Copernicus Atmosphere Monitoring Service

Validation of the CAMS regional services: concentrations above the surface

Status update for the period June-August 2020

Issued by: KNMI

Date: 17 December 2020

Ref: CAMS84_2018SC2_D4.1.1-JJA2020.pdf

IMPLEMENTED BY

ECMWF

Validation of the CAMS regional services: concentrations above the surface

Status update for the period June-August 2020

AUTHORS:
D. Akritidis (AUTH), T. Antonakaki (AA), Y. Bennouna (CNRS-LA),
A.-M. Blechschmidt (IUP-UB), T. Bösch (IUP-UB), H. Clark (CNRS-LA),
C. Gielen (BIRA-IASB), F. Hendrick (BIRA-IASB), J. Kapsomenakis (AA),
S. Kartsios (AUTH), E. Katragkou (AUTH), D. Melas (AUTH), A. Mortier (MetNo),
E. Peters (IUP-UB), K. Petersen (MPI), A. Piters (KNMI), A. Richter (IUP-UB),
M. van Roozendael (BIRA-IASB), M. Schulz (MetNo), N. Sudarchikova (MPI),
A. Wagner (MPI), P. Zanis (AUTH), C. Zerefos (AA)

EDITORS:
J. Douros (KNMI), H.J. Eskes (KNMI)

**REPORT OF THE COPERNICUS ATMOSPHERE MONITORING SERVICE,
VALIDATION SUBPROJECT (CAMS-84).**

Executive Summary

The Copernicus Atmosphere Monitoring Service (CAMS, http://atmosphere.copernicus.eu) is a component of the European Earth Observation programme Copernicus. The CAMS service consists of two major forecast and analysis systems. First, the CAMS global near-real time (NRT) service provides daily analyses and forecasts of reactive trace gases, greenhouse gases and aerosol concentrations, and is based on the ECMWF Integrated Forecast System (called CAMS-global in this document). Secondly, nine models in Europe, perform air quality forecasts and analyses on a daily basis, nested within CAMS-global. Based on these individual forecasts and analyses, an ensemble forecast of air quality over Europe is produced and disseminated by Météo-France (called ENSEMBLE or CAMS-regional below). The regional members use the global forecasting results as boundary conditions at the sides and top of the domain.

This document reports on two validation activities, namely
- an evaluation of the consistency between the global and regional modelling components of CAMS, focussing on the boundaries of the regional domain, and
- an evaluation of the regional ENSEMBLE and the nine individual models contributing to the ensemble with independent observations, focusing on the concentrations above the surface.

The current analysis includes ozone (O_3), nitrogen dioxide (NO_2), aerosol (PM10/PM2.5/AOD) and carbon monoxide (CO) forecasts covering the period up to August 2020.

The forecasts from the regional models were compared with the following set of observations:
- aerosol lidar observations from the EARLINET network;
- aerosol AOD observations from the AERONET network;
- IAGOS routine aircraft measurements of ozone and CO;
- ozone sonde profiles;
- MAX-DOAS NO_2 tropospheric columns;
- GOME-2/MetOp-A NO_2 satellite tropospheric column retrievals (IUP-UB v1.0 product);
- high-altitude ozone surface stations;
- CO and O_3 from GAW mountain stations; and
- CO observations from the MOPITT satellite instrument.

These observations are available to CAMS within one month after the observations were made.

This report is based on regional model data available for the months May 2016 to August 2020, with a focus on June – August 2020 (JJA2020). The report is updated every 3 months. The main results are summarised below. This summary is focusing on the performance of the regional ENSEMBLE. Detailed results, also for the individual models, are presented in sections 3 to 11, and each of these sections starts with a summary of the main results. Model specific findings are summarised in section 12.
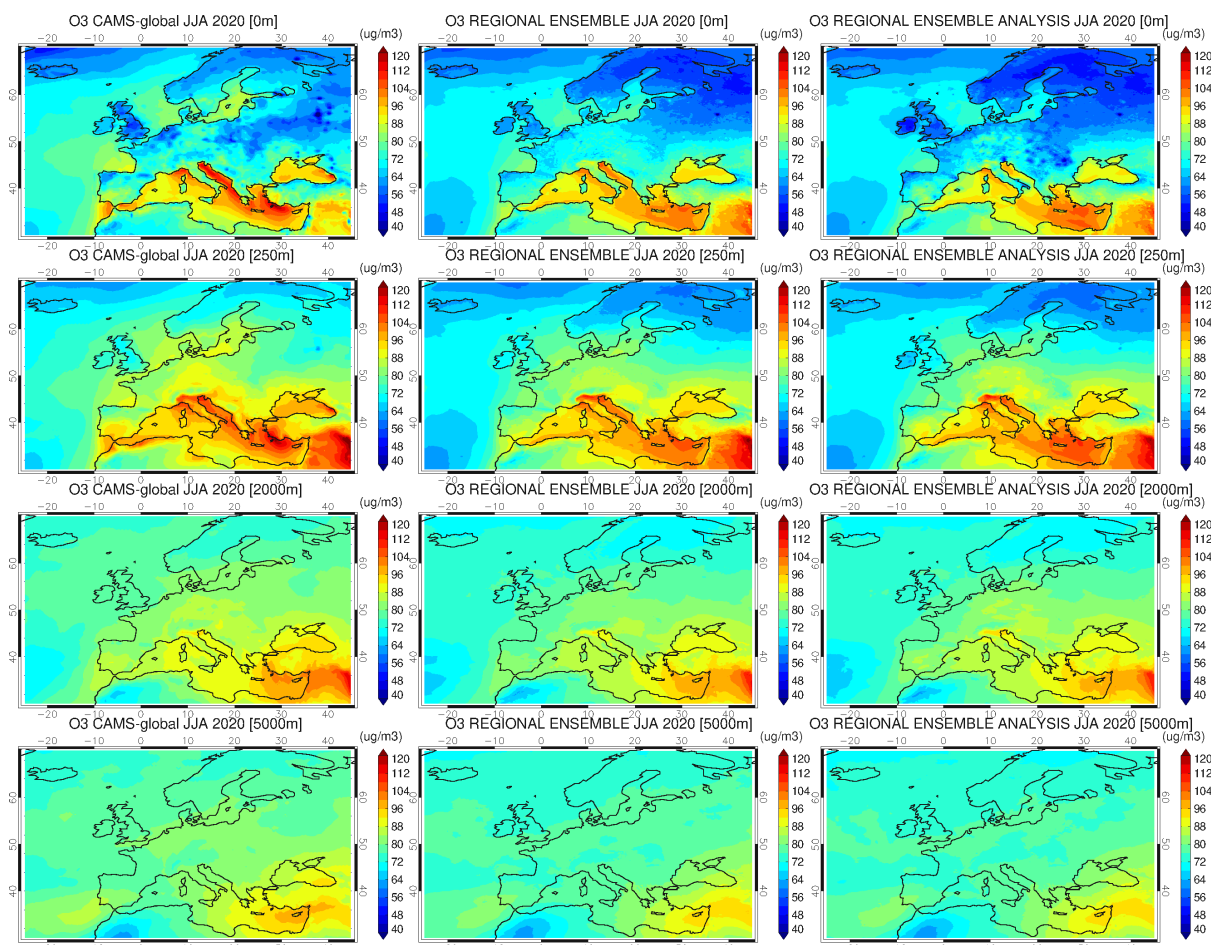
Figure S.1. CAMS global ozone forecast for day 1 (left), CAMS regional ENSEMBLE ozone forecast for day 1 (middle), CAMS regional ENSEMBLE ozone analysis (right). From top to bottom: 0, 250, 2000, 5000m altitude level. The results are averaged over the June - August 2020 period.

