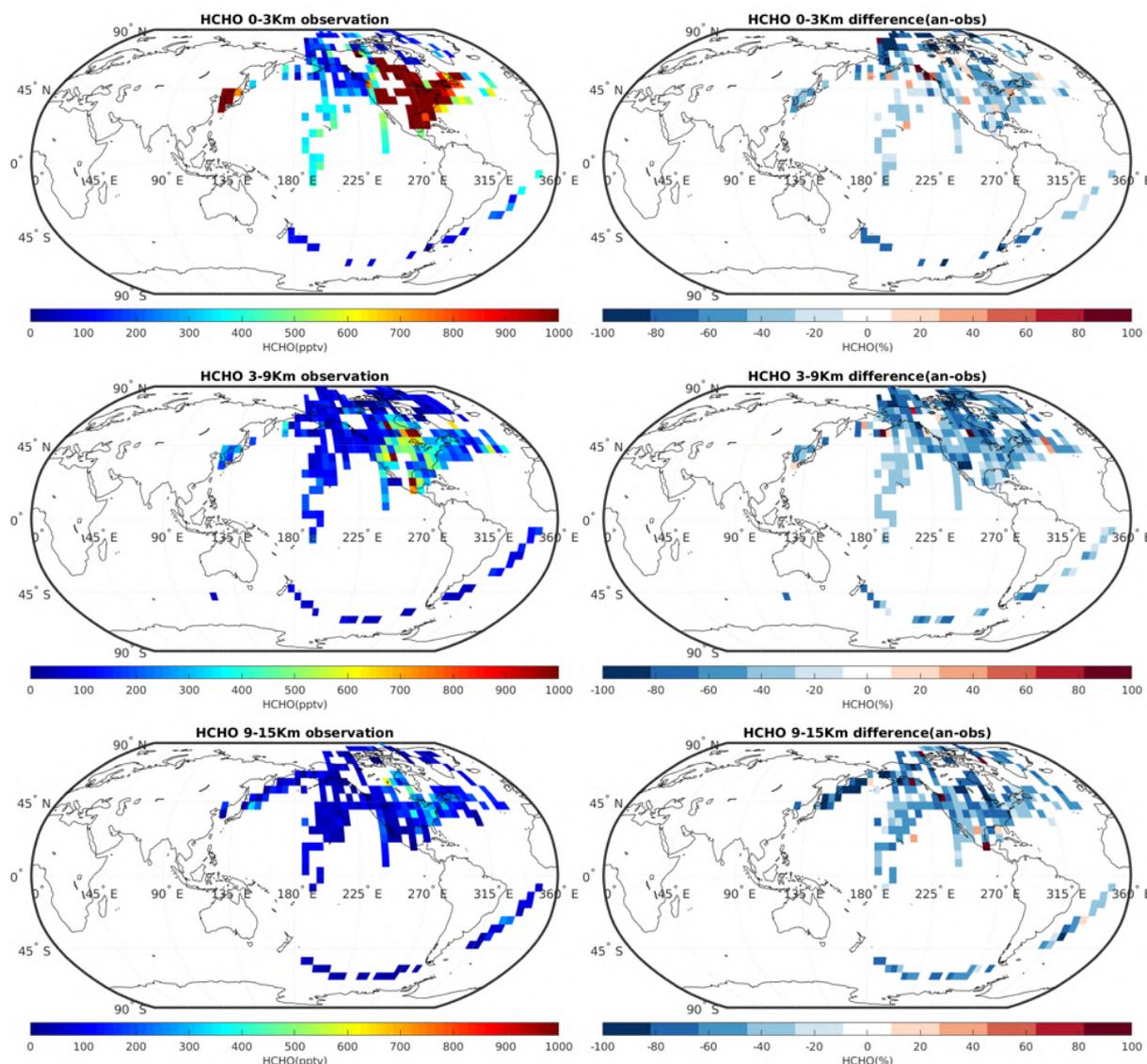


Figure 7.2.5: The campaign observation of HCHO in pptv (left, see table 7.1.1 and 7.1.2), the difference between the reanalysis and the observation (difference = reanalysis – observation; right). The data are averaged to 5°x5° (latitude x longitude) and to three altitude bins: 0-3 Km (top), 3-9 Km (middle), and 9-15Km (bottom).

### Formaldehyde

The comparisons between the reanalysis and measurements for HCHO are shown in Fig. 7.2.5. The calculated correlation coefficient for reanalysis and observation squared is 0.53, and the slope is 0.50. The included campaign measurements show that the HCHO concentrations are high in the lower altitudes in North America and Korea, and the reanalysis captures the pattern, so the bias distribution is generally uniform. The reanalysis underestimates HCHO when comparing to the campaign observation. The biases in the surface layer of the low and middle latitudes are about 30%, and the biases in the high latitudes are larger, about 60%. The biases in the upper layers are around 50% everywhere.

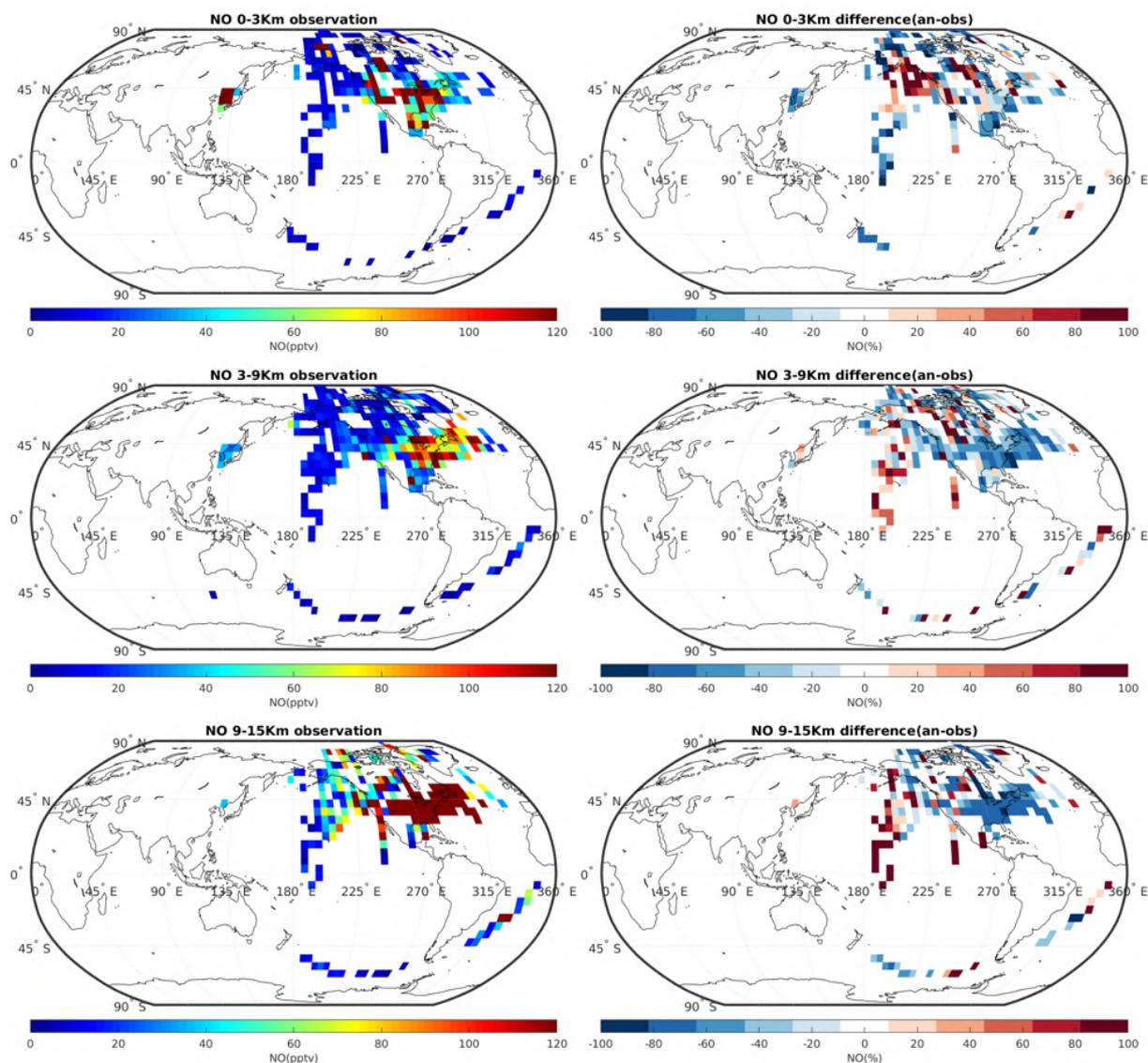


Figure 7.2.6: The campaign observation of NO in pptv (left, see table 7.1.1 and 7.1.2), the difference between the reanalysis and the observation (difference = reanalysis – observation; right). The data are averaged to 5x5 (latitude x longitude) and to three altitude bins: 0-3 Km (top), 3-9 Km (middle), and 9-15Km (bottom).

### Nitrogen oxides

The comparison of NO, NO<sub>2</sub>, and NO<sub>x</sub> (NO<sub>x</sub> = NO + NO<sub>2</sub>) are shown in Fig. 7.2.6, Fig. 7.2.7, and Fig. 7.2.8, respectively.

As shown in Fig. 7.2.6, in the surface layer, the reanalysis overestimates NO to the west of North America, while underestimating NO in the other regions. The mean bias for NO is about 70% for the averaged grids in this layer. In the free troposphere, the models are higher than the observation in some regions of Canada and Arctic and the north Pacific, while lower in the west coast of America, which is opposite to the surface layers. The bias pattern in the high altitudes is similar to the mid-troposphere, but the values are larger than the lower layers.

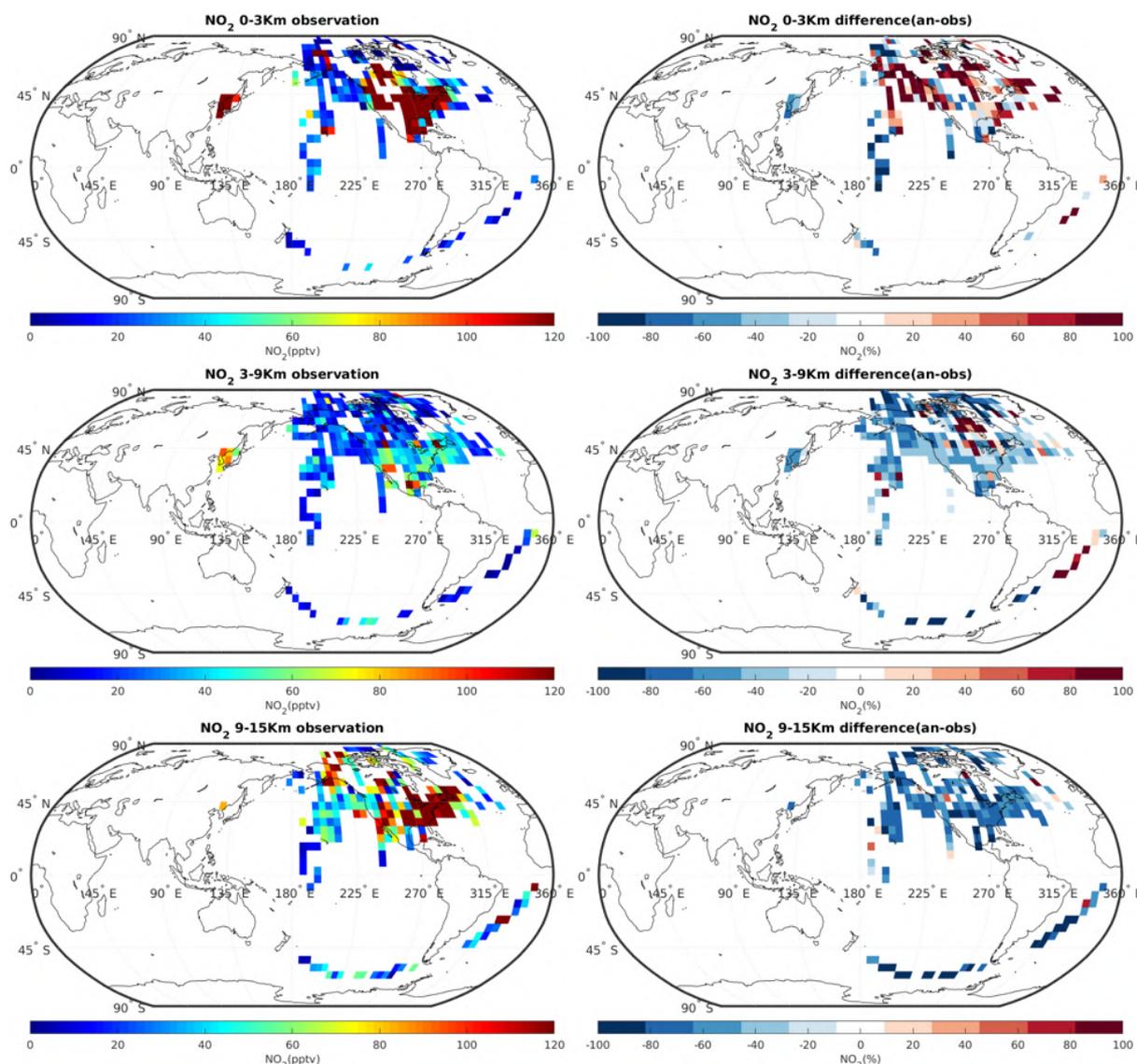


Figure 7.2.7: The campaign observation of NO<sub>2</sub> in pptv (left, see table 7.1.1 and 7.1.2), the difference between the reanalysis and the observation (difference = reanalysis – observation; right). The data are averaged to 5x5 (latitude x longitude) and to three altitude bins: 0-3 Km (top), 3-9 Km (middle), and 9-15Km (bottom).

The bias pattern for NO<sub>2</sub> is different from NO (Fig. 7.2.7). In the surface layer, the models overestimate NO<sub>2</sub> all over the North America as well as the northeast Pacific and the Atlantic. The models underestimate NO<sub>2</sub> over Korea, the middle Pacific and the Southern Ocean. The mean bias is about 120%. In the middle layer, the models are only higher than the measurements in Central Canada and the south Atlantic. In the other regions the CAMS reanalysis underestimates NO<sub>2</sub>. The biases of the CAMS reanalysis are negative at the high altitudes everywhere.

When combining NO and NO<sub>2</sub>, the agreement between the models and the measurements improves a little (see Fig. 7.2.8). The mean bias of NO<sub>x</sub> is about 70%.

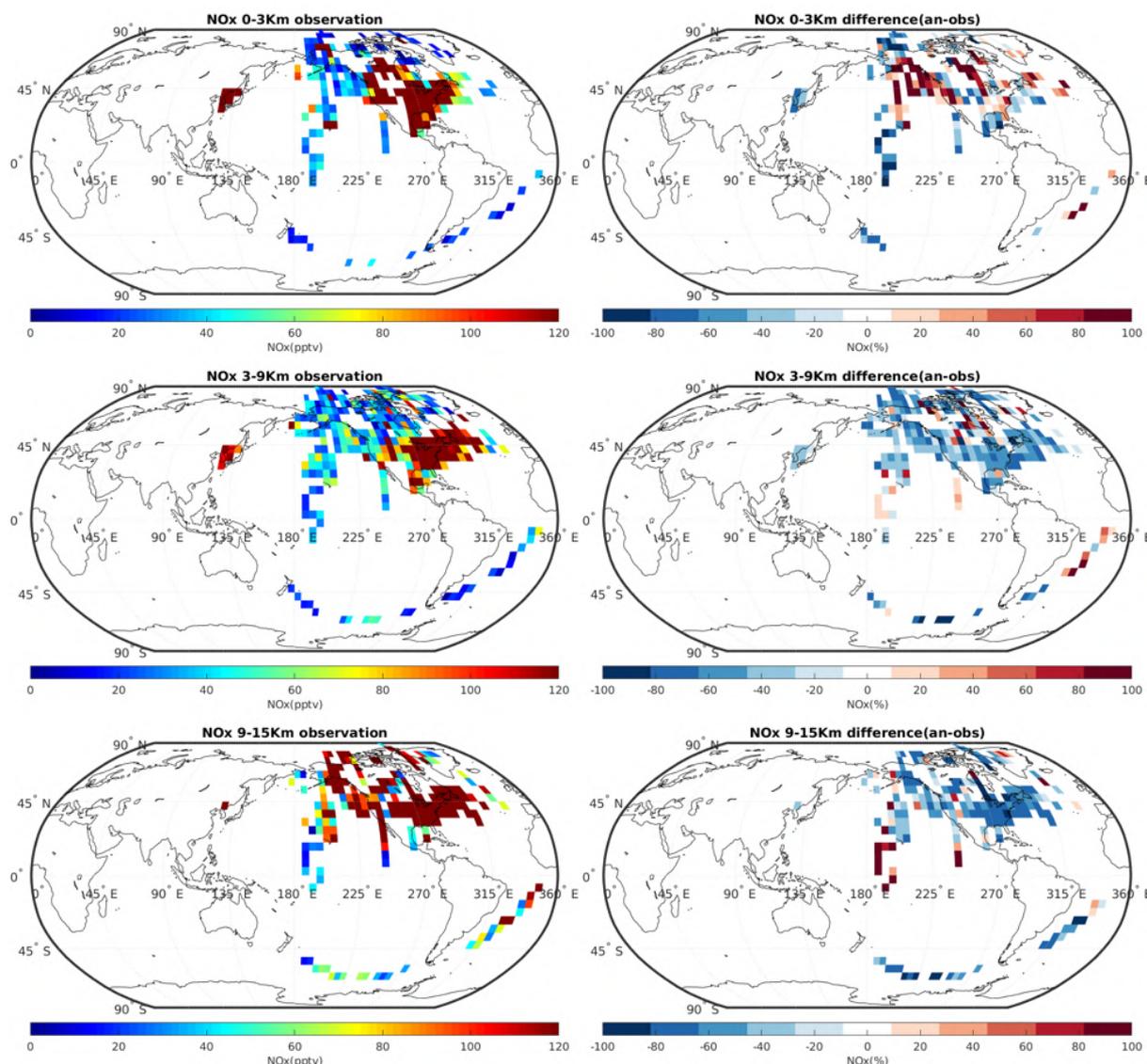


Figure 7.2.8: The campaign observation of NO<sub>x</sub> in pptv (left, see table 7.1.1 and 7.1.2), the difference between the reanalysis and the observation (difference = reanalysis – observation; right). The data are averaged to 5x5 (latitude x longitude) and to three altitude bins: 0-3 Km (top), 3-9 Km (middle), and 9-15Km (bottom).

### Hydroxyl and Hydroperoxyl radical

The reanalysis is compared to OH measurements and results are shown in Fig. 7.2.9. In the troposphere, the biases are negative in the Arctic and the Southern Ocean, and positive elsewhere. The mean bias is about 60%, but the sampling is limited, and this number cannot be easily generalised to a global mean. In the upper layers, the bias pattern is similar to the troposphere, but the biases are slightly smaller (50%).

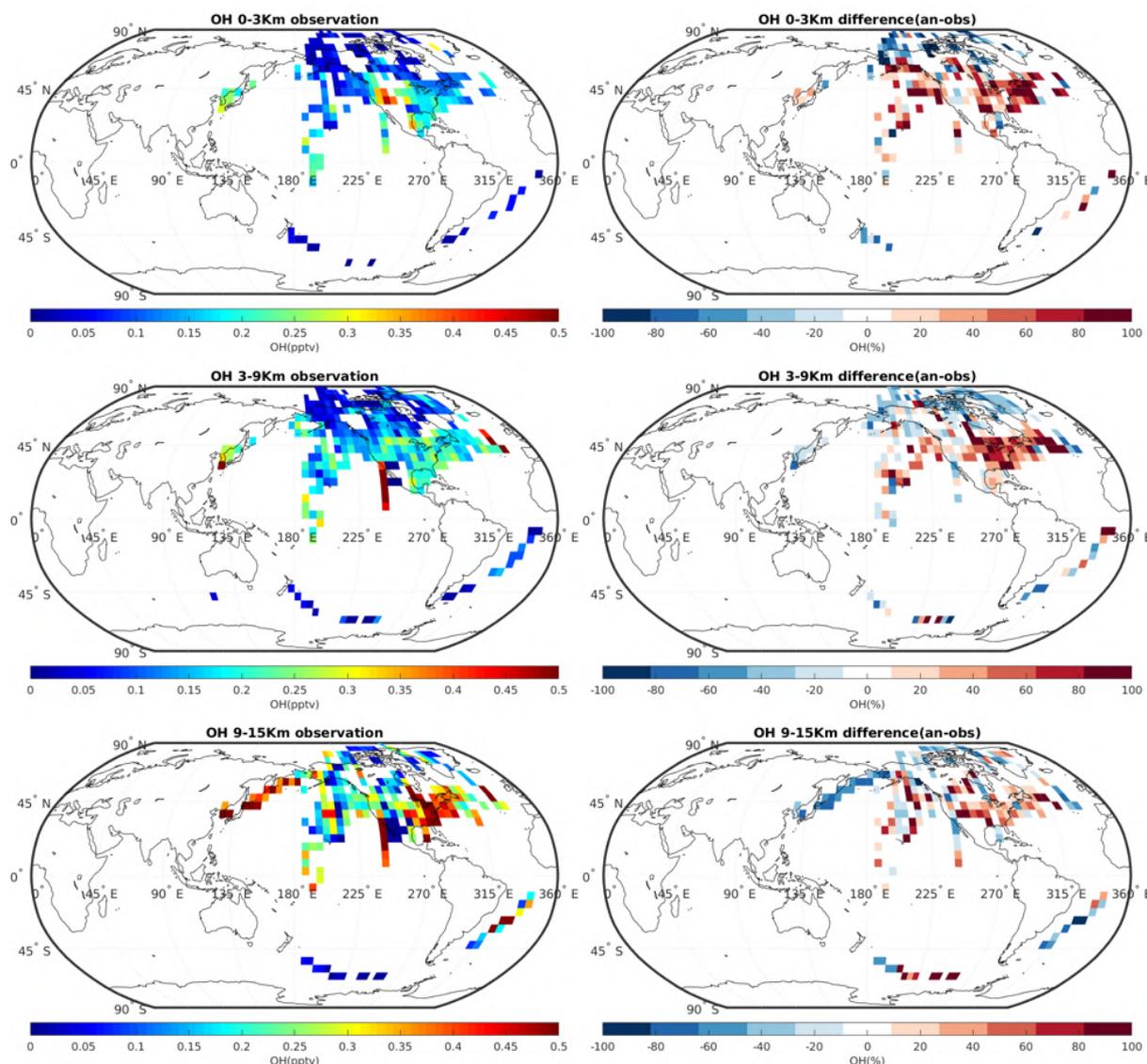


Figure 7.2.9: The campaign observation of OH in pptv (left, see table 7.1.1 and 7.1.2), the difference between the reanalysis and the observation (difference = reanalysis – observation; right). The data are averaged to 5x5 (latitude x longitude) and to three altitude bins: 0-3 Km (top), 3-9 Km (middle), and 9-15Km (bottom).