The last upgrade implemented in CAMS-global was on 9 July 2019, moving from 60 to 137 vertical levels, while the lust upgrade for the CAMS regional models happened on 4 February 2020, implementing new anthropogenic emissions and making available three new aerosol species.

General conclusions for the regional ENSEMBLE

The comparison of the European regional CAMS ENSEMBLE air quality forecasts and analyses against above-surface observations of O_3, NO_2, CO for the period up to 1st of September 2020 demonstrates that overall, the biases observed are small, often within the uncertainty of the validation approach while temporal correlations for ozone and CO are reasonable. Performance of the ENSEMBLE analysis product is found to be generally superior to that of the ENSEMBLE forecasts. The ENSEMBLE performs generally better than any of the individual models for ozone, NO_2 and CO, showing the strength of the ensemble approach adopted in CAMS.
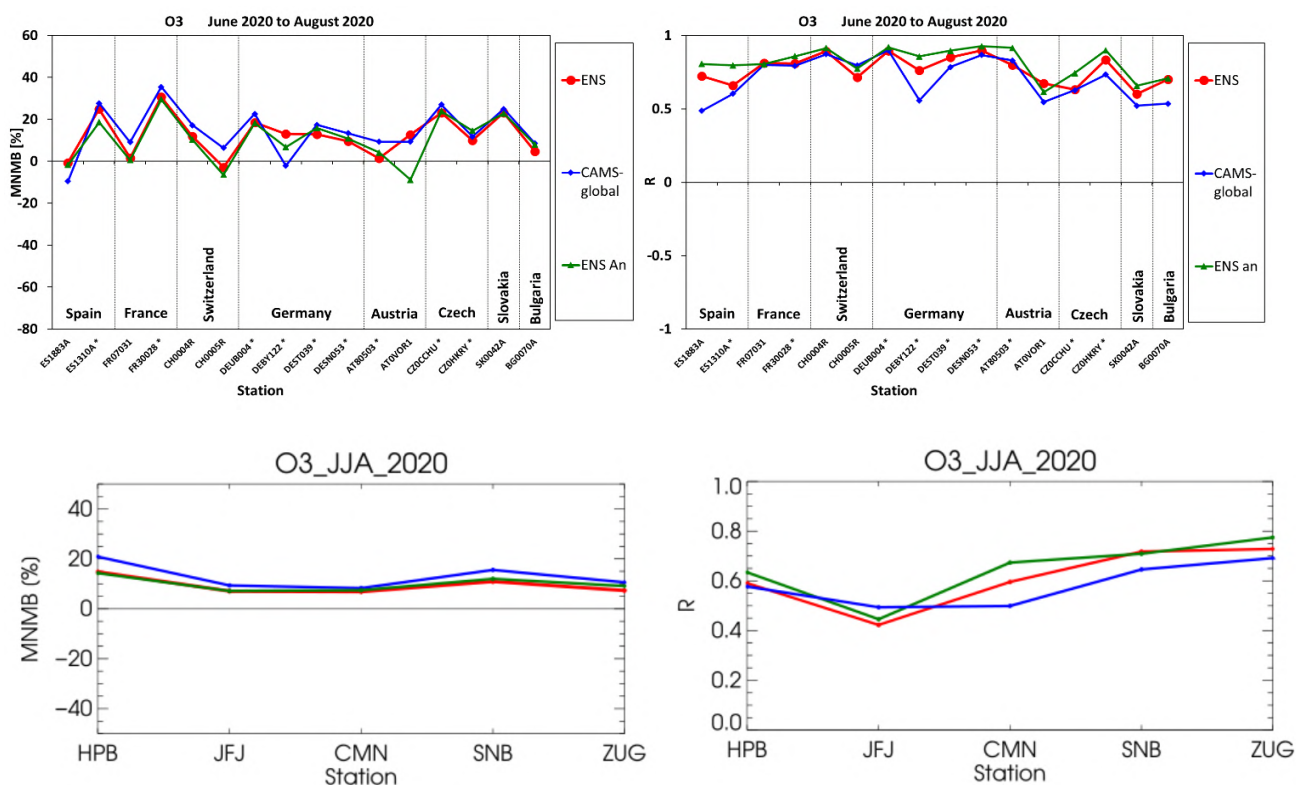
Figure S.2. Normalised bias (left) and correlation coefficient (right) for ozone for the high-altitude (above 1 km) EEA Air Quality e-reporting stations (top) and the 5 high-altitude European GAW stations (bottom). Lines represent ENSEMBLE forecast (solid red), ENSEMBLE analysis (solid green) and CAMS-global system (blue) for June - August 2020. The horizontal axis is the station identifier referring to Hohenpeissenberg (HPB), Jungfraujoch (JFJ), Monte Cimone (CMN), Sonnblick (SNB) and Zugspitze (ZUG).
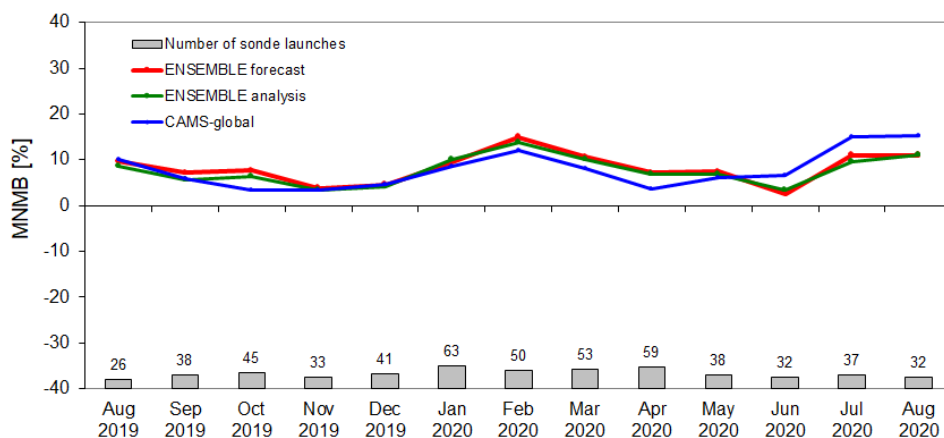