HO<sub>2</sub> is only evaluated for the reanalysis (see Fig. 7.2.10). The mean bias of HO<sub>2</sub> is about 40 % and smaller than the OH bias in the surface layer. The reanalysis is higher than the measurements in Korea and the east of America, and it is lower in the other regions. In the free troposphere, the bias pattern changes: the biases become positive in the Arctic, the middle of Pacific, and the Southern Ocean, but the biases change to negative over Korea. In the higher altitudes, the biases over the Pacific change back to negative again.

The calculated squared correlation coefficients are 0.44 for OH and 0.41 for HO<sub>2</sub>, respectively.

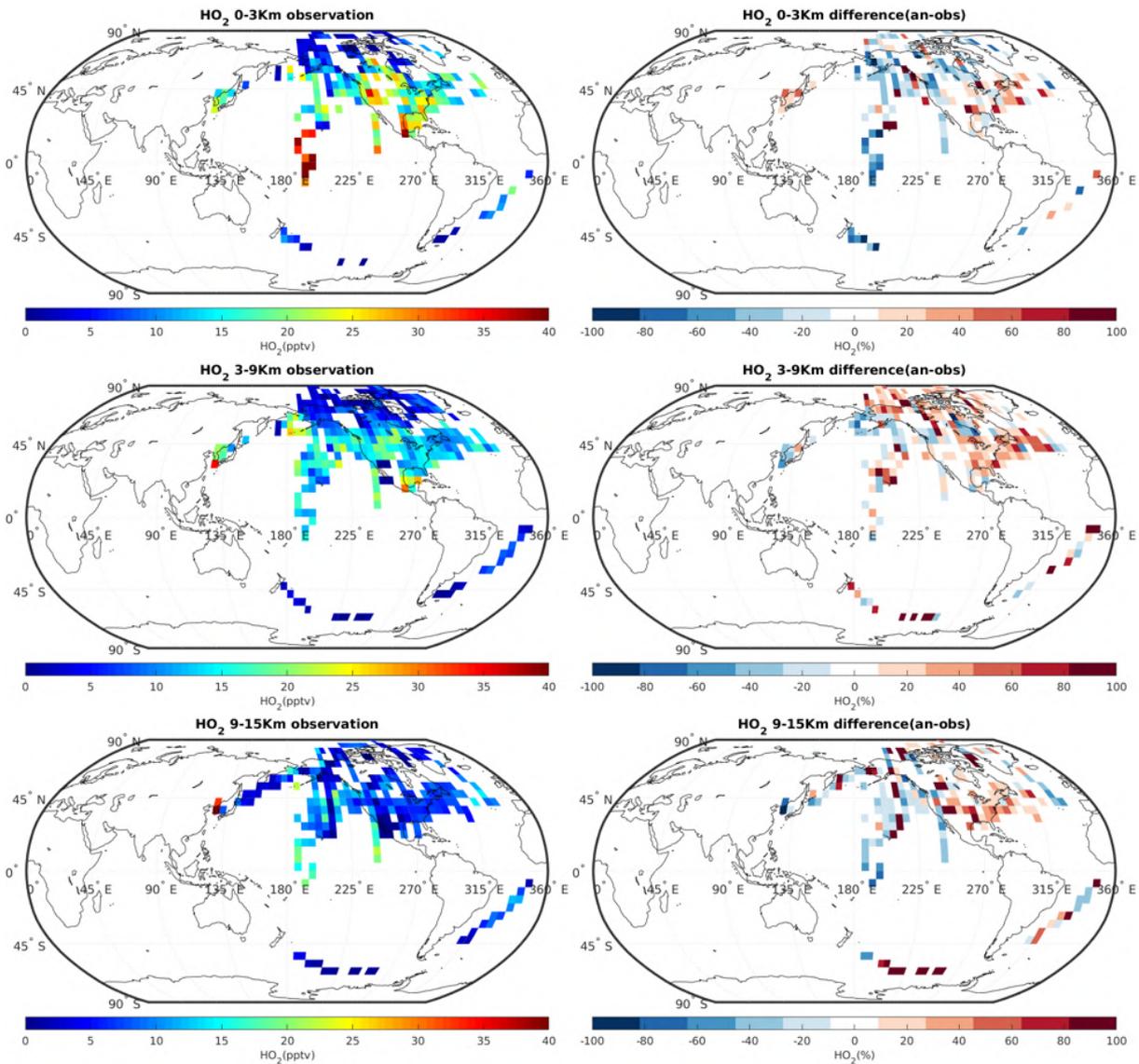


Figure 7.2.10: The campaign observation of HO<sub>2</sub> in pptv (left, see table 7.1.1 and 7.1.2) and the difference between the reanalysis and the observation (difference = reanalysis – observation; right). The data are averaged to 5x5 (latitude x longitude) and to three altitude bins: 0-3 Km (top), 3-9 Km (middle), and 9-15Km (bottom).

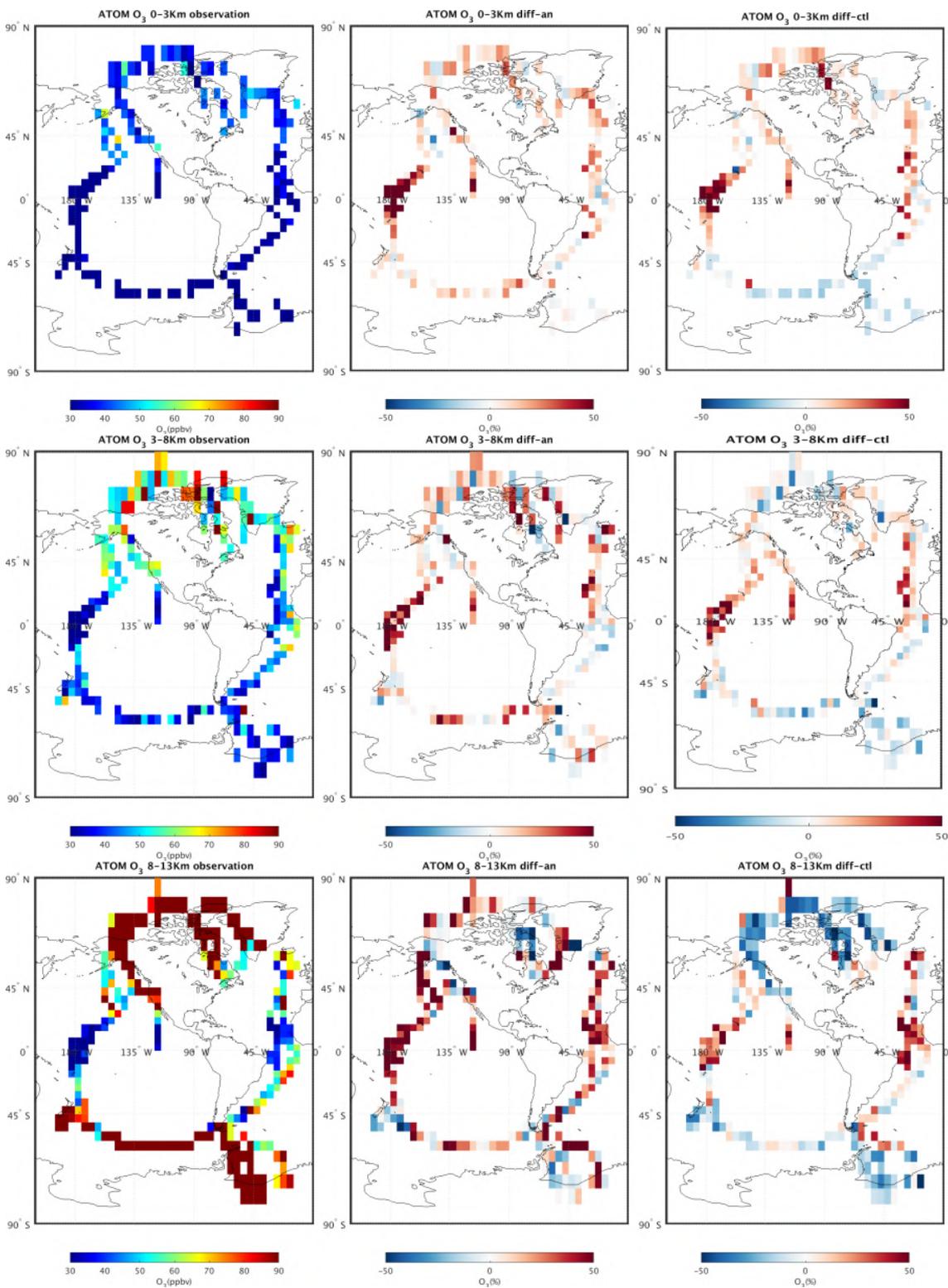


Figure 7.3.1: ATOM campaign observations of O<sub>3</sub> in ppbv (left, see table 7.1.1 and 7.1.2), the difference between the reanalysis and the observation (difference = reanalysis – observation; centre), and the difference between the control run and the observation (difference = control – observation; right). The data are averaged to 5°x5° (latitude x longitude) and to three altitude bins: 0-3 km (top), 3-9 km (middle), and 9-15 km (bottom).



### 7.3 Comparison with the ATOM campaign

#### Ozone

The reanalysis and control O<sub>3</sub> are generally in good agreement with the observations (Fig. 7.3.1). The reanalysis mostly overestimates O<sub>3</sub> compared to the observations. In the low altitudes (0-3 km), the bias is within 10% except for the low O<sub>3</sub> values over the tropical Pacific Ocean where the bias reaches 50%. The overestimation patterns and values for the control run (Fig. 7.3.1) are similar to the reanalysis over the Northern Hemisphere and different over the Southern Hemisphere showing underestimation within 10%. In the free troposphere (3-8km), the biases of the reanalysis and control runs are larger than in the lower layer. The biases in the upper layer (8-13km) are even larger, especially over the regions with low values over the Pacific Ocean and reach 50%. In contrast to the reanalysis, the control run shows negative bias (up to 50%) over the north polar region.

#### Carbon monoxide

The CO from the CAMS reanalysis shows a good agreement with the observations (Fig. 7.3.2). In the lower layer the bias is slightly positive over the Northern Hemisphere and slightly negative over the Southern Hemisphere (within 10%). The control run is similar over the Northern Hemisphere but shows a higher positive bias over the Southern Hemisphere (Fig. 7.3.2). In the free troposphere the bias pattern and values from both runs are similar compared to the results for the lowest layer, apart from the north polar region where positive biases reach 50%. In the upper layers over southern polar regions the positive biases are larger in both CAMS configurations.

#### Formaldehyde

The comparisons of the HCHO between the reanalysis and measurements, and between the control run and the measurements are very similar (Fig. 7.3.3), because there is no assimilation of HCHO data in the CAMS reanalysis. The agreements of both runs with the observations are reasonable. Both runs capture the high HCHO values over the Atlantic and Pacific oceans in the lower layer. In general, both runs underestimate the observations with larger biases in the upper layer.

#### Nitrogen oxides

As shown in Fig. 7.3.4, the NO in the reanalysis for all layers shows strong underestimation over the Atlantic Ocean (up to 100%). The bias changes from negative in the low layers to positive in the upper layer over the Pacific Ocean, except for southern polar regions where the bias is constantly negative. The control run shows different patterns, indicating negative NO bias over the Southern Hemisphere and positive bias over the Northern Hemisphere, except for negative bias in northern polar regions, for all the layers.

The bias pattern for NO<sub>2</sub> is different from NO (Fig. 7.3.5). For both reanalysis and control run, the biases are predominantly negative in all the layers, except for the Atlantic Ocean where biases are positive in the reanalysis and mostly negative in the control run. In the lower layers the bias is also positive west of North America in both runs.

Taking the sum of NO and NO<sub>2</sub> (NO<sub>x</sub>) slightly improves the agreement with the observations (Fig.7.3.6).

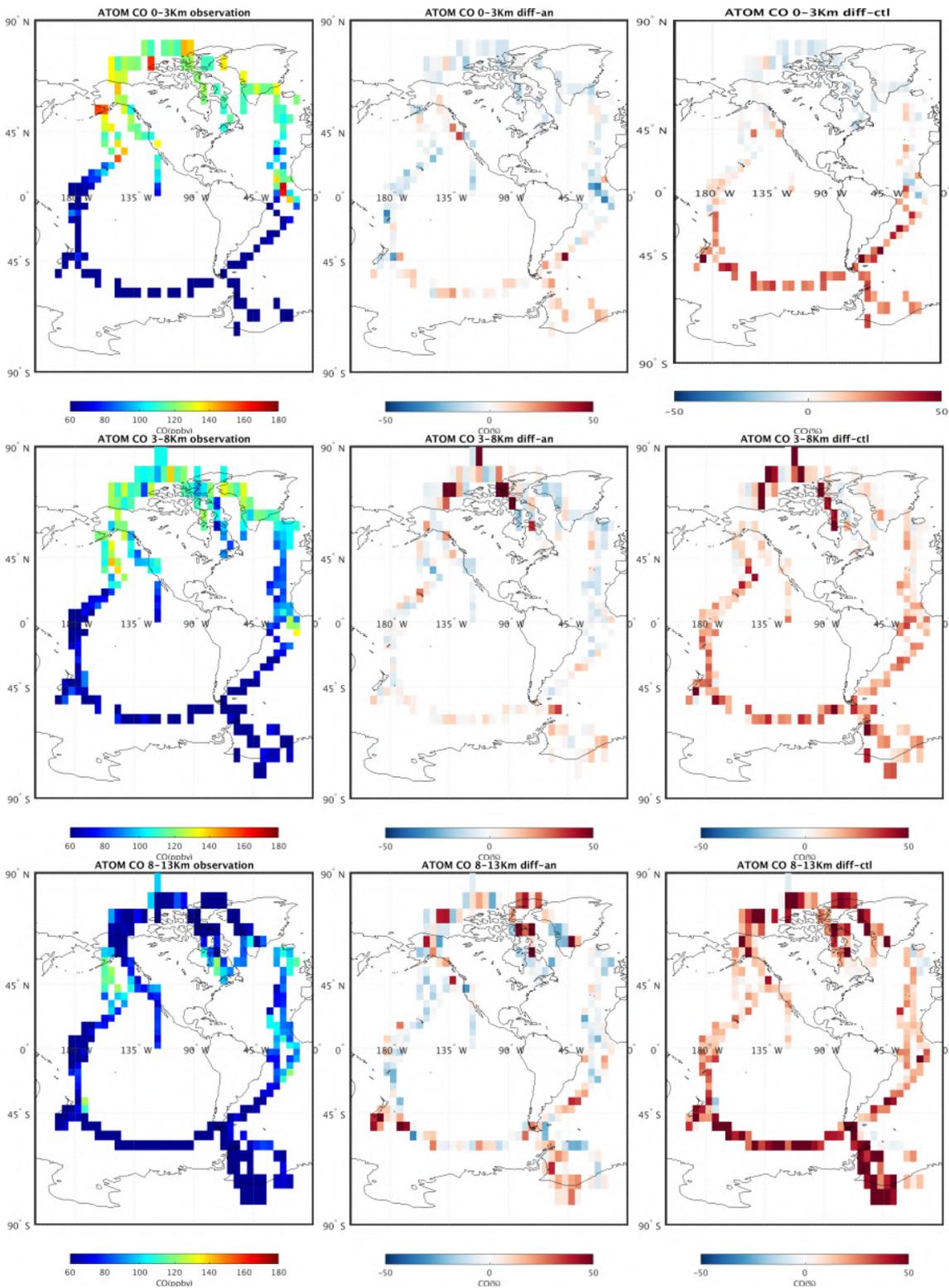


Figure 7.3.2: ATOM campaign observations of CO in ppbv (left, see table 7.1.1 and 7.1.2), the difference between the reanalysis and the observation (difference = reanalysis – observation; centre), and the difference between the control run and the observation (difference = control – observation; right). The data are averaged to 5°x5° (latitude x longitude) and to three altitude bins: 0-3 km (top), 3-9 km (middle), and 9-15 km (bottom).

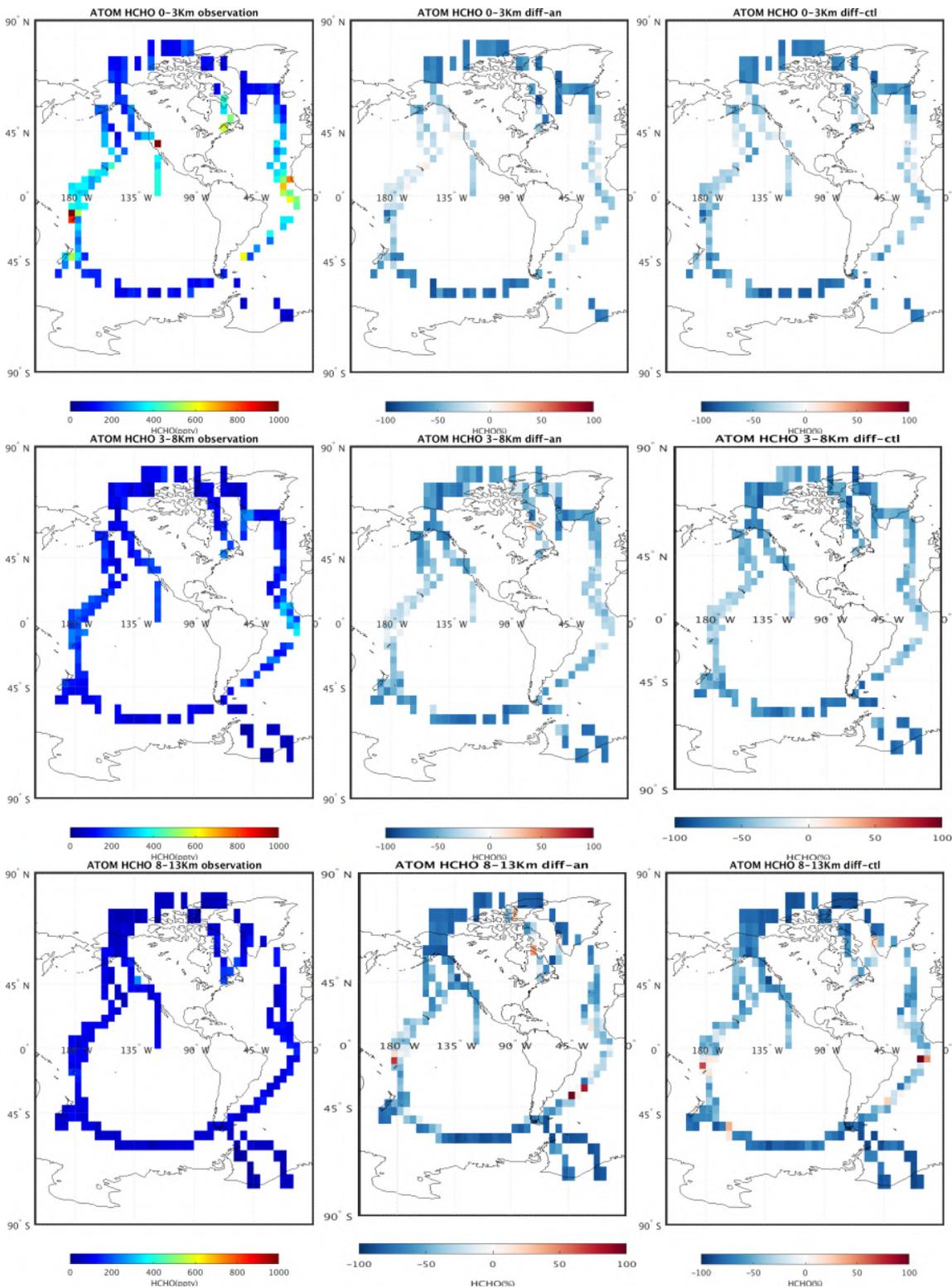


Figure 7.3.3: ATOM campaign observations of HCHO in pptv (left, see table 7.1.1 and 7.1.2), the difference between the reanalysis and the observation (difference = reanalysis – observation; centre), and the difference between the control run and the observation (difference = control – observation; right). The data are averaged to 5°x5° (latitude x longitude) and to three altitude bins: 0-3 km (top), 3-9 km (middle), and 9-15 km (bottom).

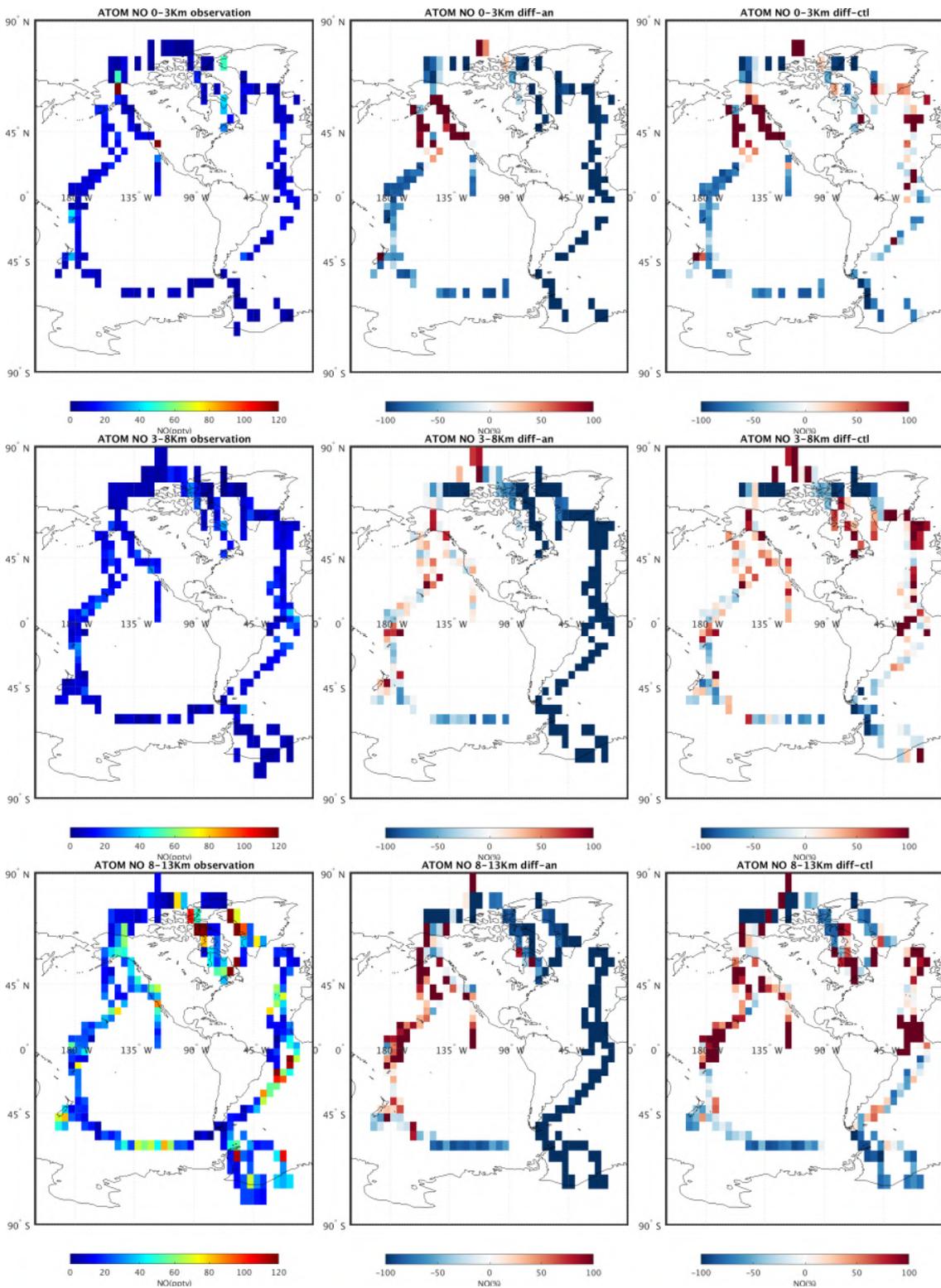


Figure 7.3.4: ATOM campaign observations of NO in pptv (left, see table 7.1.1 and 7.1.2), the difference between the reanalysis and the observation (difference = reanalysis – observation; centre), and the difference between the control run and the observation (difference = control – observation; right). The data are averaged to 5°x5° (latitude x longitude) and to three altitude bins: 0-3 km (top), 3-9 km (middle), and 9-15 km (bottom).