Figure S.3. Modified normalised mean bias (MNMB) against ozone sondes for the regional ENSEMBLE forecasts (red) and analyses (green) from August 2019 to August 2020 (horizontal axis). Ozone was averaged over the lower-middle free troposphere region, 500 hPa < p < 850 hPa.
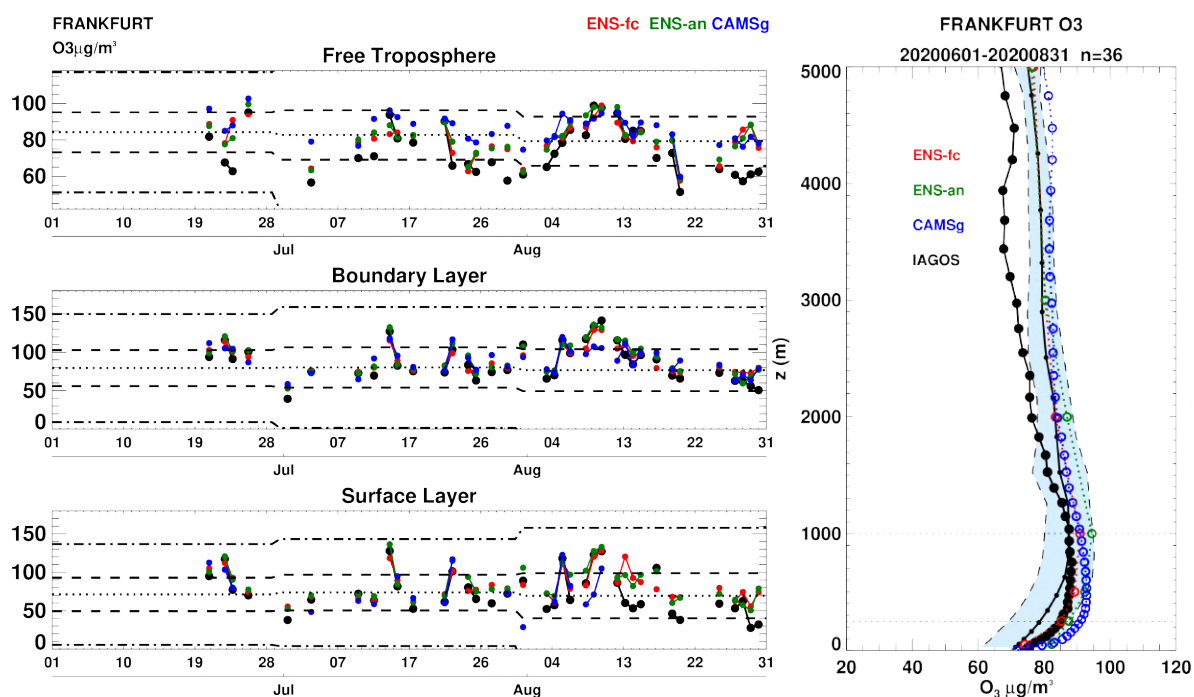
Figure S.4. Left: Daily time series of ozone at Frankfurt for June to August 2020 The black dotted line is the monthly mean of the observations over the period 2003-2016 (IAGOS/MOZAIC, Level 2 data), the black dashed line shows 1 standard deviation from the monthly mean and the black doted-dashed line shows 3 standard deviations from the monthly mean. Right: Mean profile of ozone at Frankfurt for the period June - August 2020. The shaded area indicates the range of the mean climatology of the observations plus/minus one standard deviation during the same period for all years between 2003 and 2016 (IAGOS/MOZAIC, level 2 data). In both panels, IAGOS observations are shown in black, the regional ENSEMBLE and associated analysis are shown in red and green respectively, and the global o-suite is shown in blue.

Ozone

At the boundaries of the regional domain, the ENSEMBLE agrees well with CAMS-global, indicating that the implementation of the boundary conditions was done properly. Over the full domain there is a good match between the global and regional CAMS analyses and forecasts between 1 and 5 km altitude. These results are similar to previous quarters.

The differences between the global and regional systems reveal themselves in the boundary layer and at the surface over land, as expected. A comparison of the regional analysis product with the regional day 1 forecast (Fig. S.1) shows some differences between the regional ENSEMBLE forecast and analysis, with the analysis generally having lower concentrations near the surface.

For high altitude stations, Fig. S.2, mostly an overestimation between up to 25% is observed for the ENSEMBLE first day forecast, while comparison against ozonesondes shows a milder overestimation in the 3-15% range (also Fig. S.3). Time correlations at high-altitude and GAW stations range between 0.45 and 0.93 during this period with noticeably improved figures for the ENSEMBLE analyses as compared to the ENSEMBLE forecast in the case 3 out of the 5 GAW stations.
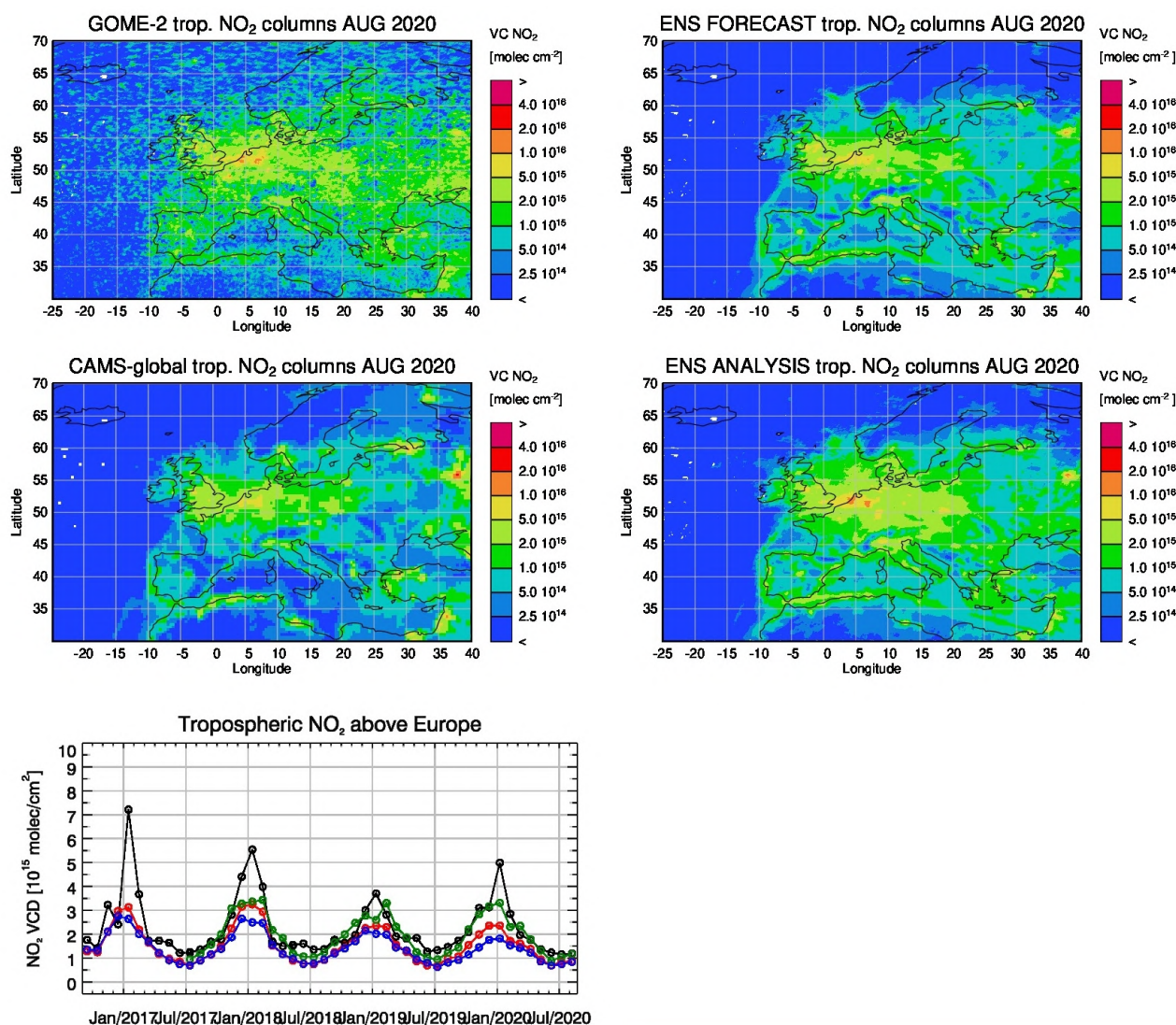
Figure S.5. Maps of satellite retrieved and model simulated tropospheric NO$_2$ columns [molecules cm^{-2}] for August 2020 for GOME-2 (top left), regional ENSEMBLE forecasts (top right) CAMS-global forecasts (middle left) and regional ENSEMBLE analyses (middle right). The panel at the bottom shows corresponding time series of average tropospheric NO$_2$ columns [10^{15} molecules cm^{-2}] from GOME-2 (black), regional ENSEMBLE forecasts (red), CAMS-global forecasts (blue) and regional ENSEMBLE analyses (green). GOME-2 data were gridded to regional model resolution (i.e. 0.1° x 0.1°). Model data were treated with the same reference sector (25°W - 20°E) subtraction approach as the satellite data and linearly interpolated to the satellite overpass time (9:30 LT).