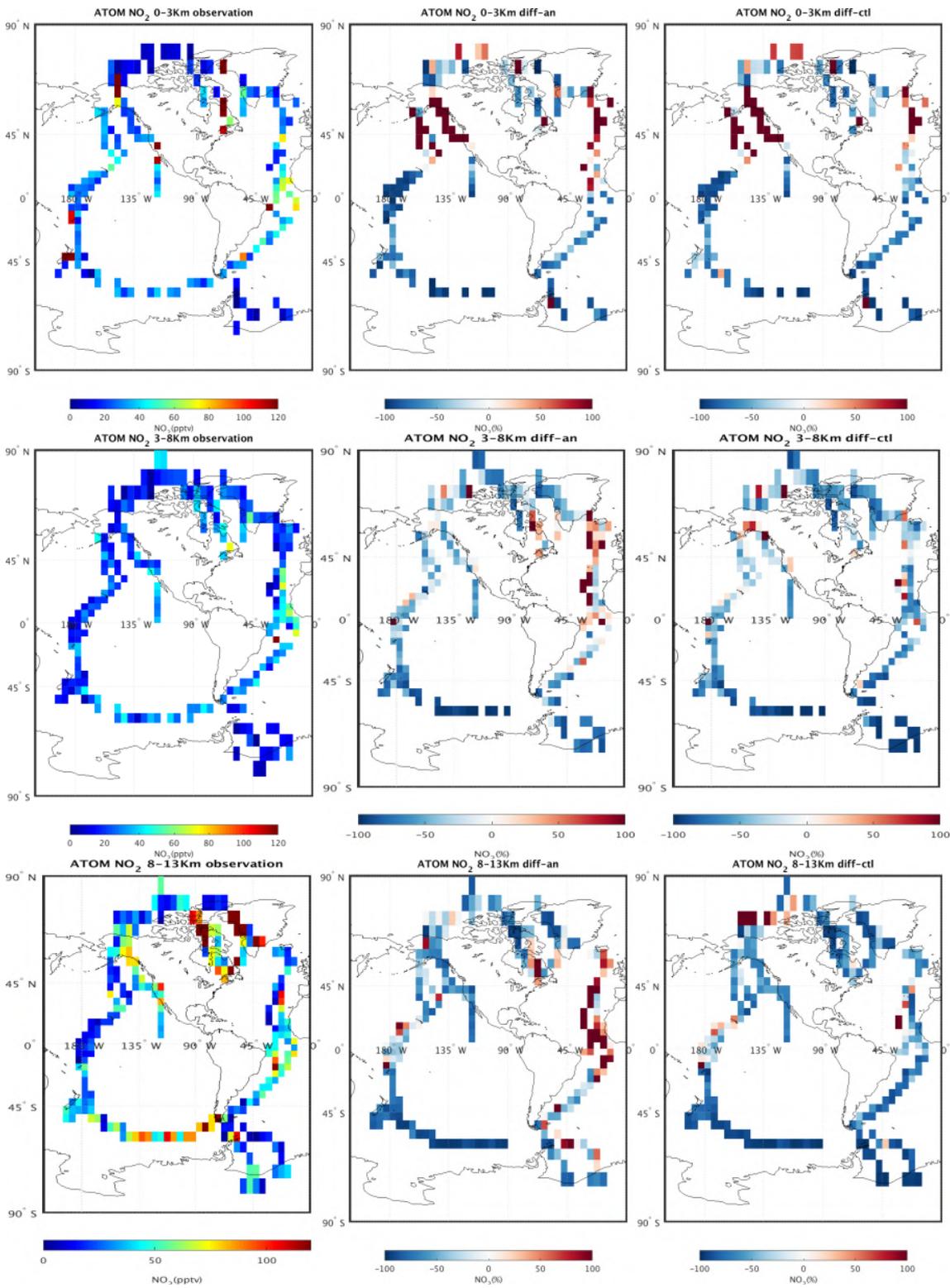


Figure 7.3.5: ATOM campaign observations of NO<sub>2</sub> in pptv (left, see table 7.1.1 and 7.1.2), the difference between the reanalysis and the observation (difference = reanalysis – observation; centre), and the difference between the control run and the observation (difference = control – observation; right). The data are averaged to 5°x5° (latitude x longitude) and to three altitude bins: 0-3 km (top), 3-9 km (middle), and 9-15 km (bottom).

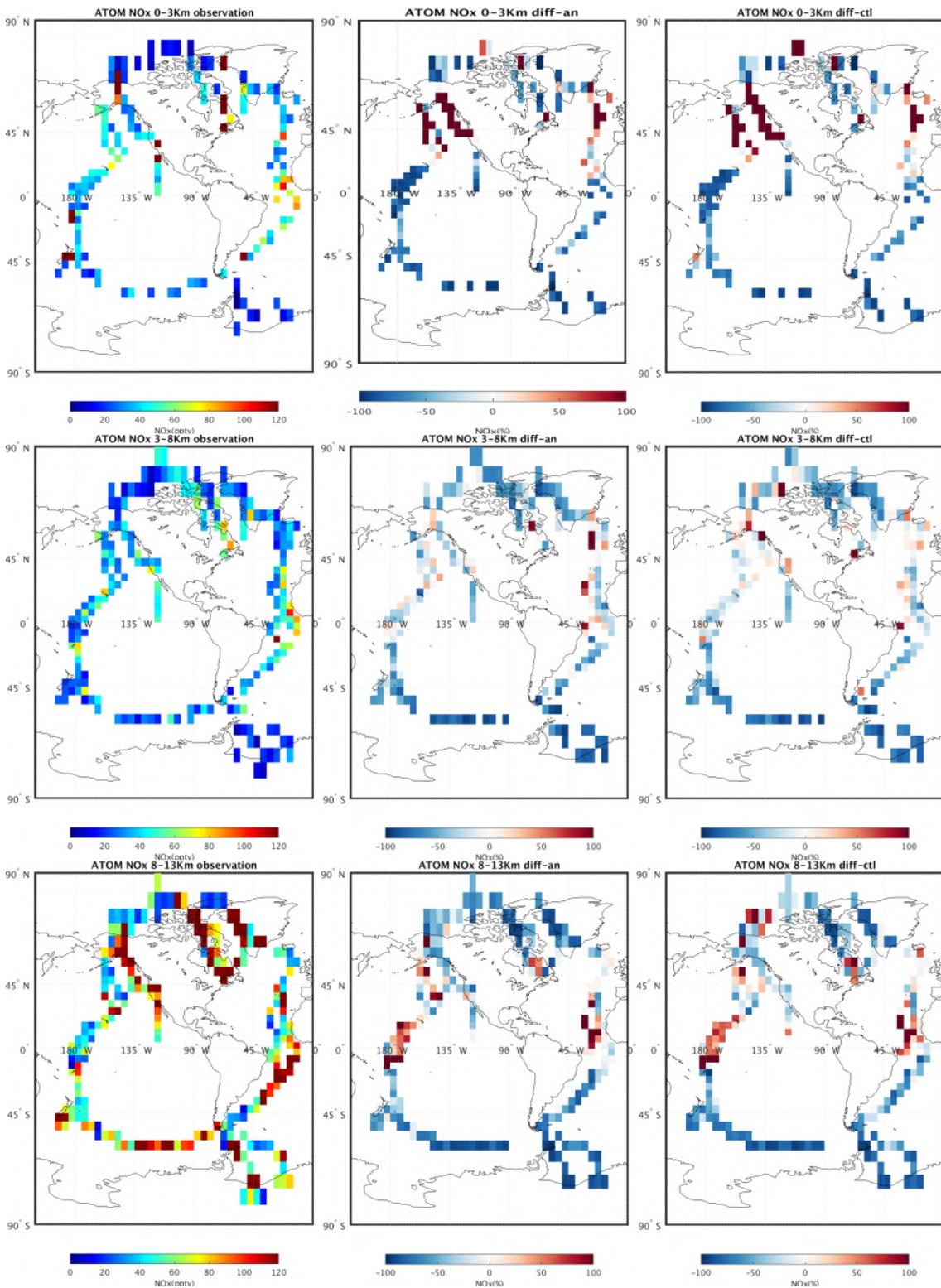


Figure 7.3.6: ATOM campaign observations of NO<sub>x</sub> in pptv (left, see table 7.1.1 and 7.1.2), the difference between the reanalysis and the observation (difference = reanalysis – observation; middle), and the difference between the control run and the observation (difference = control – observation; right). The data are averaged to 5°x5° (latitude x longitude) and to three altitude bins: 0-3 km (top), 3-9 km (middle), and 9-15 km (bottom).



### **Hydroxyl radicals**

In Fig. 7.3.7, the reanalysis OH comparison shows similar patterns through all layers, indicating a strong underestimation over the Atlantic Ocean (up to 100%) and overestimation over the Pacific Ocean, except for the southern Pacific Ocean where the bias is negative. The control run is positively biased over the northern low latitudes and part of the mid-latitudes, and negatively biased elsewhere.

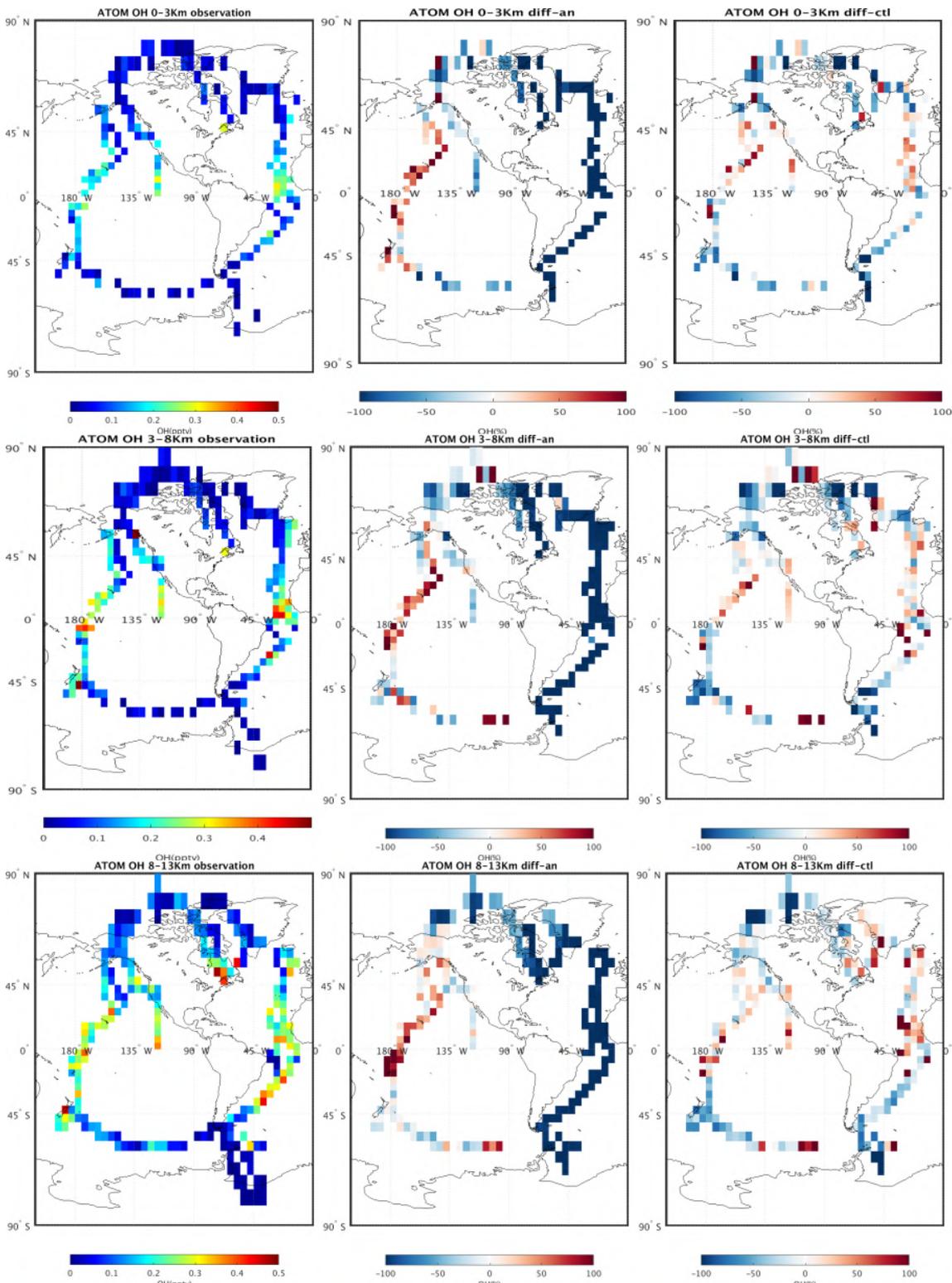


Figure 7.3.7: ATOM campaign observations of OH in pptv (left, see table 7.1.1 and 7.1.2), the difference between the reanalysis and the observation (difference = reanalysis – observation; centre), and the difference between the control run and the observation (difference = control – observation; right). The data are averaged to 5°x5° (latitude x longitude) and to three altitude bins: 0-3 km (top), 3-9 km (middle), and 9-15 km (bottom).



## 8. Stratosphere

### 8.1 Validation against ozone sondes

Stratospheric ozone from the CAMS reanalysis is evaluated against ozone soundings from the NDACC, WOUDC, NILU and SHADOZ databases. The sondes have a reported 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 reanalysis are compared to balloon sonde measurement data of 44 stations for the period January 2003 -December 2020. The validation covers the vertical range between 90 and 10hPa, (for the Tropics 60 and 10hPa). A description of the applied methodologies and a map with the sounding stations can be found in Eskes et al. (2018).

Over the Northern Midlatitudes, O<sub>3</sub> partial pressures are reproduced accurately with mostly slightly positive MNMBs between -1% and 10%, see Fig. 8.1.1. Over the Arctic, O<sub>3</sub> partial pressures are mostly slightly underestimated (MNMBs are between ±15%). Largest negative MNMBs appear during the first half of 2004 due to a lack of satellite data for data assimilation. From January 2016 onwards, MNMBs are getting positive (MNMBs up to 15%). This increase could be associated with an instrument related drop-off of stratospheric ozone reported for a variety of global sonde stations. Although the exact reason for the caveat is still under investigation, it is likely related to a change in ECC instruments after 2014 (see e.g., <https://tropo.gsfc.nasa.gov/shadoz/Archive.html>, or Stauffer et al. 2020). The drop-off in stratospheric ozone could thus entail larger positive MNMBs in the comparison with the CAMS reanalysis. Over the other regions, this effect does not show up as clearly, likely because for other regions the majority of stations used in the averaging are unaffected by the ozone drop-off. Over the Tropics, stratospheric ozone is slightly overestimated throughout 2003-2020 with MNMBs <11%. Over the Antarctic, MNMBs are mostly <10%, except during the ozone hole season 2003 and 2006, where MNMBs reach up to 25% (18% for 2006). During November 2016 and January 2017 there is an episode with slightly larger negative MNMBs of down to -14%. The overestimation of ozone in the model in October 2020 relates to the very low values between 100-40hPa, hardly visible in the profile comparison (Fig. 8.1.10).

The ozone hole situation between August and November over the Antarctic (Figures 8.1.2 to 8.1.10) could be very well reproduced with MNMBs < 8% in the years 2004, 2005, and 2007, 2008, 2010, and 2012-2015, 2017-2019 and 2020. In other years, the bias is slightly higher. In 2003, MNMBs rise up to 25% in November (Fig. 8.1.2) due to the lack of MLS satellite data for data assimilation. The ozone hole 2006 was one of the largest and deepest in record, and also 2020 was exceptional. Although the model generally reproduces the mean ozone profiles between August and November, the low ozone mixing ratios are overestimated during September and October 2006, which results in MNMBs of up to 18%. In October 2009, the model underestimates ozone mixing ratios in the lower stratosphere by around -10%. During the ozone hole season in 2011, ozone mixing ratios are overestimated by around 10% by the model.

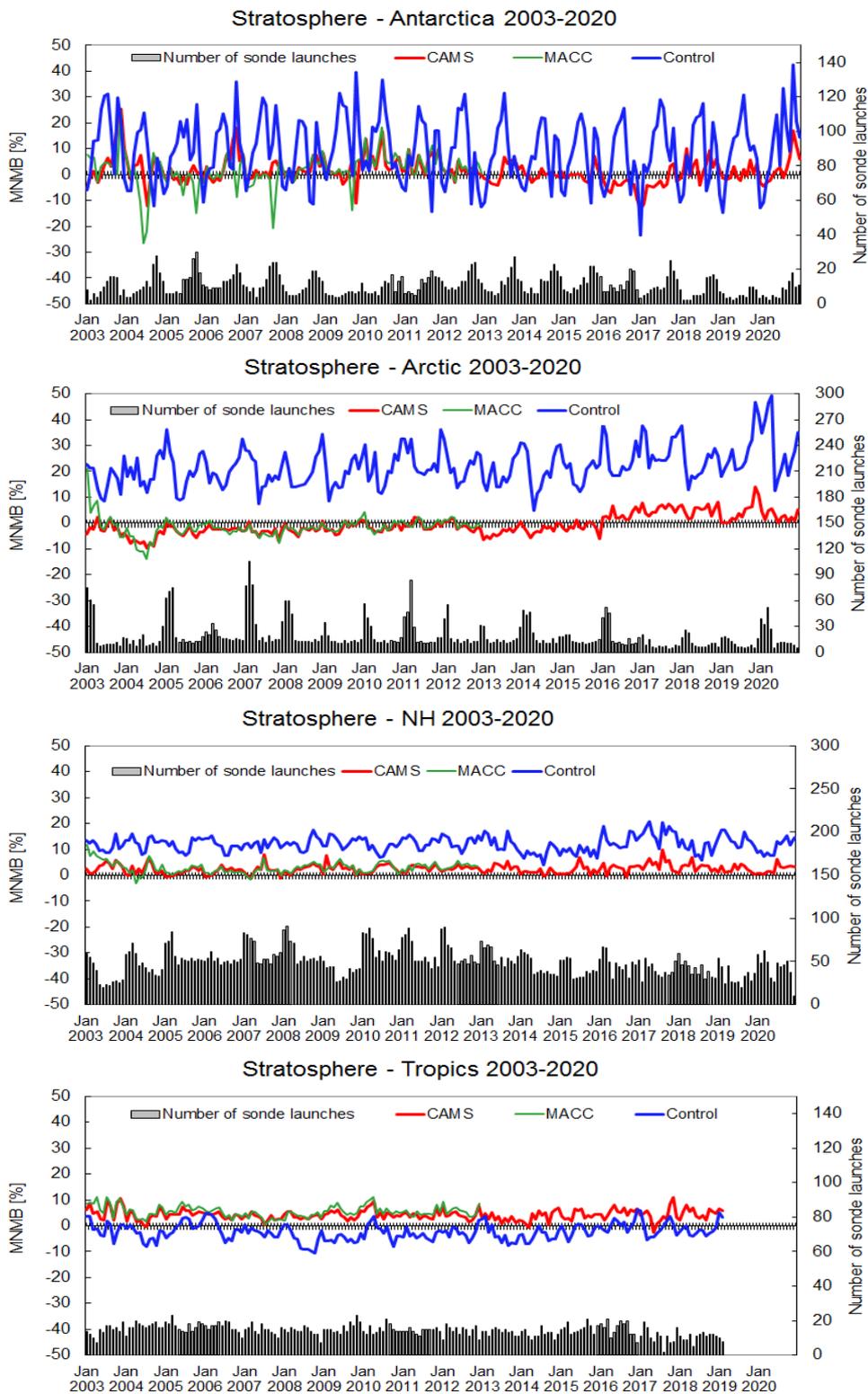


Fig. 8.1.1: Normalised bias of the reanalysis versus ozone sondes, for 4 regions (Antarctica, Arctic, Northern Midlatitudes, Tropics) in the stratosphere. Red: CAMS reanalysis, blue: control run, green: MACC reanalysis. The biases are averaged between 90 and 10hPa in the extra-tropics and between 60 and 10hPa in the Tropics.