Comparisons with IAGOS aircraft observations at Frankfurt, see Fig. S.4, reveals that ozone is mostly well reproduced by both the regional ENSEMBLE and CAMS-global in the low troposphere while large overestimations are obtained in the free troposphere for all models, in particular when ozone values are low. The comparison with the climatological variability shows that there is a positive anomaly of ozone in the low troposphere that was found in the previous quarter is not present in JJA2020.
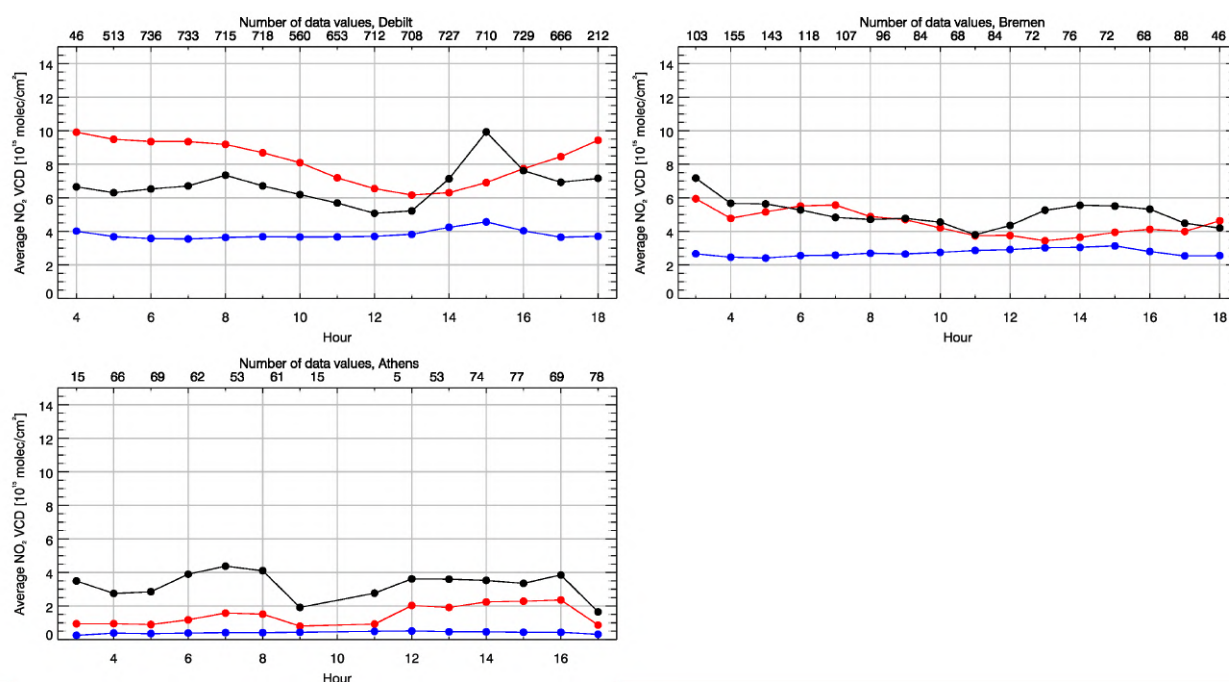
Figure S.6: Diurnal cycles (averages over hourly bins) of tropospheric NO_2 VCDs [10^{15} molec. cm^{-2}] from MAX-DOAS and models for De Bilt (the Netherlands) (top left), Bremen (Germany) top right and Athens (Greece) (bottom). The coloured lines show (black) MAX-DOAS retrievals, (red) regional ENSEMBLE forecasts and (blue) CAMS-global. Period: June – August 2020.

Nitrogen dioxide (NO_2)

The overall spatial distribution of tropospheric NO_2 as observed from space by GOME-2 is reproduced by the ENSEMBLE during JJA2020 (Fig. S.5) with better results by the analysis compared to the forecast. As described in previous reports, winter values over European emission hotspots simulated by the regional ENSEMBLE analysis, forecasts and CAMS-global show significantly smaller values than GOME-2, the observed inter-annual variability is therefore not reproduced. Values over European emission hotspots are better represented by the regional ENSEMBLE analysis and forecasts than by CAMS-global.