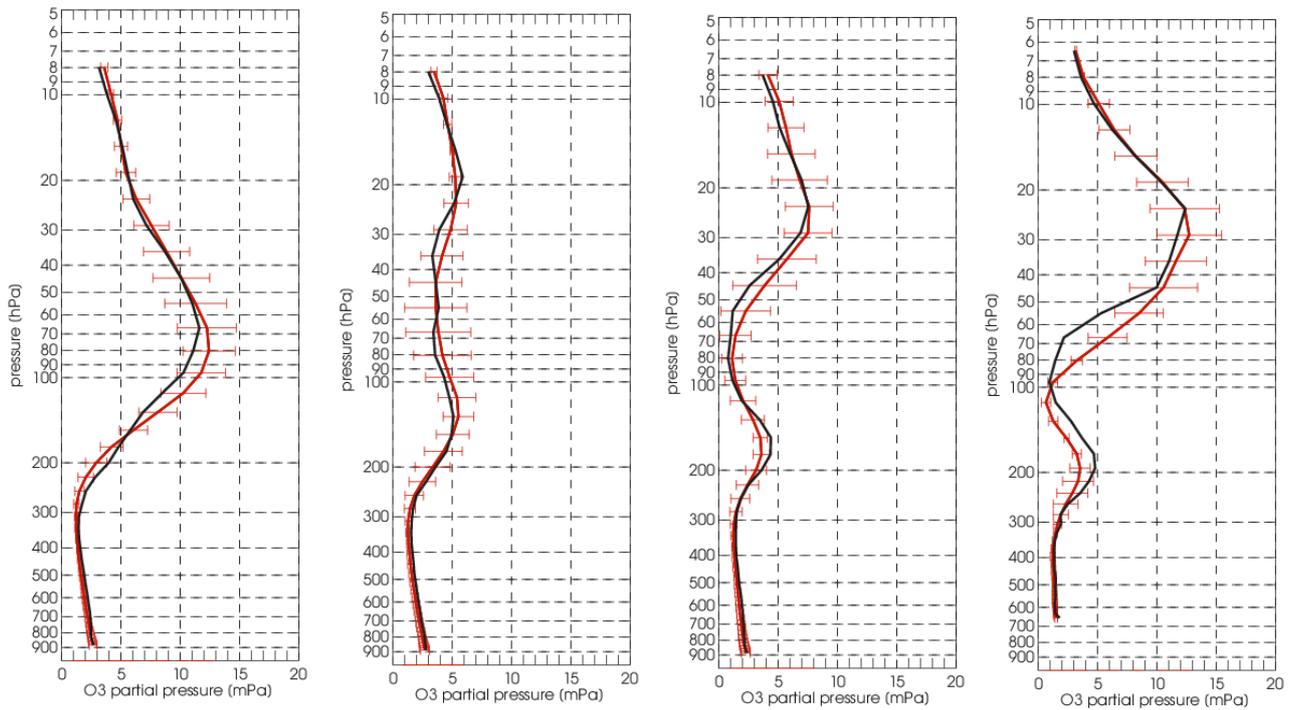


Fig. 8.1.2: Ozone in August, September, October and November 2003 (ozone hole period) over the Antarctic.

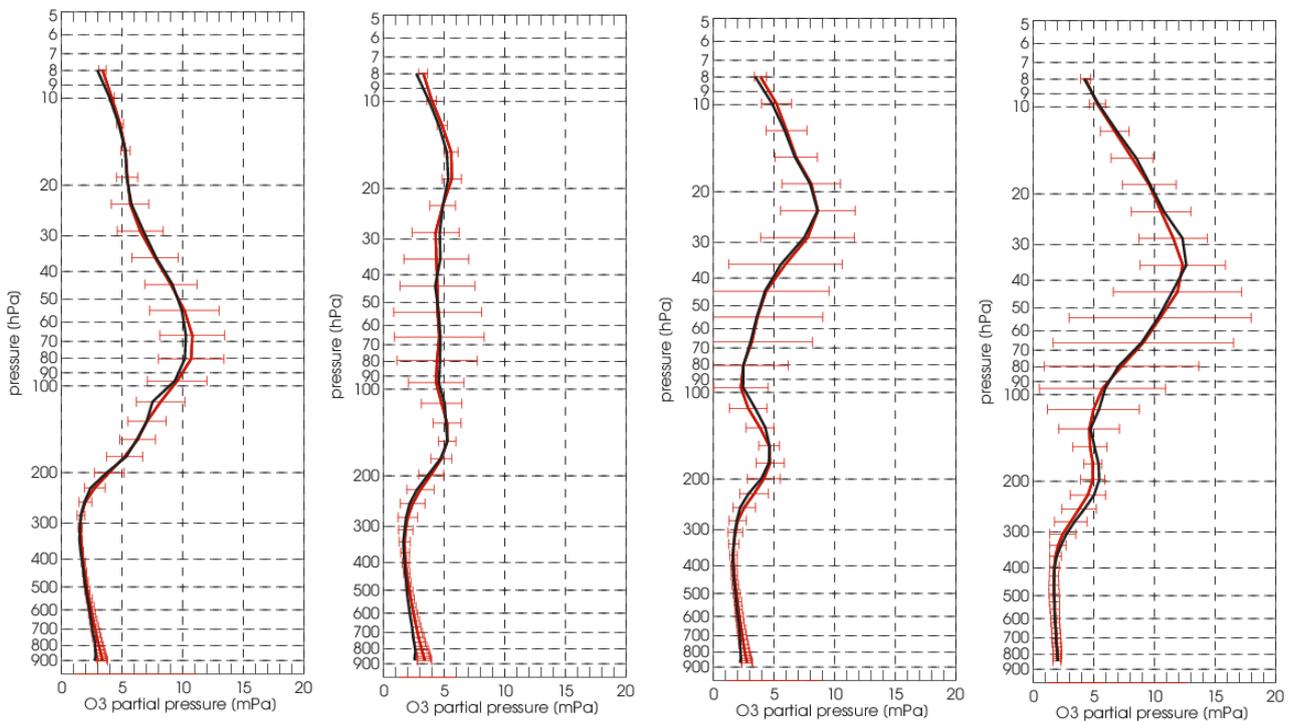


Fig. 8.1.3: Ozone profiles compared with ozone sonde observations in August, September, October and November 2005 (ozone hole period) over the Antarctic.

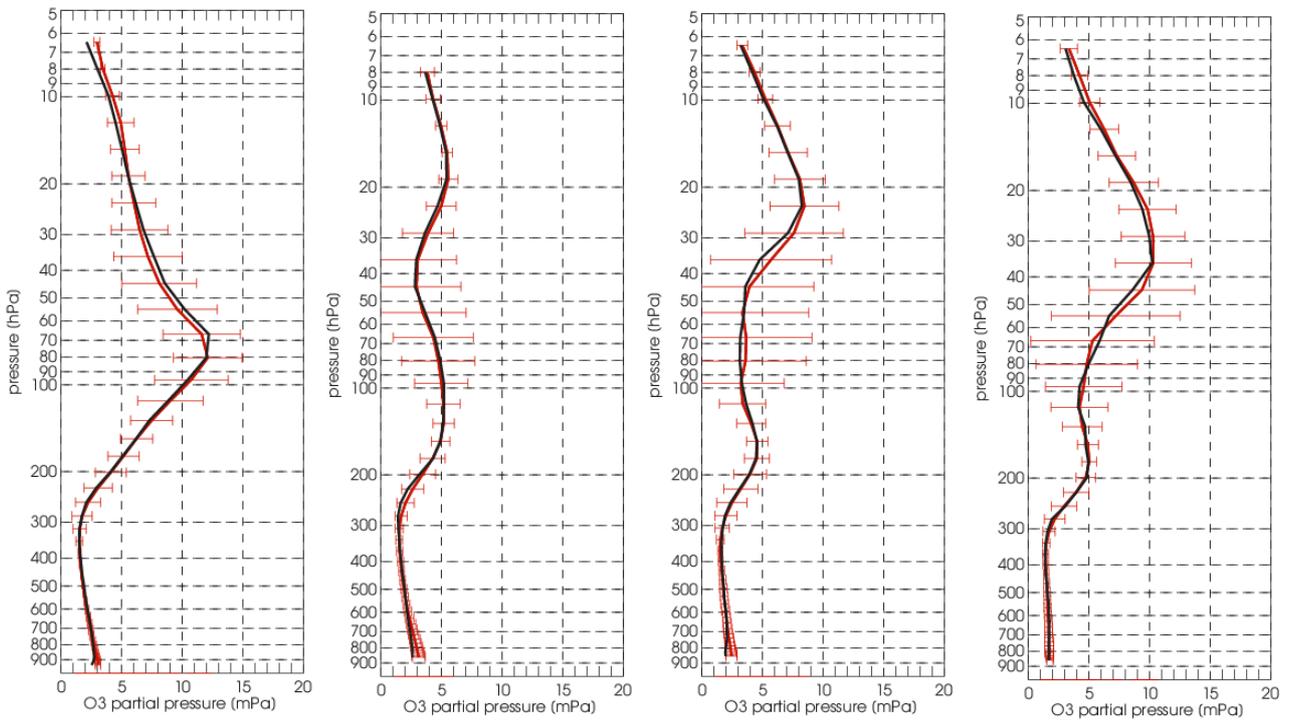


Fig. 8.1.4: Ozone in August, September, October and November 2007 (ozone hole period) over the Antarctic.

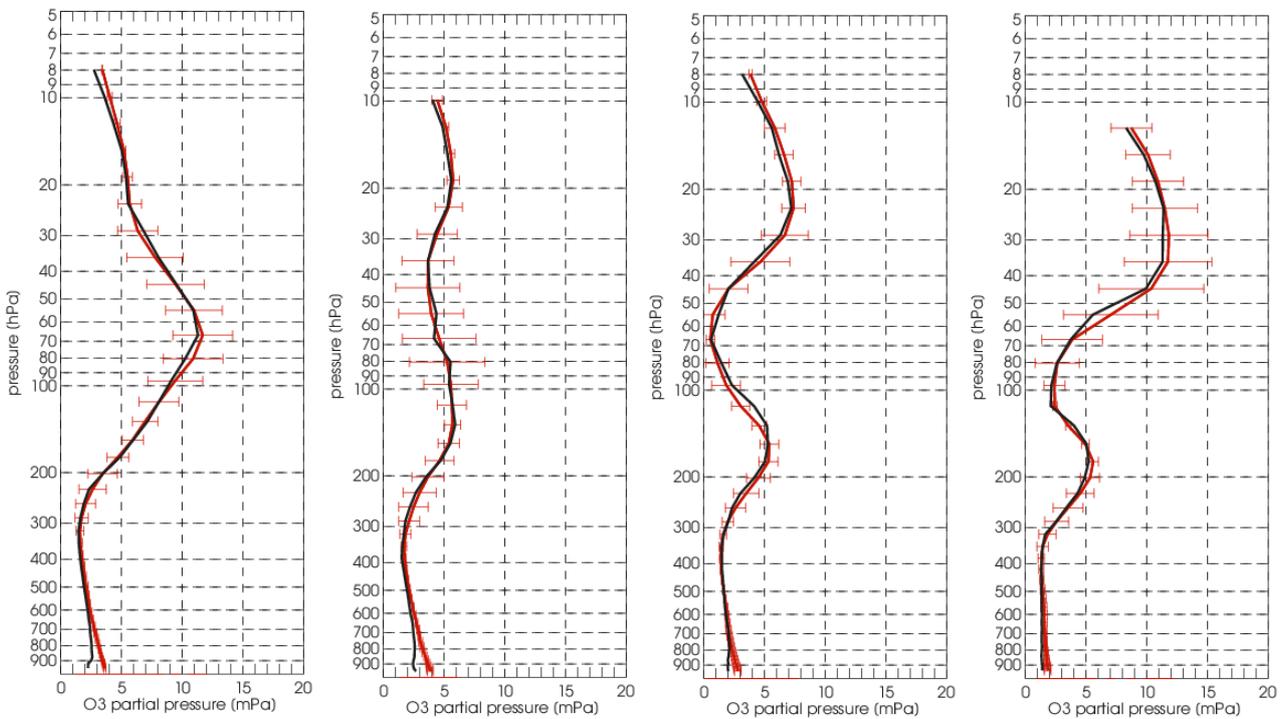


Fig. 8.1.5: Ozone in August, September, October and November 2009 (ozone hole period) over the Antarctic.

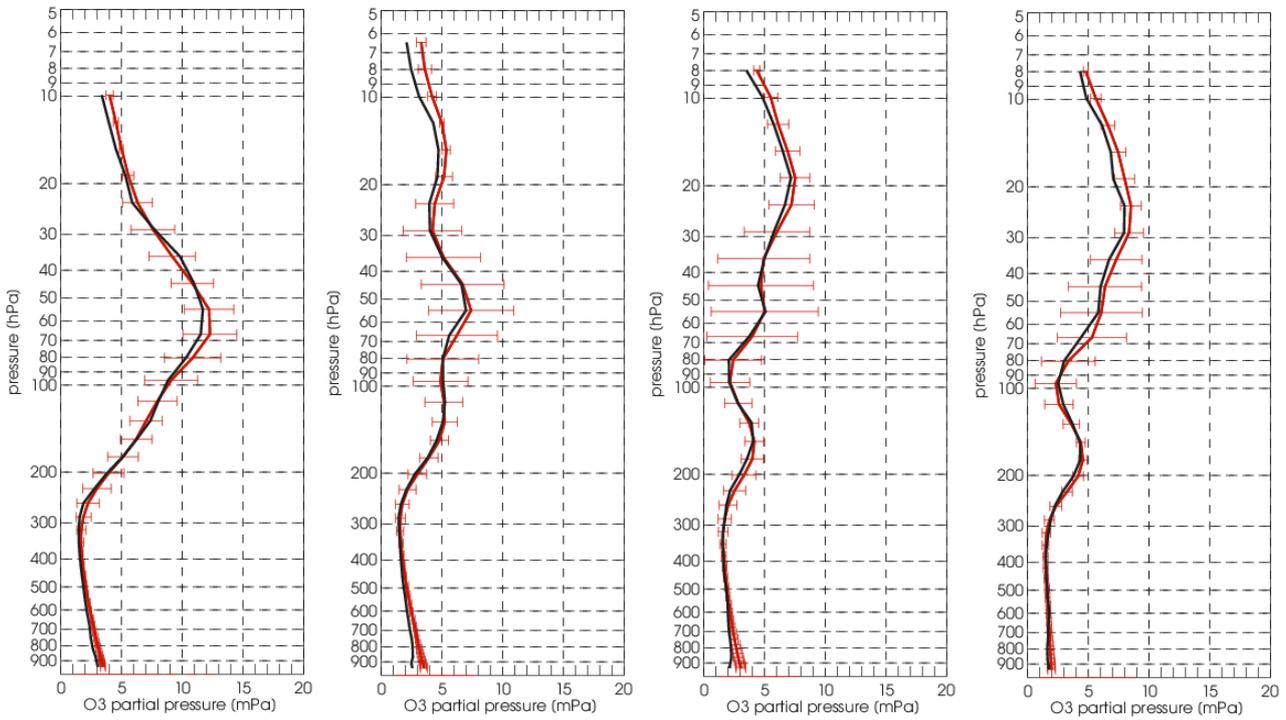


Fig. 8.1.6: Ozone in August, September, October and November 2011 (ozone hole period) over the Antarctic.

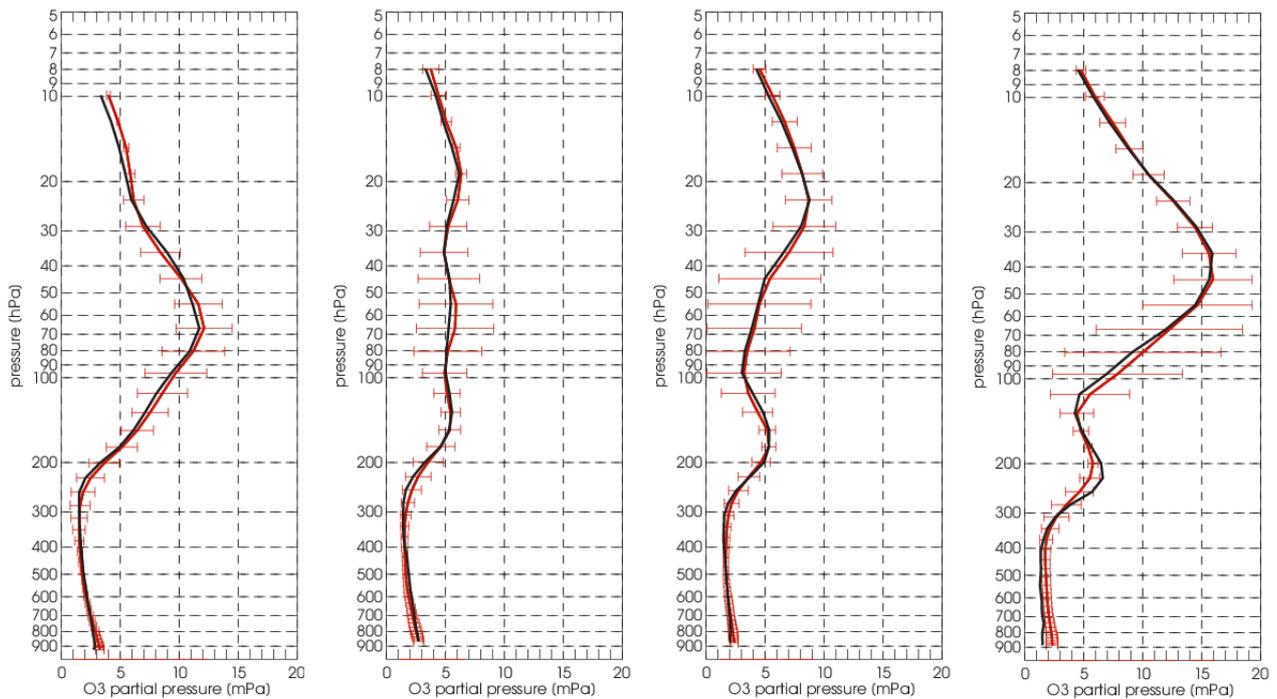


Fig. 8.1.7: Ozone in August, September, October and November 2013 over the Antarctic.

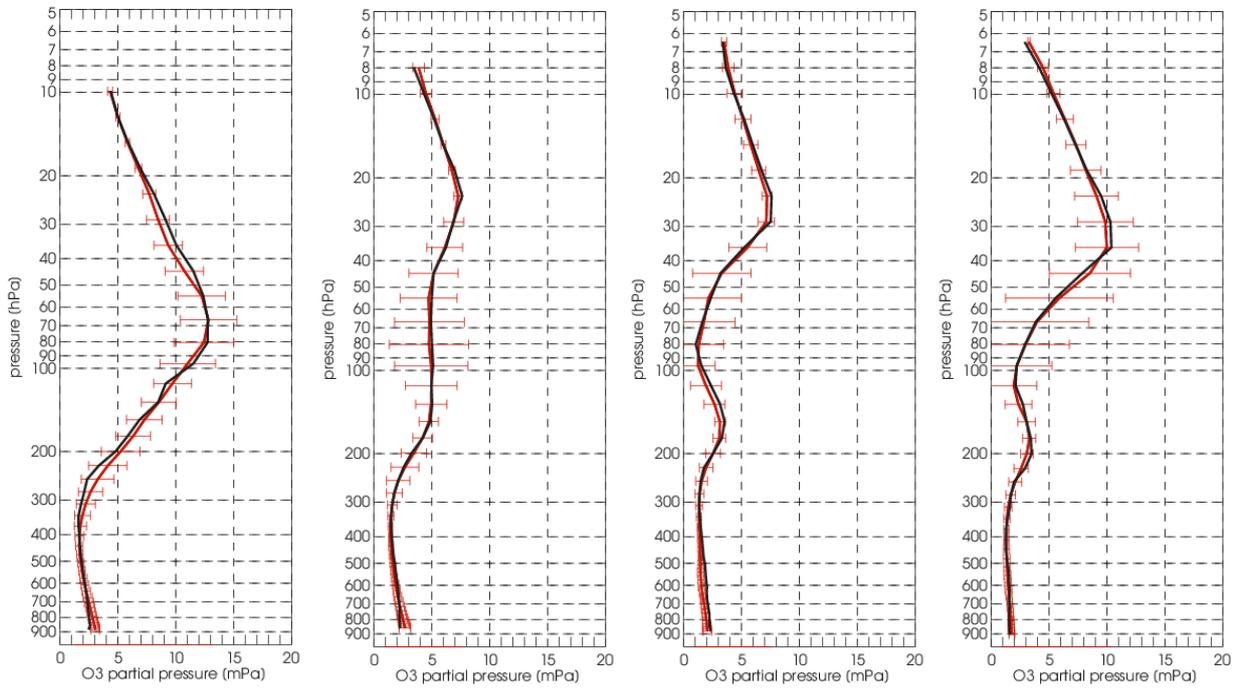


Fig. 8.1.8: Ozone in August, September, October and November 2015 over the Antarctic.

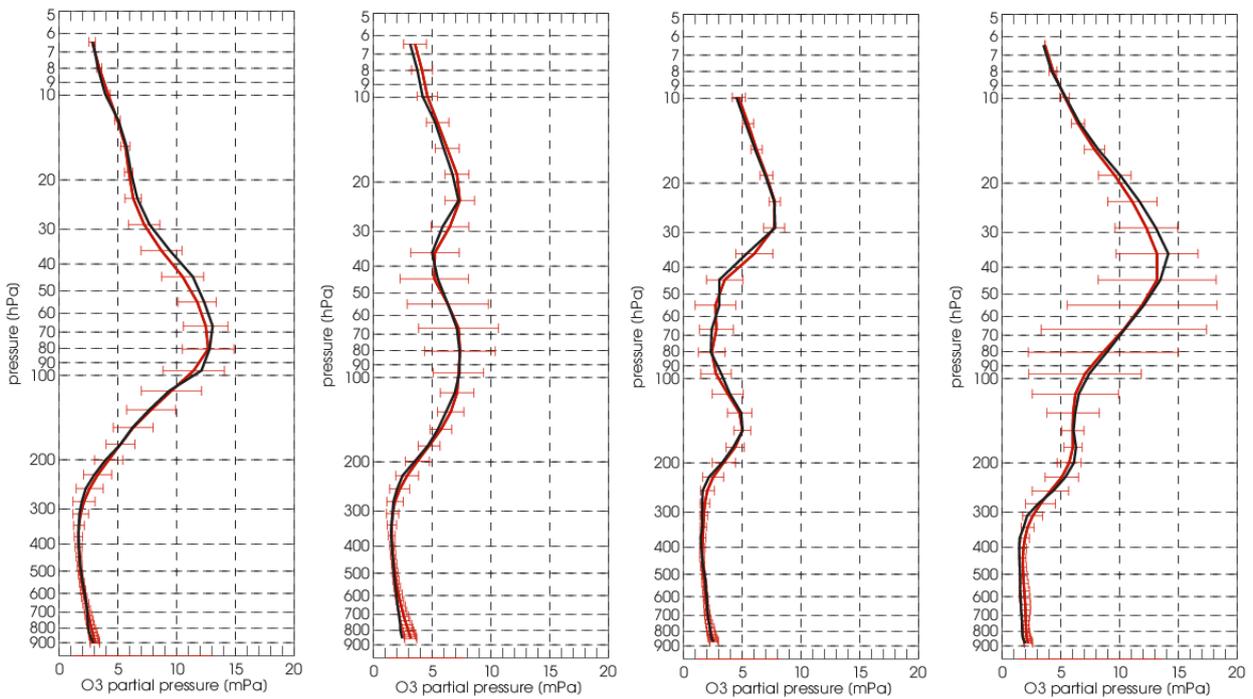


Fig. 8.1.9: Ozone in August, September, October and November 2017 over the Antarctic.

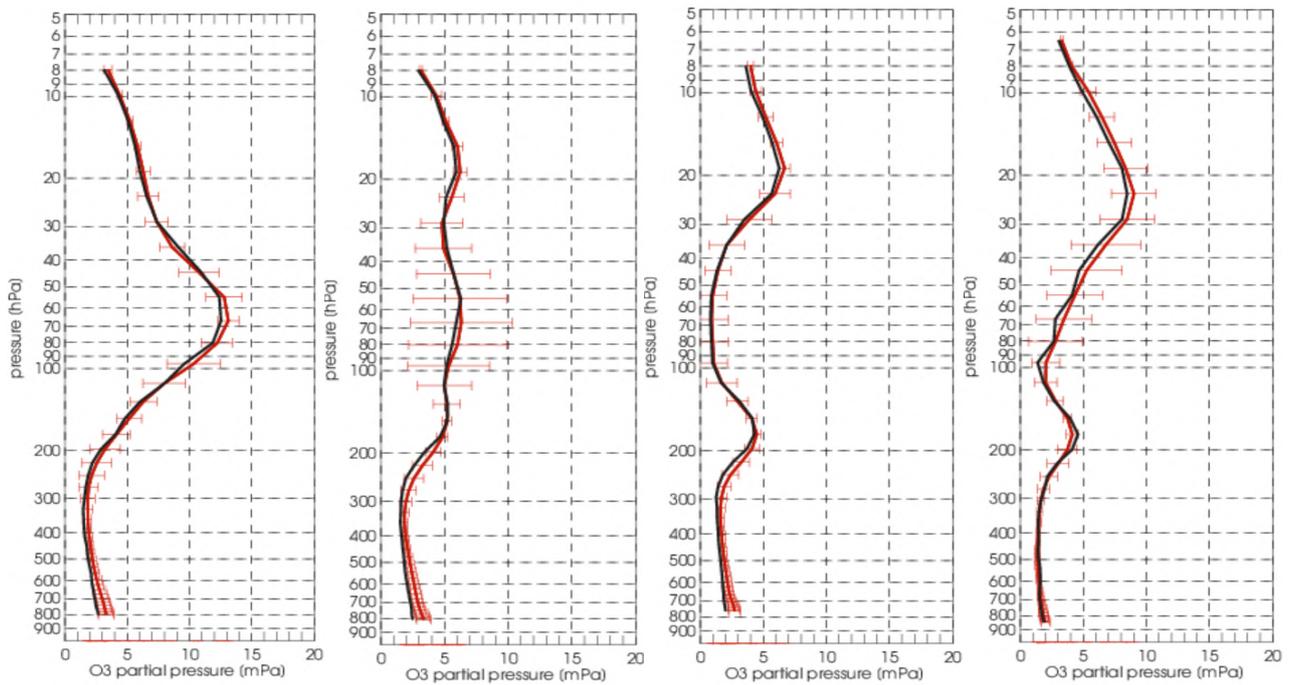


Fig. 8.1.10: Ozone profiles in August, September, October and November 2020 over the Antarctic.



## 8.2 Validation against ozone observations from the NDACC network (MWR, LIDAR)

In this section we present a comparison between the MACC and CAMS reanalyses and CAMS control run against UVVIS, MWR and LIDAR observations from the NDACC network. A detailed description of the instruments and applied methodologies for all NDACC instruments can be found in the Annex 2 and at <http://nors.aeronomie.be>. All measurement techniques are sensitive to the stratosphere. The results presented here are based on the UVVIS and FTIR NDACC data. The stratospheric column from the UVVIS zenith measurements also has a typical uncertainty of 6%. The majority of the UVVIS zenith measurements data was made available within the CAMS27 projects, which started in 2017. The FTIR total column data has a typical uncertainty of 5%. In the table and plots below the MACC reanalysis is indicated in green, the CAMS reanalysis in red and the CAMS control run in blue.

The following list summarizes the main conclusions drawn from the figures presented below.

- The control run shows a latitudinal dependence of the relative bias with a strong overestimation in the Arctic (see Figures 8.2.2 and 8.2.3).
- The Taylor diagrams in Fig. 8.2.1 show that the correlation coefficients of the CAMS reanalysis (for most stations above 0.9) are significantly higher than those of the control run.
- From the time series of the relative differences (see Fig. 8.2.2 and Figure 8.2.3), we conclude that the CAMS reanalysis seems to be stable throughout the entire time series and shows no latitude dependence of the bias, within the measurement's uncertainties.
- Total column comparisons against NDACC FTIR data in **Fout! Verwijzingsbron niet gevonden.** show a slight underestimation of the reanalysis, which falls within the typical reported uncertainty on the total column of 5%.

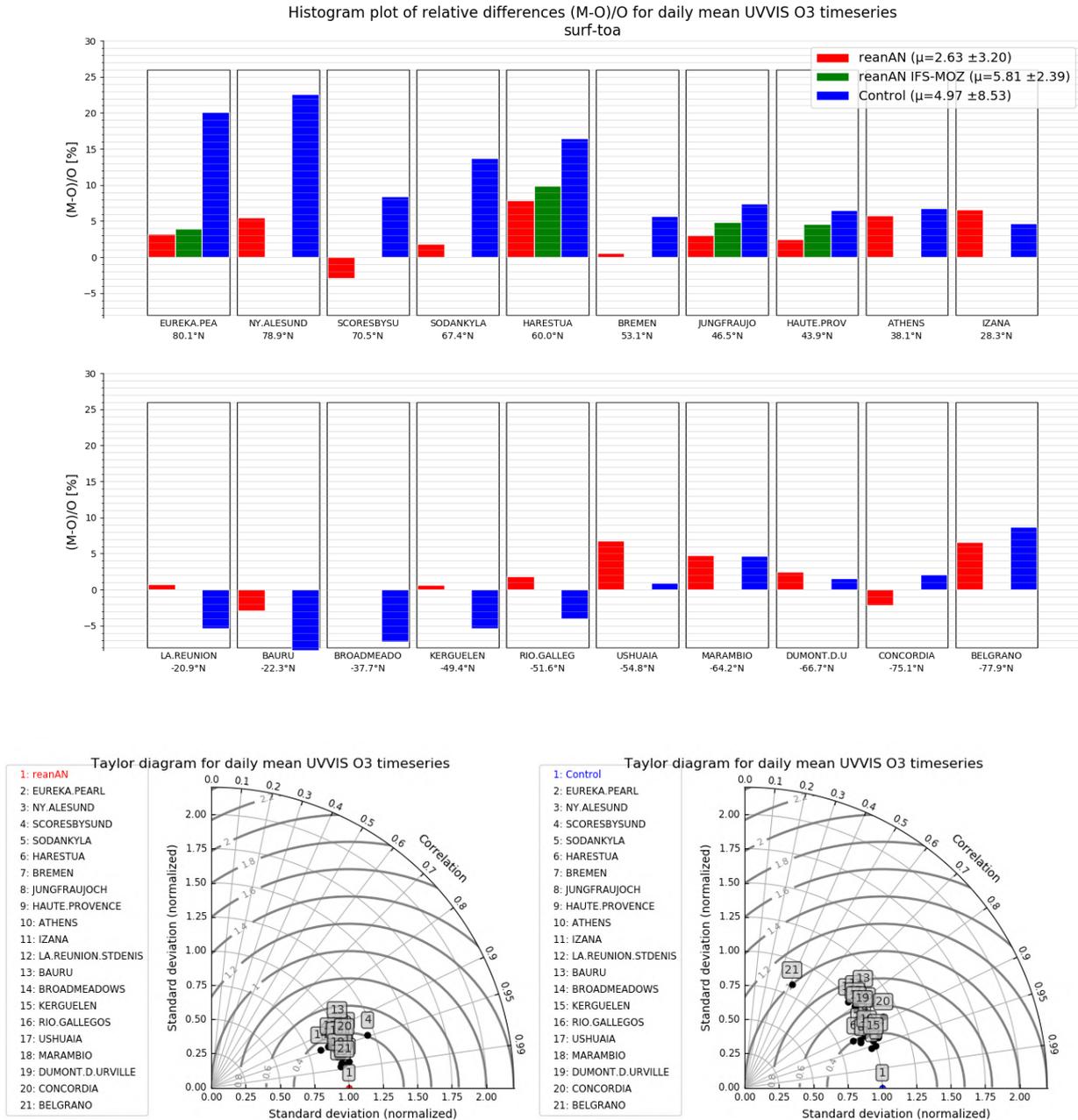


Figure 8.2.1 Averaged daily bias for UVVIS DOAS zenith sky measurements (top, sorted by decreasing latitude) for the entire time period 2003 – 2020. Only a few sites cover the full period and the majority of the sites started delivery in 2017 in the framework of the CAMS27 project. Biases are within the measurement uncertainties. The Taylor diagrams show the positive impact of the assimilation on the correlation coefficients.

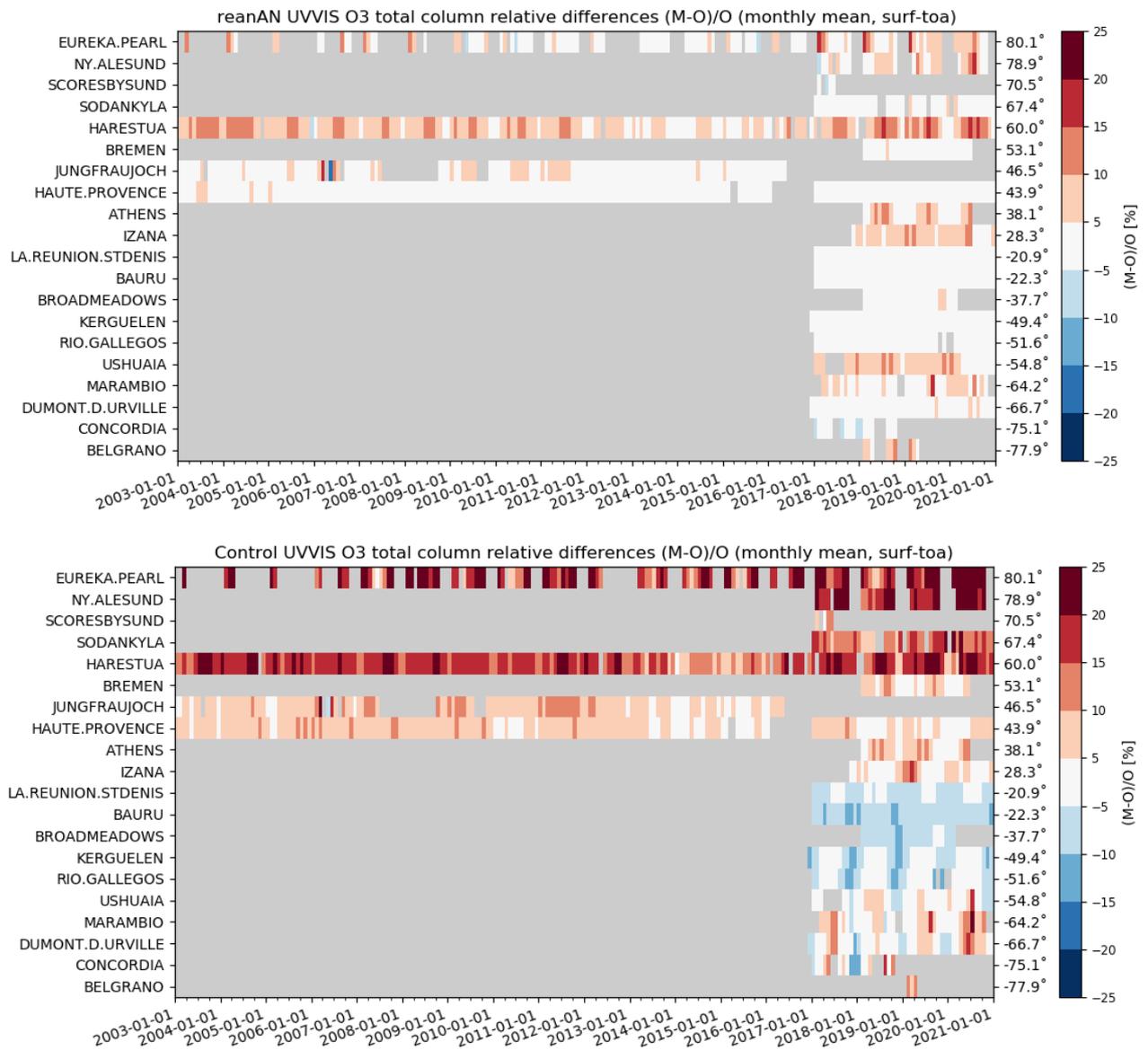


Figure 8.2.2 Time series of monthly column differences for stratospheric ozone measurements from the NDACC UVVIS instruments for the period 2003-2020. The control run has a significant latitudinal dependency in the bias.

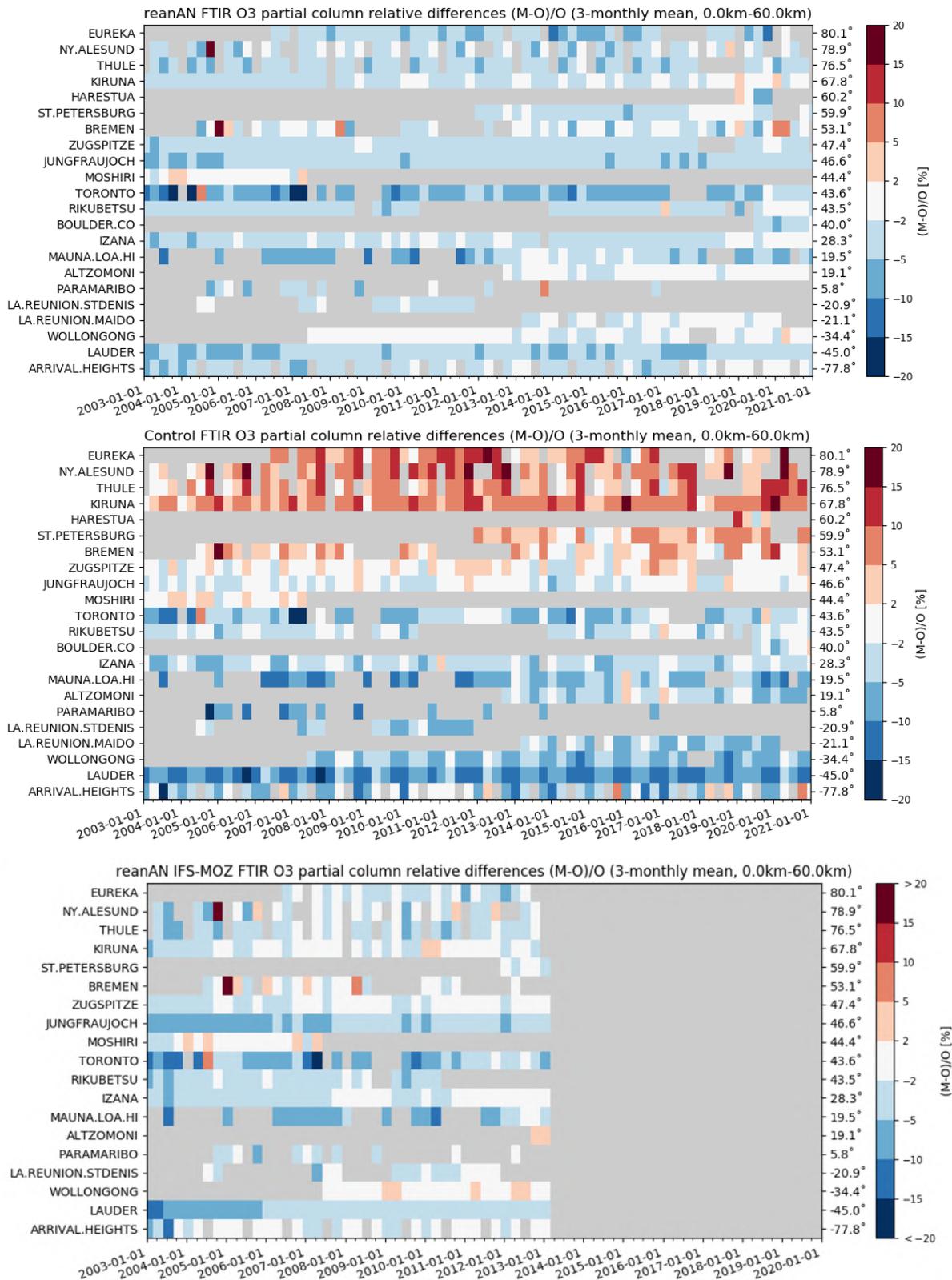


Figure 8.2.3 Time series of seasonal relative total column differences from the NDACC FTIR instruments for all three models under consideration, CAMS reanalysis (top), CAMS control (middle) and MACC (bottom) for the period 2003-2020.



### 8.3 Comparison with observations by limb-scanning satellites

This section compares the model output with observations by limb scanning satellite instruments. These instruments have a global coverage and the retrieved ozone profiles have a typical vertical resolution between 2 and 3 km from the upper troposphere up to the mesosphere. Limb observations that have been used to evaluate the reanalysis are summarized in Table 8.3.1. Among the stratospheric ozone data assimilated in the CAMS reanalysis, those based on MIPAS and MLS instrument are of particular importance because these retrieve ozone profiles at high vertical resolution. So, the measurements from these instruments will provide the largest constraint on the vertical structure of ozone in the reanalysis. Note that there are some discontinuities with the use of these datasets during the period covered by the CAMS reanalysis. Between January 2003 and March 2004, the MIPAS Fourier transform spectrometer ran at full resolution. A failure in March 2004 caused an interruption in the measurements which restarted in January 2005 with a reconfigured mission with a lower spectral resolution of the instrument and a higher vertical sampling. The retrieval version between the first (full resolution) and second (optimized resolution) period are thus different. MLS measures ozone profiles since August 2004 almost continuously except for an interruption of three weeks (26 March to 19 April 2011) that occurred during one of the largest Arctic polar ozone depletion events in 2011. Also, while assimilating off-line v4.2 from August 2004, the CAMS reanalysis switches to the MLS NRT retrieval, which is supposed to be of lower quality, in March 2018.

Table 8.3.1: List of limb instruments used for the validation/verification of the CAMS reanalysis stratospheric ozone analyses. Information provided is the version of the level-2 retrieval algorithm, the period of availability, a flag telling us if the dataset is independent of the assimilated data and the bibliographic reference.

Instrument	Version	Period	Indep. data	Ref.
Aura MLS	Off-line v4.2	Aug 2004-Dec 2020	No	Livesey et al. (2020)
ACE-FTS <sup>1</sup>	v3.6	Mar 2004-Dec 2020	Yes	Boone et al. (2013)
GOMOS <sup>2</sup>	CCI <sup>5</sup> v1	Jan 2003-Mar 2012	Yes	Sovieva et al. (2017)
MIPAS	ESA v7	Jan 2003-Mar 2004 at full resolution Jan 2005-Mar 2012 at optimized resolution	No	Raspolini et al. (2017)
SCIAMACHY <sup>3</sup>	CCI <sup>5</sup> v3.5	Jan 2003-Mar 2012	Yes	Arosio et al. (2018)
SMR <sup>4</sup>	v3 545GHz	Jan 2003-Dec 2020	Yes	Murtagh et al. (2020)

<sup>1</sup>Atmospheric Chemistry Experiment-Fourier Transform Spectrometer

<sup>2</sup>Global Ozone Monitoring by Occultation of Stars

<sup>3</sup>SCanning Imaging Absorption spectroMeter for Atmospheric Cartography

<sup>4</sup>Sub-Millimetre Radiometer

<sup>5</sup>Climate Change Initiative



Figure 8.3.1 presents time series of the monthly normalised mean bias (NMB) between the CAMS reanalysis and the assimilated limb data from MIPAS and MLS in the polar and equatorial region and at different pressure layers from the upper troposphere to 1 hPa. Since MIPAS and MLS data are assimilated, such a comparison is more a verification of the performance of the reanalyses than a validation.

First, one may note the overall better agreement of the CAMS reanalysis with observations than the MACC reanalysis, in particular for the polar regions. The MACC reanalysis also displays one discontinuity around 2007 (see e.g., at the Poles at 10-30 hPa) due to a change in the bias correction scheme for ozone observations, which does not appear in the CAMS reanalysis (Inness et al., 2013). The MACC reanalysis, on the other hand, is in better agreement with the assimilated limb data in the 1-3 hPa layer. This is likely due to the update in the CAMS reanalysis linearized ozone scheme. On March 2018, we observe a small discontinuity in the comparison between the CAMS reanalysis and MLS (e.g., at the Poles in the 10-30 hPa layer or at all latitude in the 100-200 hPa), likely due to the switch in the assimilation of MLS data from the off-line to the NRT retrieved profiles. Overall, the CAMS reanalysis agrees well with the assimilated limb profiles, with a better agreement with MLS probably due to the larger number of MLS profiles and/or lower uncertainty compared to MIPAS. The bias of the CAMS reanalysis against limb profiles is negative in the upper stratosphere (around -10% in the 1-3 hPa layer), is around +/-5% between 3 and 70 hPa, and is positive in the upper troposphere lower stratosphere (UTLS, <10% between 70 to 200 hPa). These levels of agreement are good.

Similar to Fig. 8.3.1, Fig. 8.3.2 presents time series of the differences between the CAMS reanalysis and independent observations from ACE-FTS, GOMOS, SCIAMACHY and SMR. ACE-FTS is a solar occultation instrument with high vertical coverage, relatively low uncertainty and a high stability but has a lower sampling due to the solar occultation method. GOMOS and SCIAMACHY were both instruments on board the Envisat satellite and they operated almost continuously between January 2003 and the end of Envisat in March 2012. SMR is one of the few instruments that covers the whole CAMS reanalysis period and its level-2 profiles have been recently reprocessed. Note however that until 2006, SMR was working only one day in two because its platform, the Odin satellite, was also operating astronomical instruments up to that time.

The agreement between the CAMS reanalysis and these independent data is generally good and stable over time although it shows some seasonal variability. Using ACE-FTS as reference, the bias is mostly negative in the upper stratosphere (between -20% and -10% in the 1-3 hPa layer), is between  $\pm 5\%$  between 3 and 200 hPa except in the 100-200 hPa layer in the Tropics where the variability of the bias is larger than  $\pm 15\%$ . We note several discontinuities in the time series, e.g. in the south polar region when MIPAS was off and before the start of MLS. The switch from the assimilation of the offline to the NRT MLS data is also slightly visible in the comparison against ACE-FTS (e.g., in the Tropics at 30-70 hPa or in the north polar region at 100-200 hPa). These discontinuities would prevent the use of the whole CAMS reanalysis period for stratospheric ozone trend studies.