Systematic uncertainties in the retrievals (on average on the order of 20% – 30% over polluted regions) depend on the season, with winter values and especially January in mid and high latitudes normally associated with larger error margins. Conclusions may differ for comparisons to other satellite NO_2 products (e.g. TEMIS GOME-2, http://www.temis.nl). We note that since the CAMS-global upgrade of 26 June 2018, GOME-2 observations are assimilated by the global system. This is, however, a different retrieval product than the one used for validation reported here (University of Bremen retrieval).
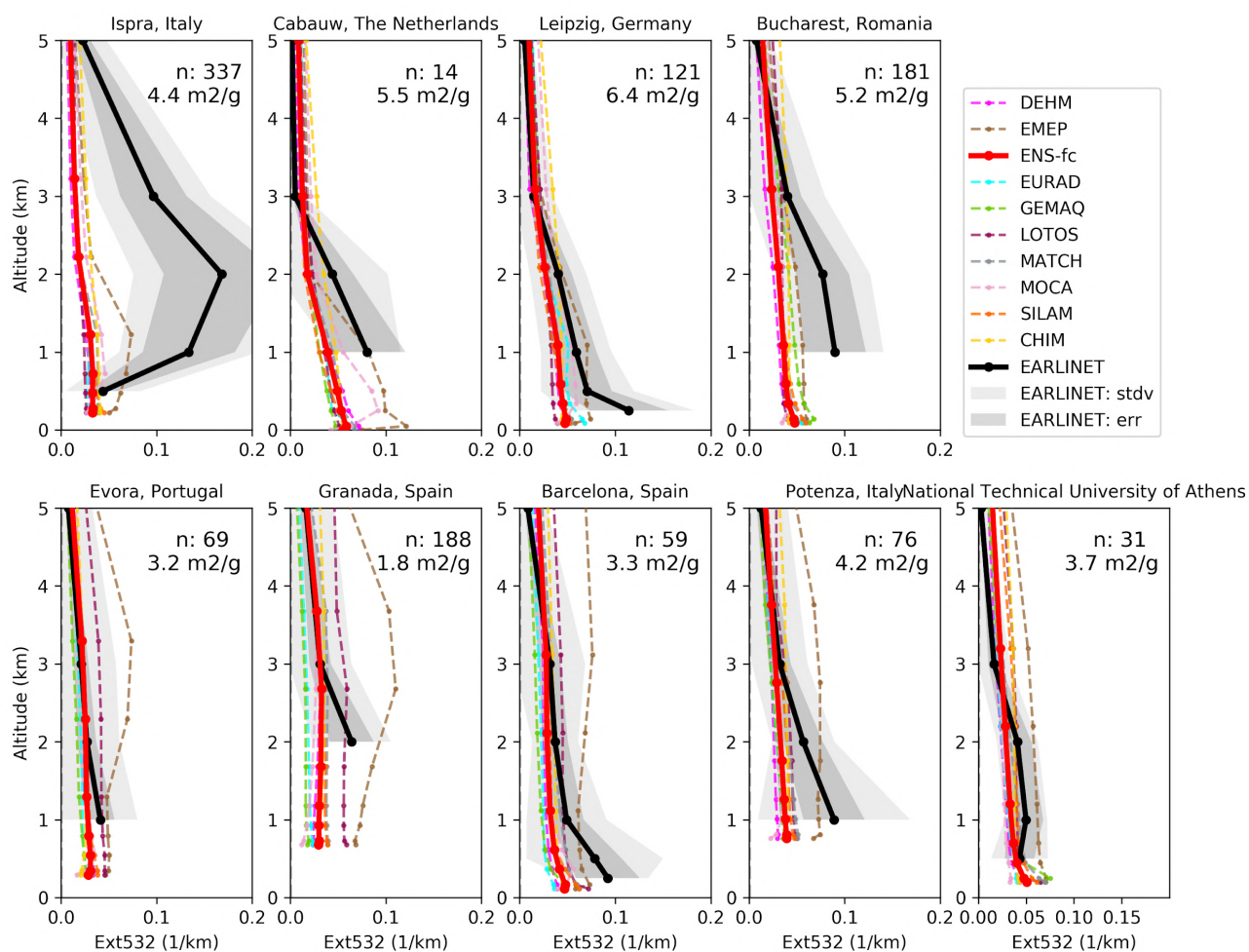
Figure S.7. Extinction profiles June - August 2020 derived from the ENSEMBLE forecast mass concentration profiles (red envelope) and from EARLINET (climatology) backscatter profiles (grey envelope: lidar ratio uncertainty, light grey: including sampling error). "n: XX means number of individual EARLINET profiles assembled (June - August 2006-2018). The EMC used for the calculation of the extinction from the concentration profiles is indicated for each station below the number of EARLINET profiles "n" used for the calculation of the climatology.
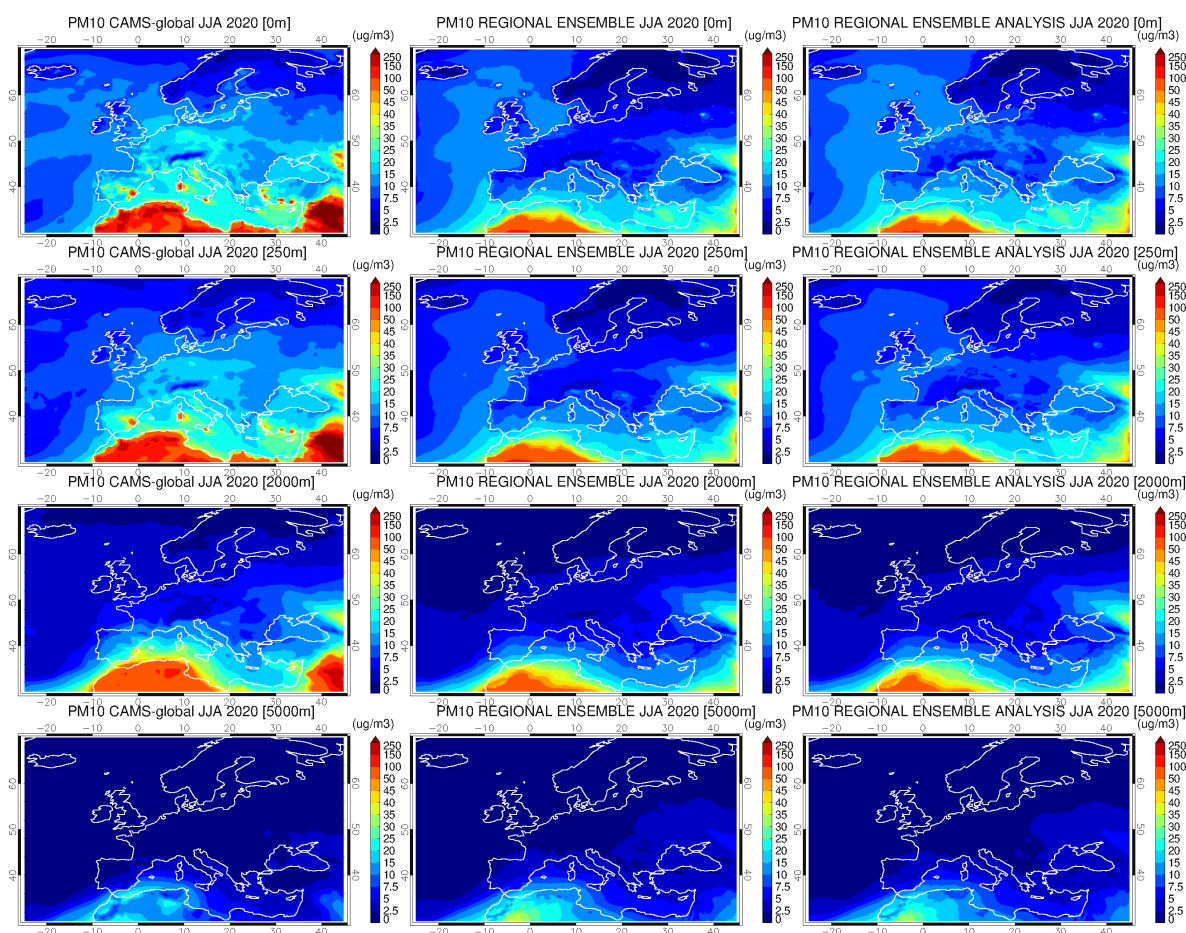
Figure S.8. CAMS global PM10 forecast for day 1 (left), CAMS regional ENSEMBLE PM10 forecast for day 1 (middle), CAMS regional PM10 analysis (right). From top to bottom: 0, 250, 2000, 5000m altitude level. Period: June to August 2020.

Comparisons to ground based remote sensing MAX-DOAS retrievals at three different European stations (see Figure S.6) for the time period of JJA 2020 show that regional ENSEMBLE forecasts are closer to the urban station observations of Bremen and Athens than CAMS-global, mainly attributed to the difference in spatial resolution. For De Bilt, CAMS-global performs better than the regional ENSEMBLE. However, this may change with the next reports when more months of data will be included. The performance of simulations for diurnal cycles of tropospheric NO_2 columns depends on the location, but generally shows a moderate performance for the ENSEMBLE products for Bremen and Athens.
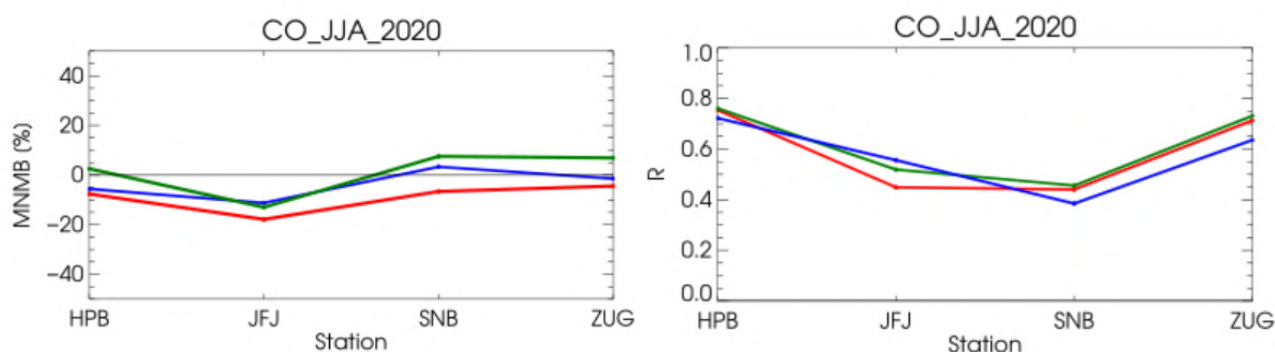
Figure S.9. MNMBs [%] (top) and correlation coefficients (bottom) for CO regional ENSEMBLE forecasts (red), analyses (green) and CAMS-global (blue) compared to observations at GAW stations. Period: June - August 2020.

Aerosol / PM

The regional models are compared with EARLINET climatological lidar profiles for the same season (data from 2006-2020), Fig. S.7. The standard dissemination of CAMS-regional forecasts does not include information on composition, size and humidity growth of the aerosol in the models. This introduces considerable uncertainty to the PM derived extinction, which conservatively spans a factor 10 for absolute extinction values. Relative differences among nearby stations and the form of extinction profiles are more certain. The order of magnitude in extinction is similar between the ENSEMBLE and the lidar profiles, with Ispra, Italy, close to the Alps being the most notable exception. The decrease in extinction with height seems to be generally steeper in the observations.

The PM10 concentrations in the regional ENSEMBLE forecasts and analyses are similar, but larger PM values are observed for the analysis over some areas of Central and Western Europe at the surface level, see Figure S.8. CAMS-global continues to show higher dust loads over the Mediterranean. An increased high bias in PM10-dust over the Mediterranean appeared with the upgrade of CAMS-global to 46R1 in July 2019 (Wagner et al., 2020). Also, notable hotspots are seen over Spain, Sardinia, Sicily and Turkey, similar to what was found during the previous quarters. The agreement between CAMS-global and the ENSEMBLE products is better for PM10 and PM2.5 in the upper layers compared to the lower layers, as expected. For PM2.5 the difference between the CAMS-global and CAMS-regional is more pronounced than for PM10 within the PBL especially for dust near the southern boundary and over the Middle East (exceeding 50 $\mu g/m^3$).

Carbon monoxide (CO)

Comparison at the GAW stations reveal biases between -12% and 2% for the analysis and between -17% and -4% for the forecast, while temporal correlations coefficients are between 0.45 and 0.76 for the analysis and between 0.44 and 0.75 for the forecast (Fig. S.9). The comparison against IAGOS aircraft observations shows that CO is mostly underestimated by all models, but more so by the regional ENSEMBLE products, especially in the free troposphere.
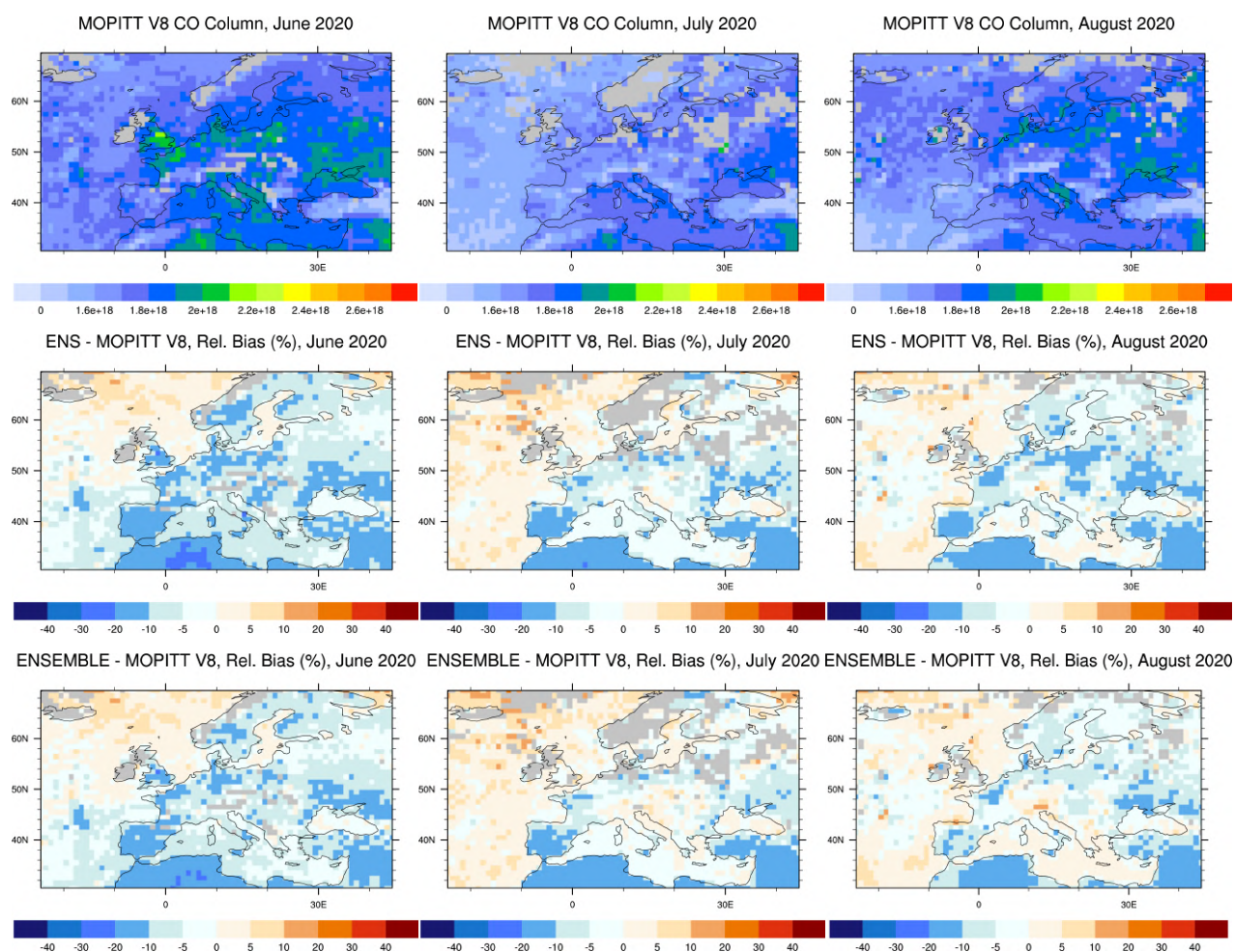
Figure S.10. CO total column for MOPITT v8 satellite retrievals (top row, in molecules/cm^2), relative difference between the regional ENSEMBLE forecasts and MOPITT (middle row) and regional ENSEMBLE analyses and MOPITT (bottom row) for June (left column), July (middle column) and August 2020 (right column). Grey colour indicates missing values.