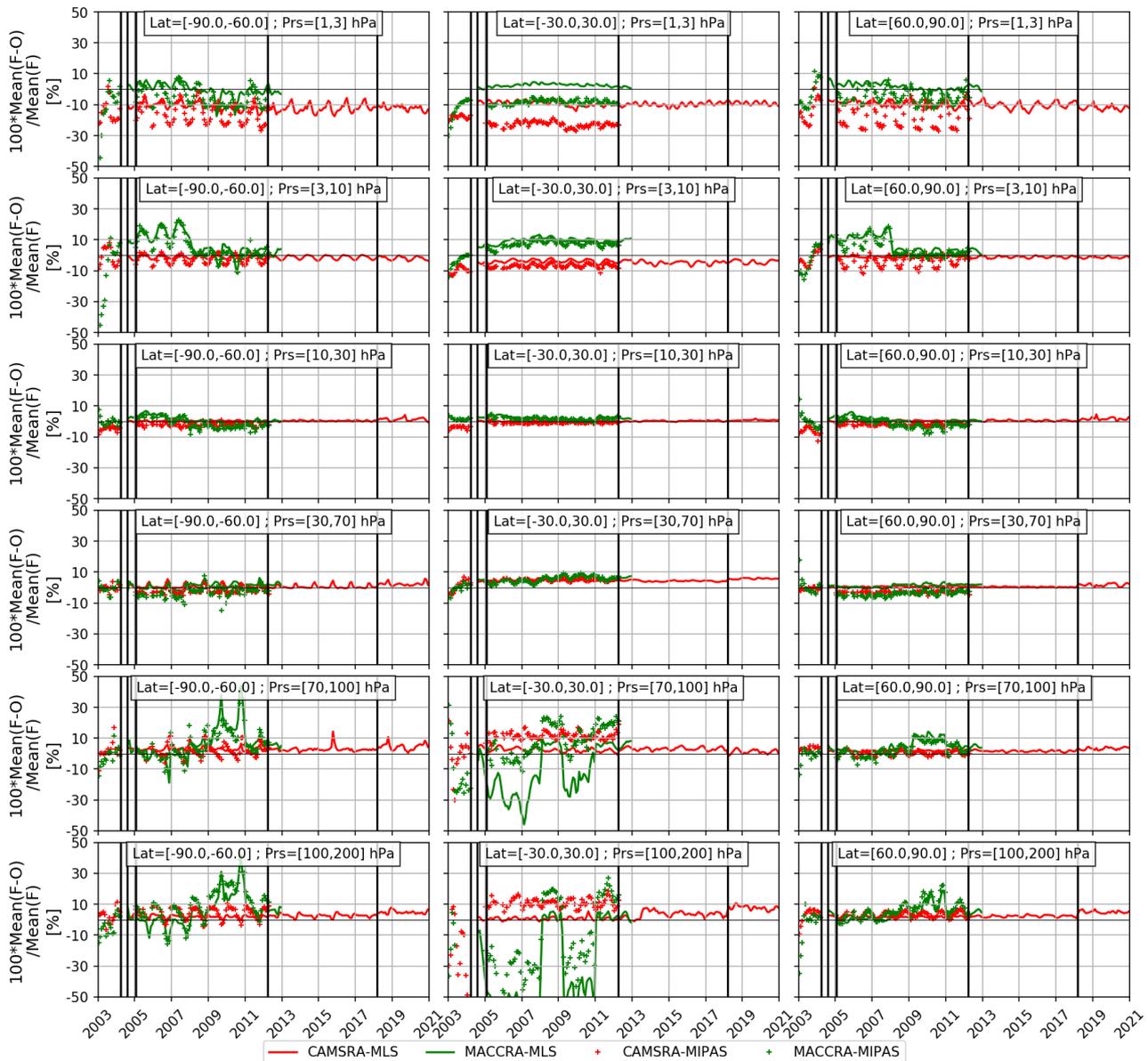


Fig. 8.3.1: Time series comparing the normalized mean bias monthly  $\text{Mean}(\text{rean-obs})/\text{Mean}(\text{rean})$  (%) of ozone with assimilated observations from Aura-MLS v4.2 (solid lines) and MIPAS ESA v7 (crosses) for CAMS reanalysis (red) and MACC-RA (green). The comparison is shown for the 2003-2020 CAMS reanalysis period in several latitude bins and pressure layers in the stratosphere. Latitudes are, from left to right: 90°S-60°S, 30°S-30°N and 60°N-90°N. Pressure layers are, from top to bottom: 1-30, 3-10, 10-30, 30-70, 70-100 and 100-200 hPa. Vertical black lines denote discontinuities in the assimilated limb profiles, from left to right: MIPAS interruption in March 2004, start of MLS in August 2004, restart of MIPAS in January 2005, end of MIPAS in March 2012, switch from the assimilation of MLS off-line to MLS NRT retrieval.

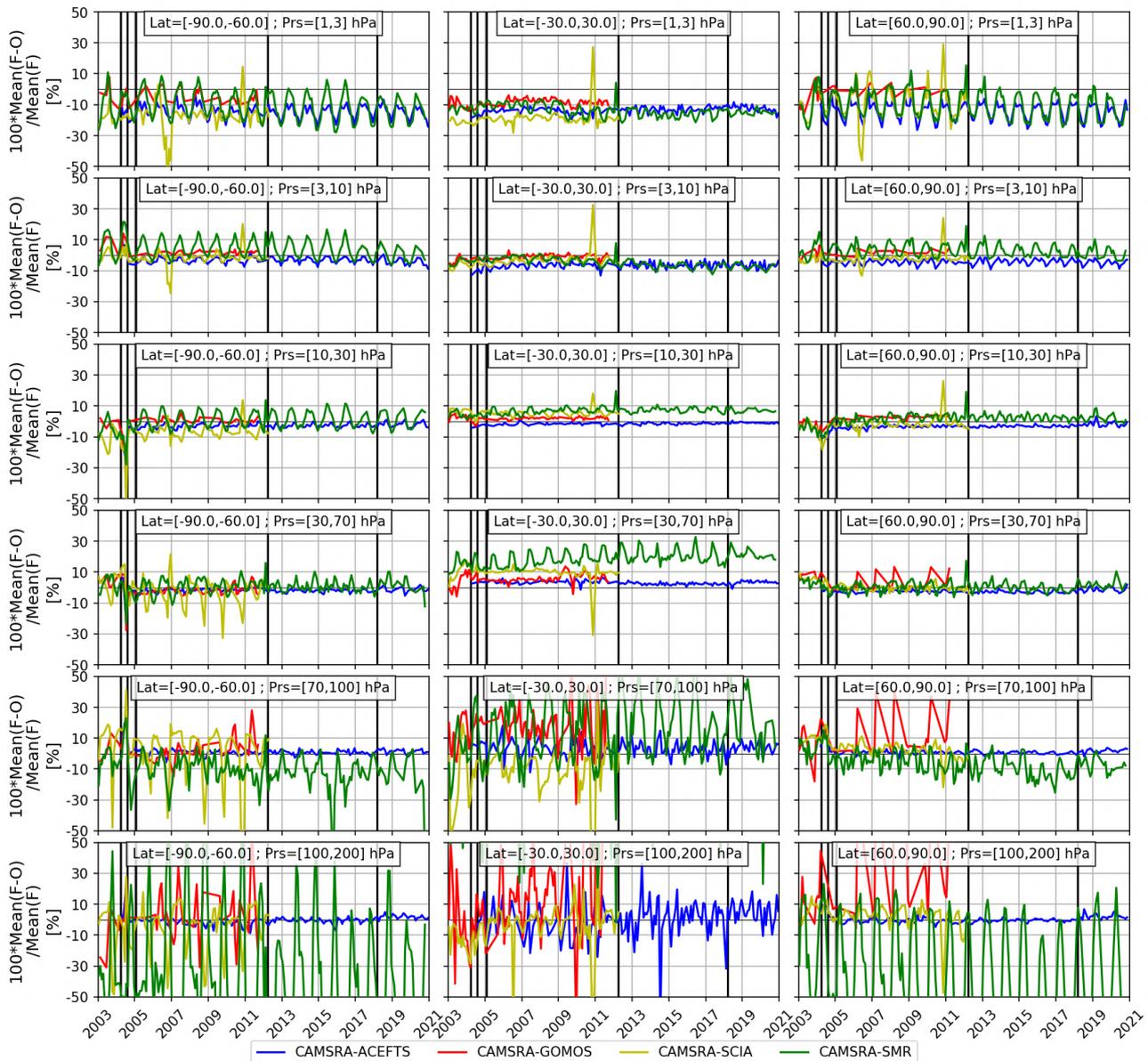


Fig. 8.3.2: As Fig. 8.3.1 but for comparison with independent observations from ACE-FTS (blue lines), GOMOS (red lines), SCIAMACHY (yellow lines) and SMR (green lines) for the period 2003-2020.

During the CAMS reanalysis period, there have been several exceptional events affecting stratospheric ozone: two Arctic winters (2011 and 2020) marked by a very stable vortex which led to the largest ozone depletion ever measured and the longest Antarctic ozone hole (2020). Based on the stability of the time series shown in Fig. 8.3.2, one could note that CAMS reanalysis performance has not been impacted by these events, which is good.

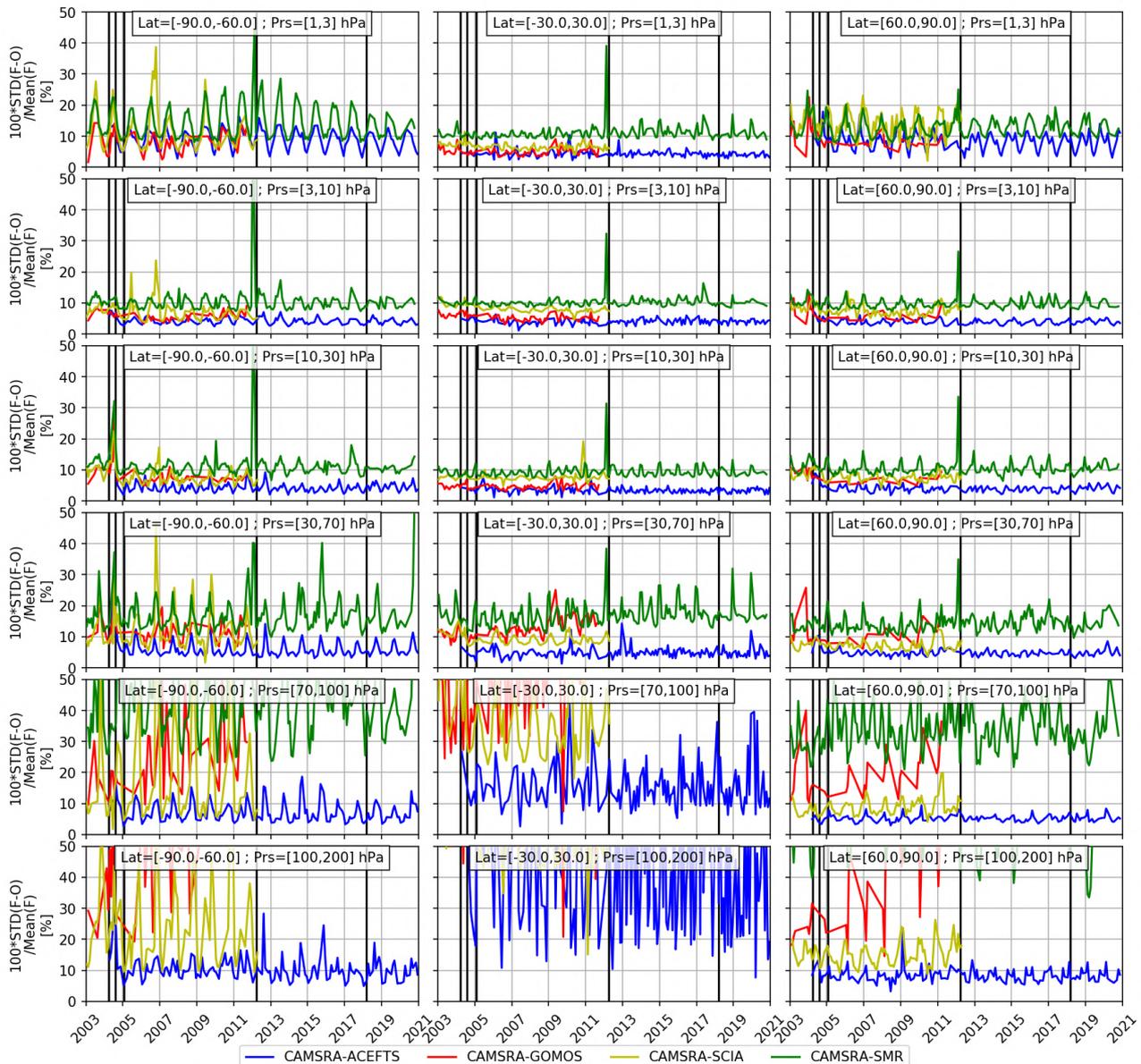


Fig. 8.3.3: As Fig. 8.3.2 but showing the normalized standard deviation of the differences for the period 2003-2020.

The time series of the normalized standard deviation (NSD) of the differences against independent observations (see Fig. 8.3.3) are also relatively good and stable over time. Against ACE-FTS, values are usually below 10% in the upper and middle stratosphere. The comparisons strongly increase (degrade) in the tropical upper troposphere, with a standard deviation often larger than 50%. However, satellite limb data are known to suffer in this region such that comparison with ozonesondes should be more reliable.

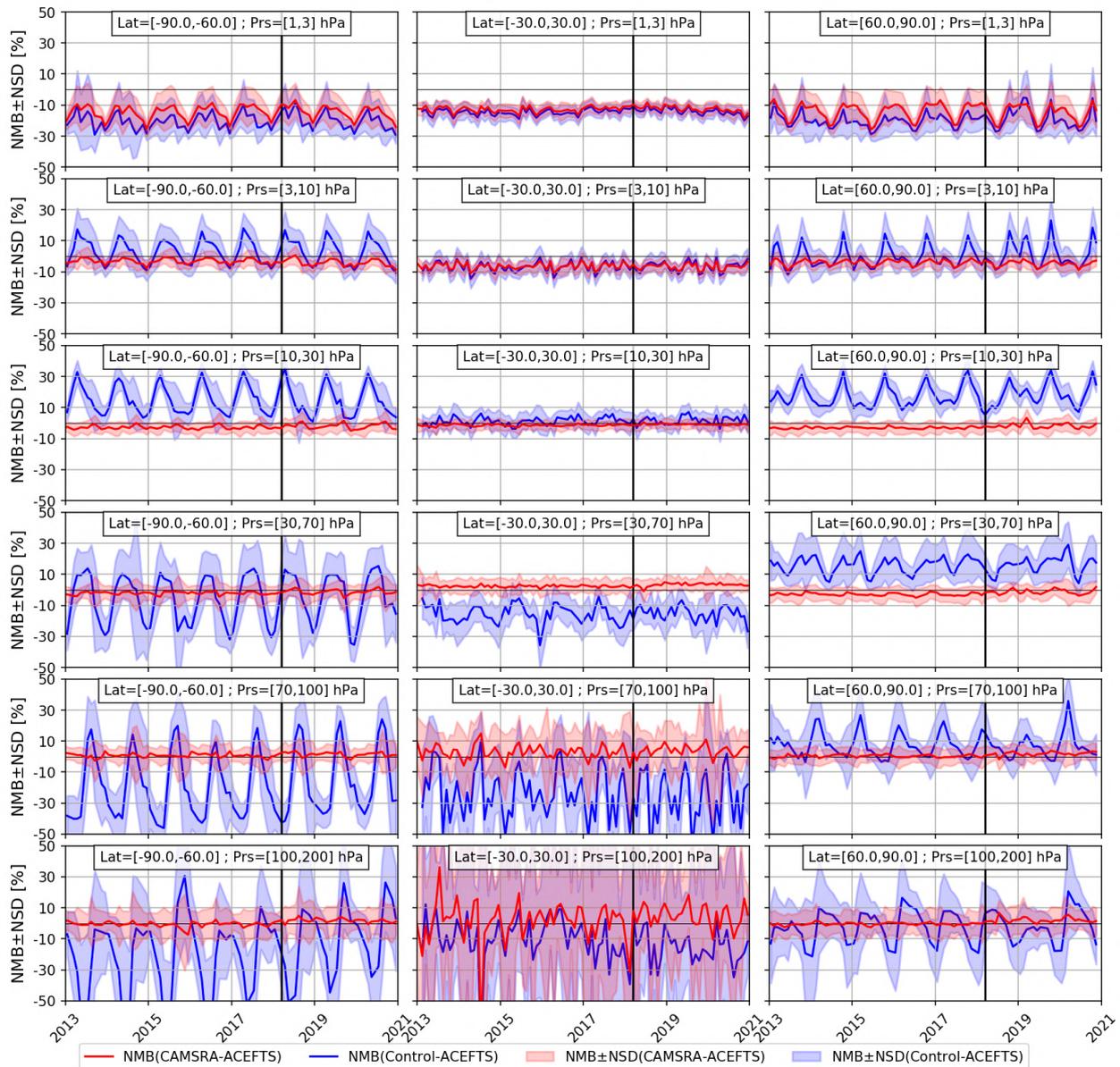


Fig. 8.3.4: As Fig. 8.3.1 but for comparison of the CAMS reanalysis (red lines) and the control run (blue lines) against ACE-FTS for 2013-2020.

Figure 8.3.4 displays the time series of the differences of ACE-FTS against the CAMS reanalysis and the control run. For pressures higher than 10 hPa, it demonstrates how the data assimilation allows the CAMS reanalysis to reduce the bias and spread of the difference against ACE-FTS. For pressures lower than 3 hPa, and in the Tropics in the 3-10 hPa layer, the benefit of the assimilation is less pronounced. This is particularly the case in the 1-3 hPa layer in the Tropics. There, ozone is at photochemical equilibrium such that improving ozone analysis could not be obtained by optimizing the model ozone initial conditions which is the aim of the data assimilation setup. Optimizing model parameters controlling the ozone photochemistry would be more efficient in this region. However, little progress has been made in this area concerning stratospheric ozone and maybe the next reanalysis could exclude ozone data above around 5 hPa.

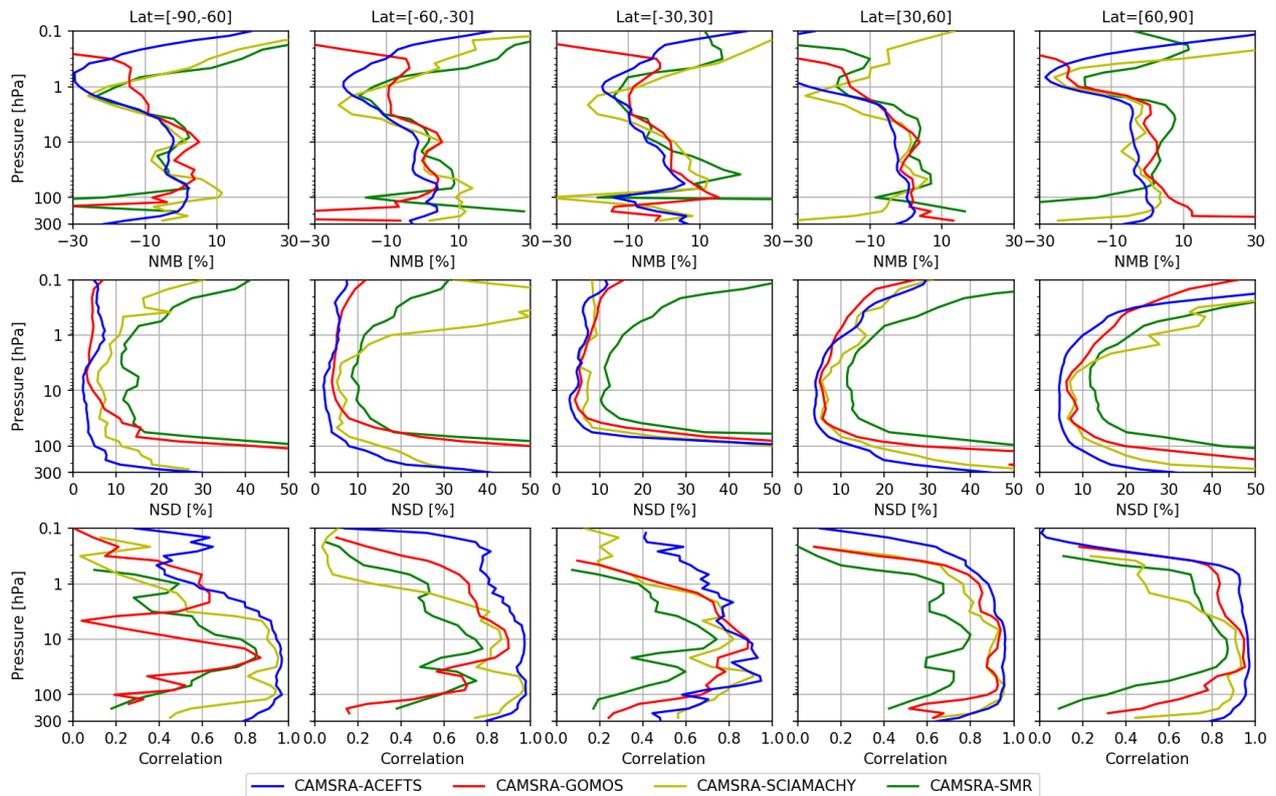


Fig. 8.3.5: Top row: Normalized mean bias (NMB,  $\text{Mean}(\text{CAMSRA-Obs})/\text{Mean}(\text{CAMSRA})$ ) between the CAMS reanalysis and independent observations from ACE-FTS (blue lines), GOMOS (red lines), SCIAMACHY (yellow lines) and SMR (green lines) for December-January-February (DJF) season and for the all years with data availability (see Table 8.3.1) in five latitude bands, from left to right: 90°S-60°S, 60°S-30°S, 30°S-30°N, 30°N-60°N and 60°N-90°N. Middle row: The normalized standard deviation of the differences (NSD,  $\text{STD\_DEV}(\text{reanalysis-Obs})/\text{Mean}(\text{reanalysis})$ ) for the same period and latitudes as for the NMB. Bottom row: The correlation between the reanalysis and the observations for the same period and latitudes as for the NMB.

Fig. 8.3.5-8 display profiles of mean differences, the associated standard deviation and the correlations between the CAMS reanalysis and independent observations for the four seasons (DJF, MAM, JJA and SON) in five latitude bands from the south to the north polar regions. While there are some differences in the comparison between the instruments, there are also some common patterns. In general, the agreement is good in the middle stratosphere [4-100 hPa at mid-latitudes and poles, 4-70 in the Tropics] where the NMB is within  $\pm 10\%$ , the NSD is lower than 10% (slightly larger against SMR) and with correlation often larger (better) than 0.8 (especially against ACE-FTS). This level of agreement is maintained during polar winter and spring (JJA and SON poleward of 60°S and DJF and MAM poleward of 60°N). In the extratropical UTLS (100-200 hPa), while NSDs increase between 20 to >50% depending on the instruments with a decrease in the correlations, NMBs are usually within  $\pm 10\%$  and the general level of agreement is relatively good. In the tropical UTLS (70-200 hPa), while NMBs stay relatively low (within  $\pm 10\%$  against ACE-FTS), NSDs increase to larger value than 50% with correlation lower than 0.7. In this region where the ozone amount is much lower than in the middle stratosphere, it is known that satellite instruments have several difficulties and evaluation against ozonesondes would be more reliable.

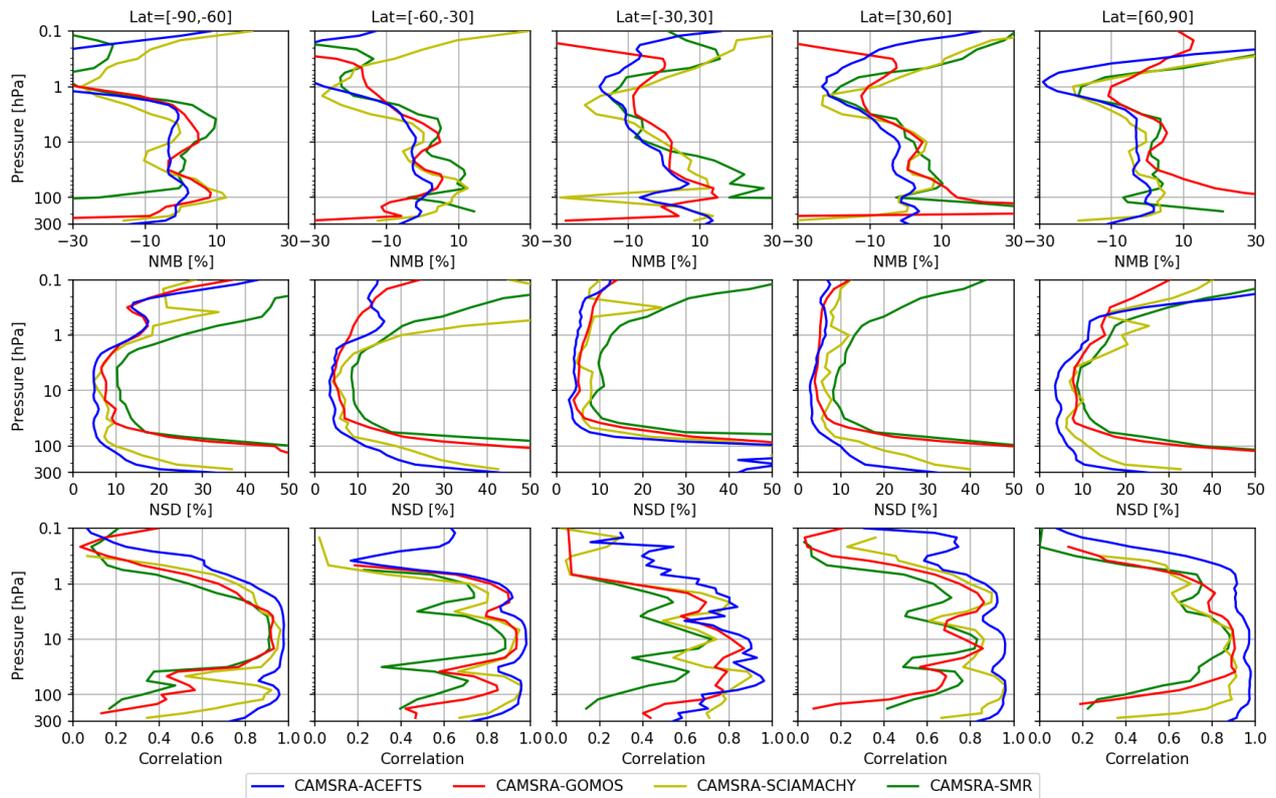


Fig. 8.3.6: As Fig. 8.3.5 but for the March-April-May (MAM) season.

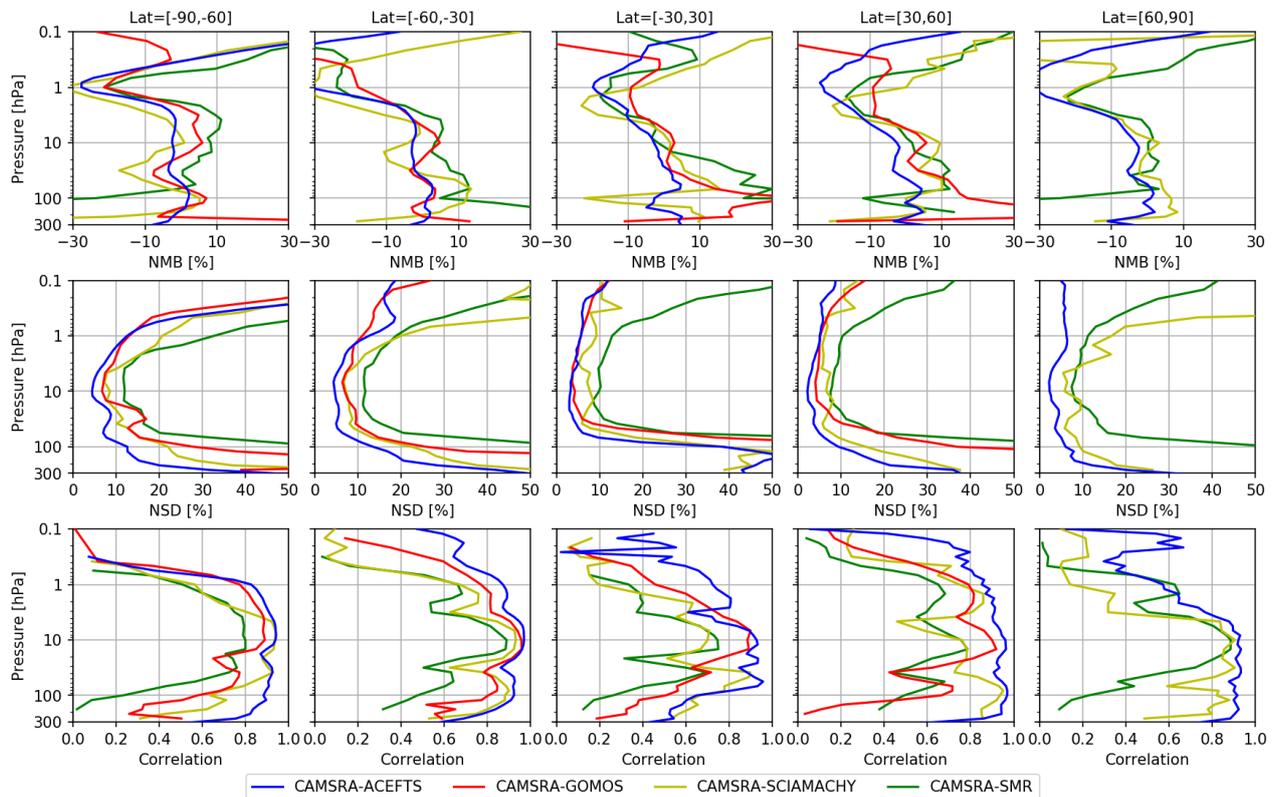


Fig. 8.3.7: As Fig. 8.3.5 but for the June-July-August (JJA) season.

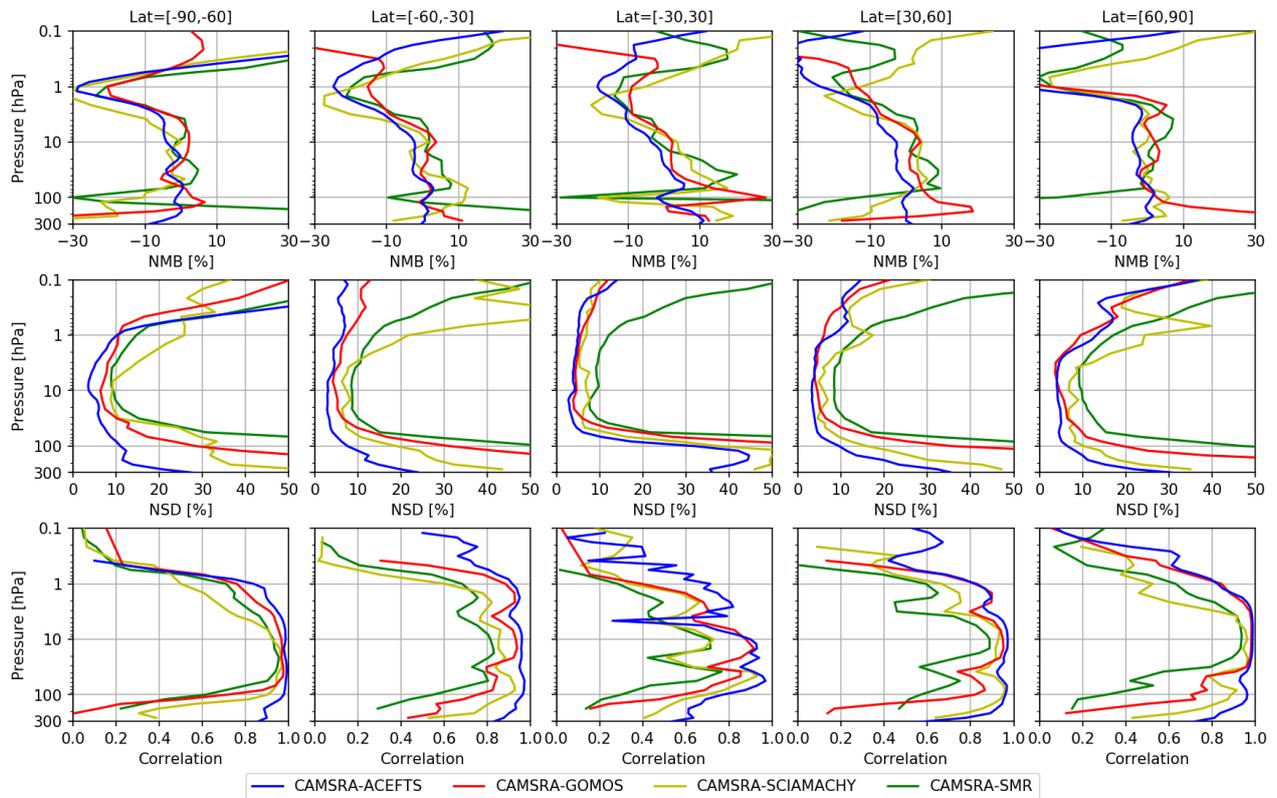


Fig. 8.3.8: As Fig. 8.3.5 but for the September-October-November (SON) season.

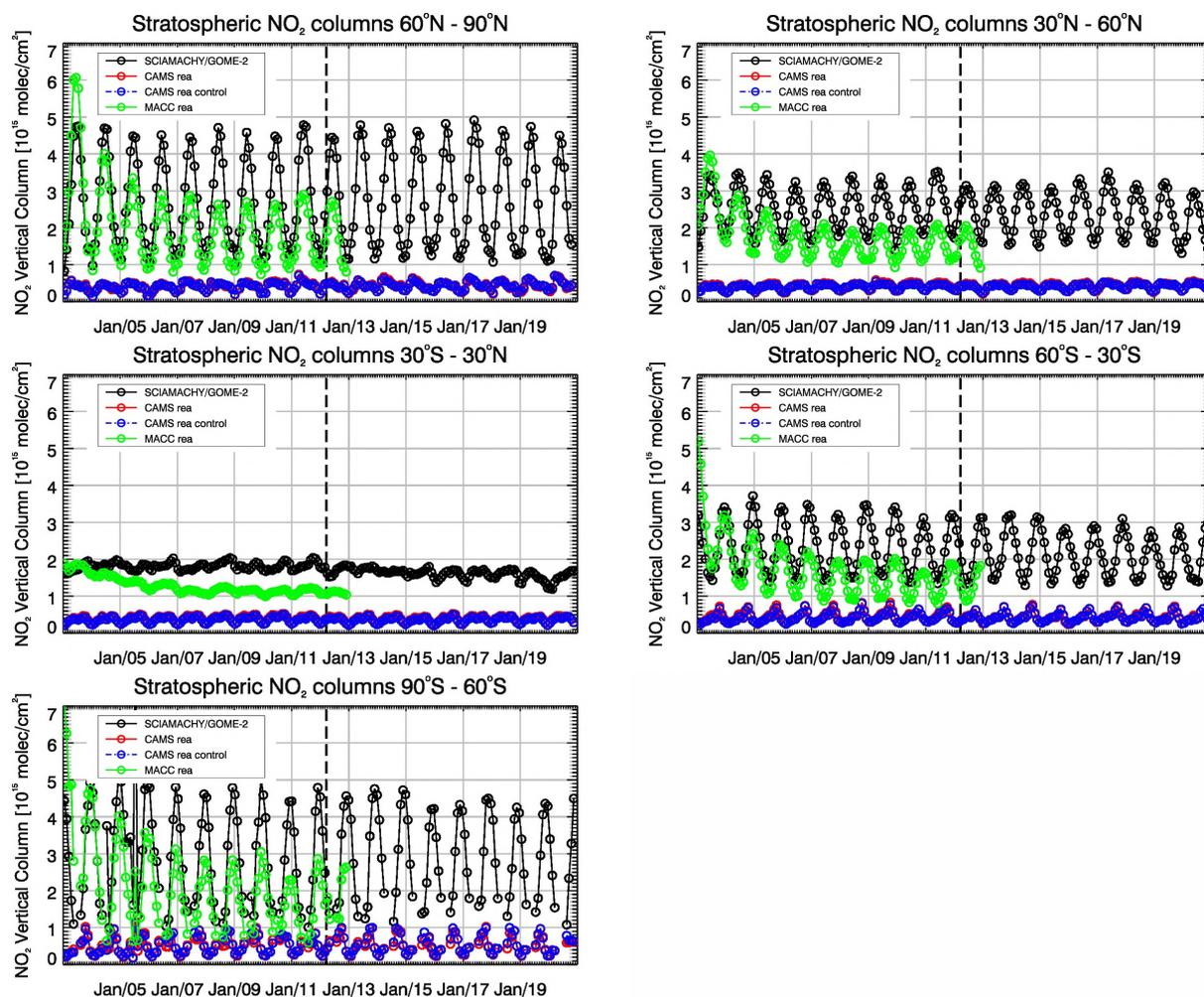


Fig. 8.4.1: Time series of average stratospheric NO<sub>2</sub> columns [ $10^{15}$  molec.  $\text{cm}^{-2}$ ] from SCIAMACHY/GOME-2A compared to model results for different latitude bands for the period 2003 to 2020. The switch from SCIAMACHY to GOME-2A is indicated by the vertical black dashed lines.

## 8.4 Stratospheric NO<sub>2</sub>

Due to the lack of stratospheric chemistry in the IFS-CB05 scheme, the only useful product in the stratosphere is ozone. Stratospheric NO<sub>2</sub> has also been evaluated but the results are only indicative.

Nitrogen dioxide from SCIAMACHY/Envisat satellite retrievals (IUP-UB v0.7) and GOME-2/MetOp-A satellite retrievals (IUP-UB v1.0) were compared to simulated stratospheric NO<sub>2</sub> columns. Time series for different latitude bands (Fig. 8.4.1) show that (while the MACC reanalysis performed better) the CAMS reanalysis fails to reproduce observed stratospheric NO<sub>2</sub> columns, due to the missing stratospheric chemistry in CAMS-IFS.



## 9. References

- Arosio, C., Rozanov, A., Malinina, E., Weber, M., and Burrows, J. P.: Merging of ozone profiles from SCIAMACHY, OMPS and SAGE II observations to study stratospheric ozone changes, *Atmos. Meas. Tech.*, 12, 2423–2444, <https://doi.org/10.5194/amt-12-2423-2019>, 2019.
- Benedetti, A., J.-J. Morcrette, O. Boucher, A. Dethof, R. J. Engelen, M. Fisher, H. Flentjes, N. Huneus, L. Jones, J. W. Kaiser, S. Kinne, A. Mangold, M. Razinger, A. J. Simmons, M. Suttie, and the GEMS-AER team: Aerosol analysis and forecast in the ECMWF Integrated Forecast System. Part II : Data assimilation, *J. Geophys. Res.*, 114, D13205, doi:10.1029/2008JD011115, 2009.
- Benedictow, A., A.M. Blechschmidt, I. Bouarar, E. Botek, S. Chabrilat, Y. Christophe, E. Cuevas, H. Clark, H. Flentje, A. Gaudel, J. Griesfeller, V. Huijnen, N. Huneus, L. Jones, J. Kapsomenakis, S. Kinne, B. Langerock, K. Lefever, M. Razinger, A. Richter, M. Schulz, W. Thomas, V. Thouret, M. Vrekoussis, A. Wagner, C. Zerefos: Validation report of the MACC reanalysis of global atmospheric composition, MACC-II Deliverable D\_83.6 (available from [http://macc.copernicus-atmosphere.eu/services/agac/global\\_verification/validation\\_reports](http://macc.copernicus-atmosphere.eu/services/agac/global_verification/validation_reports))
- Bennouna, Y., Schulz, M., Y. Christophe, H.J. Eskes, S. Basart, A. Benedictow, A.-M. Blechschmidt, S. Chabrilat, H. Clark, E. Cuevas, H. Flentje, K.M. Hansen, U. Im, J. Kapsomenakis, B. Langerock, K. Petersen, A. Richter, N. Sudarchikova, V. Thouret, A. Wagner, Y. Wang, T. Warneke, C. Zerefos, Validation report of the CAMS global Reanalysis of aerosols and reactive gases, years 2003-2019, Copernicus Atmosphere Monitoring Service (CAMS) report, CAMS84\_2018SC2\_D5.1.1-2019.pdf, April 2020, doi:10.24380/2v3p-ab79.
- Cammas, J.P., Brioude J., Chaboureaud 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 Teyssède, H.: A revised linear ozone photochemistry parameterization for use in transport and general circulation models: multi-annual simulations, *Atmos. Chem. Phys.*, 7, 2183-2196, doi:10.5194/acp-7-2183-2007, 2007.
- Ciais, P., M. Reichstein, N. Viovy, A. Granier, J. Ogee, V. Allard, M. Aubinet, N. Buchmann, C. Bernhofer, A. Carrara, F. Chevallier, N. De Noblet, A.D. Friend, P. Friedlingstein, T. Grunwald, B. Heinesch, P. Keronen, A. Knohl, G. Krinner, D. Loustau, G. Manca, G. Matteucci, F. Miglietta, J.M. Ourcival, D. Papale, K. Pilegaard, S. Rambal, G. Seufert, J.F. Soussana, M.J. Sanz, E.D. Schulze, T. Vesala, and R. Valentini, Europe-wide reduction in primary productivity caused by the heat and drought in 2003, *Nature*, 437 (7058), 529-533, 2005.
- Clerbaux, C., Boynard, A., Clarisse, L., George, M., Hadji-Lazaro, J., Herbin, H., Hurtmans, D., Pommier, M., Razavi, A., Turquety, S., Wespes, C., and Coheur, P.-F.: Monitoring of atmospheric composition using the thermal infrared IASI/MetOp sounder, *Atmos. Chem. Phys.*, 9, 6041-6054, <https://doi.org/10.5194/acp-9-6041-2009>, 2009.
- Cuevas, E., Milford, C., Barreto, A., Bustos, J. J., García, R. D., Marrero, C. L., Prats, N., Bayo, C., Ramos, R., Terradellas, E., Suárez, D., Rodríguez, S., de la Rosa, J., Vilches, J., Basart, S., Werner, E., López-Villarrubia, E., Rodríguez-Mireles, S., Pita Toledo, M. L., González, O., Belmonte, J., Puigdemunt, R., Lorenzo, J.A., Oromí, P., and del Campo-Hernández, R.: Desert Dust Outbreak in the Canary Islands (February 2020): Assessment and Impacts. (Eds. Cuevas, E., Milford, C. and Basart, S.), State Meteorological Agency (AEMET), Madrid, Spain and World Meteorological Organization, Geneva, Switzerland, WMO Global Atmosphere Watch (GAW) Report No. 259, WWRP 2021-1, 2021.



- 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.
- Deshler, T., J.L. Mercer, H.G.J. Smit, R. Stubi, G. Levrat, B.J. Johnson, S.J. Oltmans, R. Kivi, A.M. Thompson, J. Witte, J. Davies, F.J. Schmidlin, G. Brothers, T. Sasaki (2008) Atmospheric comparison of electrochemical cell ozonesondes from different manufacturers, and with different cathode solution strengths: The Balloon Experiment on Standards for Ozonesondes. *J. Geophys. Res.* 113, D04307, doi:10.1029/2007JD008975
- Emmons, L. K., D. P. Edwards, M. N. Deeter, J. C. Gille, T. Campos, P. Nédélec, P. Novelli, and G. Sachse, Measurements of Pollution In The Troposphere (MOPITT) validation through 2006 *Atmos. Chem. Phys.*, 9, 1795-1803, 2009
- Errera, Q., Daerden, F., Chabrilat, S., Lambert, J. C., Lahoz, W. A., Viscardy, S., Bonjean, S., and Fonteyn, D., 4D-Var Assimilation of MIPAS chemical observations: ozone and nitrogen dioxide analyses, *Atmos. Chem. Phys.*, 8, 6169-6187, 2008.
- Eskes, H., Huijnen, V., Arola, A., Benedictow, A., Blechschmidt, A.-M., Botek, E., Boucher, O., Bouarar, I., Chabrilat, S., Cuevas, E., Engelen, R., Flentje, H., Gaudel, A., Griesfeller, J., Jones, L., Kapsomenakis, J., Katragkou, E., Kinne, S., Langerock, B., Razinger, M., Richter, A., Schultz, M., Schulz, M., Sudarchikova, N., Thouret, V., Vrekoussis, M., Wagner, A., and Zerefos, C.: Validation of reactive gases and aerosols in the MACC global analysis and forecast system, *Geosci. Model Dev.*, 8, 3523-3543, doi:10.5194/gmd-8-3523-2015, 2015.
- Eskes, H.J., S. Basart, A. Benedictow, Y. Bennouna, A.-M. Blechschmidt, S. Chabrilat, 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, doi: 10.24380/7cjp-dn95.
- 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.
- 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
- Hubert, D., Lambert, J.-C., Verhoelst, T., Granville, J., Keppens, A., Baray, J.-L., Bourassa, A. E., Cortesi, U., Degenstein, D. A., Froidevaux, L., Godin-Beekmann, S., Hoppel, K. W., Johnson, B. J., Kyrölä, E., Leblanc, T., Lichtenberg, G., Marchand, M., McElroy, C. T., Murtagh, D., Nakane, H., Portafaix, T., Querel, R., Russell III, J. M., Salvador, J., Smit, H. G. J., Stebel, K., Steinbrecht, W., Strawbridge, K. B., Stübi, R., Swart, D. P. J., Taha, G., Tarasick, D. W., Thompson, A. M., Urban, J., van Gijsel, J. A. E., Van Malderen, R., von der Gathen, P., Walker, K. A., Wolfram, E., and Zawodny, J. M.: Ground-based assessment of the bias and long-term stability of 14 limb and occultation ozone profile data records, *Atmos. Meas. Tech.*, 9, 2497-2534, doi:10.5194/amt-9-2497-2016, 2016.



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.

Huijnen, V., H.J. Eskes, A. Wagner, M. Schulz, Y. Christophe, M. Ramonet, S. Basart, A. Benedictow, A.-M. Blechschmidt, S. Chabrillat, H. Clark, E. Cuevas, H. Flentje, K.M. Hansen, U. Im, J. Kapsomenakis, B. Langerock, A. Richter, N. Sudarchikova, V. Thouret, T. Warneke, C. Zerefos, *Validation report of the CAMS near-real-time global atmospheric composition service. System evolution and performance statistics; Status up to 1 June 2016*, Copernicus Atmosphere Monitoring Service (CAMS) report, CAMS84\_2015SC1\_D.84.1.4\_2016Q3\_201609, September 2016.

Huijnen, V., M. J. Wooster, J. W. Kaiser, D. L. A. Gaveau, J. Flemming, M. Parrington, A. Inness, D. Murdiyarso, B. Main & M. van Weele, *Fire carbon emissions over maritime southeast Asia in 2015 largest since 1997*, *Scientific Reports* volume 6, Article number: 26886, <https://doi.org/10.1038/srep26886>, 2016.

Inness, A., Baier, F., Benedetti, A., Bouarar, I., Chabrillat, S., Clark, H., Clerbaux, C., Coheur, P., Engelen, R. J., Errera, Q., Flemming, J., George, M., Granier, C., Hadji-Lazaro, J., Huijnen, V., Hurtmans, D., Jones, L., Kaiser, J. W., Kapsomenakis, J., Lefever, K., Leitão, J., Razinger, M., Richter, A., Schultz, M. G., Simmons, A. J., Suttie, M., Stein, O., Thépaut, J.-N., Thouret, V., Vrekoussis, M., Zerefos, C., and the MACC team: *The MACC reanalysis: an 8 yr data set of atmospheric composition*, *Atmos. Chem. Phys.*, 13, 4073-4109, <https://doi.org/10.5194/acp-13-4073-2013>, 2013.

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.

Inness, A., Ades, M., Agusti-Panareda, A., Barré, J., Benedictow, A., Blechschmidt, A.-M., Dominguez, J. J., Engelen, R., Eskes, H., Flemming, J., Huijnen, V., Jones, L., Kipling, Z., Massart, S., Parrington, M., Peuch, V.-H., Razinger, M., Remy, S., Schulz, M., and Suttie, M.: *The CAMS reanalysis of atmospheric composition*, *Atmos. Chem. Phys.*, 19, 3515-3556, <https://doi.org/10.5194/acp-19-3515-2019>, 2019.

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.

Koppe, C., S. Kovats, G. Jendritzky and B. Menne, "Heat-waves: risks and responses", *Health and Global Environmental Change SERIES*, No. 2, WHO, 2004.

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.

Livesey, N. J., Read, W. G., Wagner, P. A., Froidevaux, L., Lambert, A., Manney, G. L., Millán, L. F., Pumphrey, H. C., Santee, M. L., Schwartz, M. J., Wang, S., Fuller, R. A., Jarnot, R. F., Knosp, B. W., Martinez, E., and Lay, R. R.: *Earth Observing System (EOS) Aura Microwave Limb Sounder (MLS) Version 4.2x Level 2 and 3 data quality and description document*, Tech. Rep. D-33509 Rev. E, JPL, 2020.



- Lopez, M., M. Schmidt, M. Ramonet, J.L. Bonne, A. Colomb, V. Kazan, P. Laj, and J.M. Pichon, Three years of semicontinuous greenhouse gas measurements at the Puy de Dome station (central France), *Atmospheric Measurement Techniques*, 8 (9), 3941-3958, 2015.
- 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.
- Murtagh, D., Skyman, A., Rydberg, B., and Erikson, P.: Odin/SMR Product Validation and Evolution Report, version 1.3.0 (27 Oct. 2020). <https://odin.rss.chalmers.se/static/documents/PVER.pdf>
- Ramonet, M., B. Langerock, T. Warneke, H. J. Eskes, Validation report of the CAMS greenhouse gas global reanalysis, years 2003-2020, Copernicus Atmosphere Monitoring Service (CAMS) report CAMS84\_2018SC3\_D5.1.2-2020.pdf, April 2021.
- Raspolini et al.: Improvements in MIPAS ESA V7 and V8 products and trends, EUMETSAT Meteorological Satellite Conference 2017, 2017.  
[https://www.researchgate.net/publication/326294673\\_Improvements\\_in\\_MIPAS\\_ESA\\_V7\\_and\\_V8\\_products\\_and\\_trends](https://www.researchgate.net/publication/326294673_Improvements_in_MIPAS_ESA_V7_and_V8_products_and_trends)
- 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 NO<sub>2</sub> retrieval for the GOME-2 satellite instrument, *Atmos. Meas. Tech.*, 4, 1147-1159, doi:10.5194/amt-4-1147-2011, 2011.
- Sindelarova, K., Granier, C., Bouarar, I., Guenther, A., Tilmes, S., Stavrou, T., Müller, J.-F., Kuhn, U., Stefani, P., and Knorr, W.: Global data set of biogenic VOC emissions calculated by the MEGAN model over the last 30 years, *Atmos. Chem. Phys.*, 14, 9317-9341, doi:10.5194/acp-14-9317-2014, 2014.
- Smit, H.G.J., W. Straeter, B.J. Johnson, S.J. Oltmans, J. Davies, D.W. Tarasick, B. Hoegger, R. Stubi, F.J. Schmidlin, T. Northam, A.M. Thompson, J.C. Witte, I. Boyd: Assessment of the performance of ECC-ozonesondes under quasi-flight conditions in the environmental simulation chamber: Insights from the Juelich Ozone Sonde Intercomparison Experiment (JOSIE), *J. Geophys. Res.* 112, D19306, doi:10.1029/2006JD007308, 2007.
- Sofieva, V. F., Rahpoe, N., Tamminen, J., Kyrölä, E., Kalakoski, N., Weber, M., Rozanov, A., von Savigny, C., Laeng, A., von Clarmann, T., Stiller, G., Lossow, S., Degenstein, D., Bourassa, A., Adams, C., Roth, C., Lloyd, N., Bernath, P., Hargreaves, R. J., Urban, J., Murtagh, D., Hauchecorne, A., Dalaudier, F., van Roozendael, M., Kalb, N., and Zehner, C.: Harmonized dataset of ozone profiles from satellite limb and occultation measurements, *Earth Syst. Sci. Data*, 5, 349-363, doi:10.5194/essd-5-349-2013, 2013, <http://www.earth-syst-sci-data.net/5/349/2013/>.
- Sofieva, V. F., Ialongo, I., Hakkarainen, J., Kyrölä, E., Tamminen, J., Laine, M., Hubert, D., Hauchecorne, A., Dalaudier, F., Bertaux, J.-L., Fussen, D., Blanot, L., Barrot, G., and Dehn, A.: Improved GOMOS/Envisat ozone retrievals in the upper troposphere and the lower stratosphere, *Atmos. Meas. Tech.*, 10, 231-246, <https://doi.org/10.5194/amt-10-231-2017>, 2017.
- Stauffer, R. M., Thompson, A. M., Kollonige, D. E., Witte, J. C., Tarasick, D. W., Davies, J., et al. (2020). A post-2013 dropoff in total ozone at a third of global ozonesonde stations: Electrochemical concentration cell instrument artifacts? *Geophysical Research Letters*, 47, e2019GL086791. <https://doi.org/10.1029/2019GL086791>



- Stein, O., Schultz, M. G., Bouarar, I., Clark, H., Huijnen, V., Gaudel, A., George, M., and Clerbaux, C.: On the wintertime low bias of Northern Hemisphere carbon monoxide found in global model simulations, *Atmos. Chem. Phys.*, **14**, 9295–9316, doi:10.5194/acp-14-9295-2014, 2014.
- Tressol, M., C. Ordonez, R. Zbinden, J. Brioude, V. Thouret, C. Mari, P. Nedelec, J.-P. Cammas, H. Smit, H.-W. Patz, and A. Volz-Thomas, Air pollution during the 2003 European heat wave as seen by MOZAIC airliners, *Atmos. Chem. Phys.*, **8**, 2133–2150, doi:10.5194/acp-8-2133-2008, 2008.
- 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/](http://www.atmos-chem-phys.net/10/11277/2010/), 2010
- 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.
- von Clarmann, T., Höpfner, M., Kellmann, S., Linden, A., Chauhan, S., Funke, B., Grabowski, U., Glatthor, N., Kiefer, M., Schiefer-decker, T., Stiller, G. P., and Versick, S.: Retrieval of temperature, H<sub>2</sub>O, O<sub>3</sub>, HNO<sub>3</sub>, CH<sub>4</sub>, N<sub>2</sub>O, ClONO<sub>2</sub> and ClO from MIPAS reduced resolution nominal mode limb emission measurements, *Atmos. Meas. Tech.*, **2**, 159–175, <https://doi.org/10.5194/amt-2-159-2009>, <http://www.atmos-meas-tech.net/2/159/2009/>, 2009.
- 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.
- Wagner, A., Bennouna, Y., Blechschmidt, A.-M., Brasseur, G., Chabrillat, S., Christophe, Y., Errera, Q., Eskes, H., Flemming, J., Hansen, K. M., Inness, A., Kapsomenakis, J., Langerock, B., Richter, A., Sudarchikova, N., Thouret, V., Zerefos, C.: Comprehensive evaluation of the Copernicus Atmosphere Monitoring Service (CAMS) reanalysis against independent observations: Reactive gases. *Elementa: Science of the Anthropocene* (2021) **9** (1): 00171, <https://doi.org/10.1525/elementa.2020.00171>, 2021.
- 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.
- Williams, J. E., van Velthoven, P. F. J., and Brenninkmeijer, C. A. M.: Quantifying the uncertainty in simulating global tropo- spheric composition due to the variability in global emission es- timates of Biogenic Volatile Organic Compounds, *Atmos. Chem. Phys.*, **13**, 2857–2891, doi:10.5194/acp-13-2857-2013, 2013.
- Wunch, D., G. C. Toon, J.-F. L. Blavier, R. A. Washenfelder, J. Notholt, B. J. Connor, D. W. T. Griffith, V. Sherlock, and P. O. Wennberg, The total carbon column observing network, *Philosophical Transactions of the Royal Society - Series A: Mathematical, Physical and Engineering Sciences*, **369**(1943), 2087–2112, doi:10.1098/rsta.2010.0240, 2011.
- Wunch, D., Toon, G. C., Sherlock, V., Deutscher, N. M., Liu, C., Feist, D. G., & Wennberg, P. O., 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, 2015.



## 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***

Yasmine Bennouna, Valerie Thouret, CNRS-LA (editor, MOZAIC, IAGOS)  
Annette Wagner, MPG (O<sub>3</sub> sondes, GAW data)  
Harald Flentje, DWD (O<sub>3</sub> sondes, GAW data)  
Anne Blechschmidt and Andreas Richter, IUB Bremen (SCIAMACHY and GOME-2 NO<sub>2</sub>, HCHO)  
John Kapsomenakis, Christos Zerefos, AA (ESRL)  
Natalia Sudarchikova and Katinka Petersen, satellite IR observations (MPG)  
Kaj Hansen, Ulas Im, AU (Arctic theme)  
Bavo Langerock, BIRA (NDACC)

### ***Tropospheric aerosol***

Michael Schulz, MetNo (editor, Aerocom, Aeronet)  
Anna Benedictow, Jan Griesfeller, MetNo (Aerocom, Aeronet)  
Sara Basart, M. Teresa Pay, Oriol Jorba, BSC-CNS (Aeronet, MODIS, AirBase, SDS-WAS NAMEE RC)  
Emilio Cuevas, AEMET (Aeronet, MODIS, AirBase, SDS-WAS NAMEE RC)

### ***Stratosphere***

Quentin Errera, Yves Christophe, Simon Chabrilat, BIRA (satellite observations)  
Annette Wagner, Natalia Sudarchikova, MPI-M (O<sub>3</sub> sondes)  
Bavo Langerock, BIRA (NDACC FTIR, MWR, UVVIS DOAS, LIDAR)  
Anne Blechschmidt and Andreas Richter, IUB-UB Bremen (SCIAMACHY NO<sub>2</sub>)

### ***Reactive gases and aerosol modeling***

Antje Inness, Zak Kipling, Melanie Ades, Johannes Flemming, Richard Engelen (ECMWF)  
Vincent Huijnen (KNMI), Samuel Remy (LMD)



### ***Acknowledgements for the validation datasets used***

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, David Giles 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 thanks go to the 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 acknowledge the European Environment Information and Observation Network (Eionet) Air Quality portal which 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.

Also, we want to acknowledge the PM10 observations obtained from the openair project (<http://www.openair-project.org/>) which is a Natural Environment Research Council (NERC) knowledge exchange project that aims to provide a collection of open-source tools for the analysis of air pollution data.

The data used in this publication were obtained as part of the Network for the Detection of Atmospheric Composition Change (NDACC) and are publicly available (see <http://www.ndacc.org>).

We wish to thank the NDACC investigators for the provision of observations at Ny Alesund, Bern, Jungfraujoch, Izaña, Reunion Mado, Hohenpeissen, Mauna Loa, Lauder.

We wish to thank ESA for providing the HARMonized dataset of Ozone profiles (HARMOZ).

We acknowledge the strong support of the European Commission, Airbus, and the airlines (Lufthansa, Air France, Austrian, Air Namibia, Cathay Pacific, Iberia and China Airlines so far) which have carried the MOZAIC or IAGOS equipment and undertaken maintenance since 1994. In its 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 is supported by AERIS (CNES and INSU-CNRS).

The authors acknowledge the provision of Global Atmosphere Watch (GAW) data by the World Data Centre for Greenhouse Gases (WDCGG).

We wish to acknowledge the provision of ozone sonde data by the World Ozone and Ultraviolet Radiation Data Centre established at EC in Toronto (<https://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 (<http://www.nilu.no/>) and by the National Aeronautics and Space Administration NASA (<https://tropo.gsfc.nasa.gov/shadoz/>).



The authors acknowledge the European Monitoring and Evaluation Programme (EMEP; <http://www.emep.int/>) 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.

We acknowledge the European Space Agency (ESA) for providing the consolidated and harmonized ozone datasets through its Chemistry-Climate Initiative (HARMOZ O3 CCI; <http://www.esa-ozone-cci.org/>). Besides the European instruments aboard Envisat (MIPAS, GOMOS, SCIAMACHY), this includes datasets by Third-Party Missions which were originally provided by NASA (HALOE, SAGE) and by the University of Saskatchewan, Canada (ODIN/OSIRIS; <http://osirus.usas.ca/>).

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 providers of the aircraft campaign data used in this report (see Table 7.1.1 to 7.1.3) are gratefully acknowledged.

We wish to thank the TCCON investigators for making available the CO observations. References for the used TCCON data are:

*Blumenstock, T., Hase, F., Schneider, M., Garcia, O. E., & Sepulveda, E. (2014). TCCON data from Izana (ES), Release GGG2014R0. Pasadena, California. doi:10.14291/tcon.ggg2014.izana01.R0/1149295*

*De Mazière, M., Sha, M. K., Desmet, F., Hermans, C., Scolas, F., Kumps, N., . . . Cammas, J.-P. (2014). TCCON data from Reunion Island (RE), Release GGG2014R0. Pasadena, California. doi:10.14291/tcon.ggg2014.reunion01.R0/1149288*

*Deutscher, N. M., Notholt, J., Messerschmidt, J., Weinzierl, C., Warneke, T., Petri, C., . . . Katrynski, K. (2014). TCCON data from Bialystok (PL), Release GGG2014R1. Pasadena, California. doi:10.14291/tcon.ggg2014.bialystok01.R1/1183984*

*Dubey, M., Henderson, B., Green, D., Butterfield, Z., Keppel-Aleks, G., Allen, N., . . . Lindenmaier, R. (2014). TCCON data from Manaus (BR), Release GGG2014R0. Pasadena, California. doi:10.14291/tcon.ggg2014.manaus01.R0/1149274*

*Dubey, M., Lindenmaier, R., Henderson, B., Green, D., Allen, N., Roehl, C., . . . Wunch, D. (2014). TCCON data from Four Corners (US), Release GGG2014R0. Pasadena, California. doi:10.14291/tcon.ggg2014.fourcorners01.R0/1149272*

*Feist, D. G., Arnold, S. G., John, N., & Geibel, M. C. (2014). TCCON data from Ascension Island (SH), Release GGG2014R0. Pasadena, California. doi:10.14291/tcon.ggg2014.ascension01.R0/1149285*

*Goo, T.-Y., Oh, Y.-S., & Velazco, V. A. (2014). TCCON data from Anmeyondo (KR), Release GGG2014R0. Pasadena, California. doi:10.14291/tcon.ggg2014.anmeyondo01.R0/1149284*

*Griffith, D. W., Deutscher, N. M., Velazco, V. A., Wennberg, P. O., Yavin, Y., Aleks, G. K., . . . Bryant, G. (2014). TCCON data from Darwin (AU), Release GGG2014R0. Pasadena, California. doi:10.14291/tcon.ggg2014.darwin01.R0/1149290*



- Griffith, D. W., Velazco, V. A., Deutscher, N. M., Murphy, C., Jones, N., Wilson, S., . . . Riegenbach, M. (2014). TCCON data from Wollongong (AU), Release GGG2014R0. Pasadena, California. doi:10.14291/tccon.ggg2014.wollongong01.R0/1149291
- Hase, F., Blumenstock, T., Dohe, S., Gross, J., & Kiel, M. (2014). TCCON data from Karlsruhe (DE), Release GGG2014R0. Pasadena, California. doi:10.14291/tccon.ggg2014.karlsruhe01.R0/1149270
- Iraci, L. T., Podolske, J., Hillyard, P. W., Roehl, C., Wennberg, P. O., Blavier, J.-F., . . . Albertson, R. (2016). TCCON data from Edwards (US), Release GGG2014R1. Pasadena, California. doi:10.14291/tccon.ggg2014.edwards01.R1/1255068
- Iraci, L. T., Podolske, J., Hillyard, P. W., Roehl, C., Wennberg, P. O., Blavier, J.-F., . . . Barney, J. (2016). TCCON data from Indianapolis (US), Release GGG2014R1. Pasadena, California. doi:10.14291/tccon.ggg2014.indianapolis01.R1/1330094
- Kawakami, S., Ohyama, H., Arai, K., Okumura, H., Taura, C., Fukamachi, T., & Sakashita, M. (2014). TCCON data from Saga (JP), Release GGG2014R0. Pasadena, California. doi:10.14291/tccon.ggg2014.saga01.R0/1149283
- Kivi, R., Heikkinen, P., & Kyrö, E. (2014). TCCON data from Sodankyla (FI), Release GGG2014R0. Pasadena, California. doi:10.14291/tccon.ggg2014.sodankyla01.R0/1149280
- Morino, I., Matsuzaki, T., & Shishime, A. (2014). TCCON data from Tsukuba (JP), 125HR, Release GGG2014R1. Pasadena, California. doi:10.14291/tccon.ggg2014.tsukuba02.R1/1241486
- Morino, I., Velazco, V. A., Akihiro, H., Osamu, U., & Griffith, D. W. (2018). TCCON data from Burgos, Ilocos Norte (PH), Release GGG2014.R0. Pasadena, California. doi:10.14291/tccon.ggg2014.burgos01.R0
- Morino, I., Yokozeki, N., Matzuzaki, T., & Shishime, A. (2014). TCCON data from Rikubetsu (JP), Release GGG2014R0. Pasadena, California. doi:10.14291/tccon.ggg2014.rikubetsu01.R0/1149282
- Notholt, J., Petri, C., Warneke, T., Deutscher, N. M., Buschmann, M., Weinzierl, C., . . . Grupe, P. (2014). TCCON data from Bremen (DE), Release GGG2014R0. Pasadena, California. doi:10.14291/tccon.ggg2014.bremen01.R0/1149275
- Notholt, J., Warneke, T., Petri, C., Deutscher, N. M., Weinzierl, C., Palm, M., & Buschmann, M. (2017). TCCON data from Ny\_Iesund, Spitsbergen (NO), Release GGG2014.R0. Pasadena, California. doi:10.14291/tccon.ggg2014.nyalesund01.R0/1149278
- Sherlock, V., Connor, B. J., Robinson, J., Shiona, H., Smale, D., & Pollard, D. (2014). TCCON data from Lauder (NZ), 125HR, Release GGG2014R0. Pasadena, California. doi:10.14291/tccon.ggg2014.lauder02.R0/1149298
- Strong, K., Mendonca, J., Weaver, D., Fogal, P., Drummond, J. R., Batchelor, R., & Lindenmaier, R. (2017). TCCON data from Eureka (CA), Release GGG2014R1. Pasadena, California. doi:10.14291/tccon.ggg2014.eureka01.R1/1325515
- Sussmann, R., & Rettinger, M. (2014). TCCON data from Garmisch (DE), Release GGG2014R0. Pasadena, California. doi:10.14291/tccon.ggg2014.garmisch01.R0/1149299
- Sussmann, R., & Rettinger, M. (2018). TCCON data from Zugspitze (DE), Release GGG2014R1. Pasadena, California. doi:10.14291/tccon.ggg2014.zugspitze01.R1
- Te, Y., Jeseck, P., & Janssen, C. (2014). TCCON data from Paris (FR), Release GGG2014R0. Pasadena, California. doi:10.14291/tccon.ggg2014.paris01.R0/1149279



- Warneke, T., Messerschmidt, J., Notholt, J., Weinzierl, C., Deutscher, N. M., Petri, C., . . . Parmentier, E. (2014). TCCON data from Orléans (FR), Release GGG2014R0. Pasadena, California. doi:10.14291/tccon.ggg2014.orleans01.R0/1149276
- Wennberg, P. O., Roehl, C., Wunch, D., Toon, G. C., Blavier, J.-F., Washenfelder, R. a., . . . Ayers, J. (2014). TCCON data from Park Falls (US), Release GGG2014R0. Pasadena, California. doi:10.14291/tccon.ggg2014.parkfalls01.R0/1149161
- Wennberg, P. O., Wunch, D., Roehl, C., Blavier, J.-F., Toon, G. C., & Allen, N. (2014). TCCON data from Caltech (US), Release GGG2014R1. Pasadena, California. doi:10.14291/tccon.ggg2014.pasadena01.R1/1182415
- Wennberg, P. O., Wunch, D., Roehl, C., Blavier, J.-F., Toon, G. C., Allen, N., . . . Martin, J. (2016). TCCON data from Lamont (US), Release GGG2014R1. Pasadena, California. doi:10.14291/tccon.ggg2014.lamont01.R1/1255070
- Wennberg, P. O., Wunch, D., Yavin, Y., Toon, G. C., Blavier, J.-F., Allen, N., & Keppel-Aleks, G. (2014). TCCON data from Jet Propulsion Laboratory (US), 2007, Release GGG2014R0. Pasadena, California. doi:10.14291/tccon.ggg2014.jpl01.R0/1149163
- Wunch, D., Mendonca, J., Colebatch, O., Allen, N., Blavier, J.-F. L., Roche, S., . . . Strong, K. (2017). TCCON data from East Trout Lake (CA), Release GGG2014R1. Pasadena, California. 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