Comparisons with MOPITT CO satellite observations (version 8, Fig. S.10) data also show an underestimation of CO values over the land part of the domain and overestimation over the ocean part are within 10% with some regional exceptions. The overestimated spot by 20% can be seen in the North of Italy in August 2020. The ENSEMBLE analysis data are very similar to the forecast data, except some overestimation (10%) in the middle part of domain in July-August and larger area with a positive bias over Northern Italy in August. Note that due to the COVID-19 measures in 2020 we may expect changes (reductions) in anthropogenic emissions which are not included in the emission inventories yet.

General conclusions for the spread in results among the individual regional models

Two new models, namely DEHM and GEM-AQ, are evaluated in this report. The performance of those two models is on a par with the other models, revealing no significant issues. Regional models largely appear to have efficiently implemented the use of global CAMS boundary conditions and provide 3D fields that are comparable to those of the global model at the boundaries and at elevated levels.

For ozone, comparisons with observations from IAGOS reveal a higher bias in the free troposphere and better performance at higher ozone values. Bias and correlation at GAW and EEA e-reporting high altitude surface stations show only a small spread among most of the models and an improved performance by the analysis product. Comparisons with ozonesondes provide an even more favourable picture as the spread is even smaller.

For NO_2, validation with MAX-DOAS shows as sizeable spread between individual models, more so when in the time series of the observations. However, most of the simulated values probably fall within the uncertainty range of the MAX-DOAS retrievals. The overall spatial distribution of tropospheric NO_2 as seen by GOME-2/MetOp-A is reproduced well by the regional models, but values over central European emission hotspots are often underestimated by most models during winter as compared to the GOME-2 retrievals.

For CO, comparisons with IAGOS for the individual models are in general very similar for all the profiles of this quarter. Some differences between the models reveal themselves in the surface and boundary layer for profiles with the highest CO values. At GAW mountain stations, bias for the individual models is mostly negative for the forecasts, but less so in the analyses, although the spread between the models is quite high, especially for the correlation coefficient. Comparisons with MOPITT retrievals show an underestimation over land and an overestimation over the ocean for most models.

For PM, observations from the EARLINET and AERONET networks reveal that the order of magnitude in extinction is similar between the models and the lidar profiles at most locations, with relatively little spread among the models.

Table of Contents

1 Introduction

The Copernicus Atmosphere Monitoring Service (CAMS, https://atmosphere.copernicus.eu) is a component of the European Earth Observation programme Copernicus. The CAMS near-real time services consist of daily analysis and forecasts with the IFS system with data assimilation of trace gas concentrations and aerosol properties. The global modelling system is also used to provide the boundary conditions for an ensemble of more detailed regional air quality models that are used to zoom in on the European domain and produce 4-day forecasts of air quality. The regional forecasting service provides daily 4-days forecasts of the main air quality species and analyses of the day before, based on the results from 9 state-of-the-art atmospheric chemistry. The ensemble represents the median of the 9 model forecasts.

Routine validation of the regional models against surface observations from the European member states (EEA Air Quality e-reporting) is provided for each model individually as well as the ENSEMBLE in separate quarterly validation reports. Validation reports of the CAMS regional products are available in the following portal:
https://atmosphere.copernicus.eu/regional-air-quality-production-systems.
This web page provides access to the quarterly reports on the daily analyses and forecast activities and verification of the regional ENSEMBLE. An overview of the regional air quality forecasting system is provided by Marécal et al (2015).

Validation reports (e.g. Wagner et al., 2020) for the CAMS global products are available at https://atmosphere.copernicus.eu/node/325, including the evaluation on Earth's troposphere, stratosphere, aerosols and greenhouse gases, with state-of-the art observational datasets (GAW, IAGOS, MOPPIT, EMEP, GOME-2, OMPS-LP, BASCOE, AERONET etc.). A published overview on the validation of reactive gases and aerosols in the global analysis and forecast system can be found in Eskes et al (2015). A validation study of the global surface ozone reanalysis for Europe is provided by Katragkou et al (2015).

Details of the various observational datasets can be found in Eskes et al. (2019), "Observations characterisation and validation methods document", also available at:
https://atmosphere.copernicus.eu/sites/default/files/2020-02/CAMS84_2018SC1_D6.1.1-2019_observations_v4_0.pdf.

This document presents an evaluation of the concentrations above the surface as modelled by the set of 9 individual regional models, the ENSEMBLE forecasts derived from those 9 individual forecasts and analyses, and the consistency between the global and regional modelling systems of CAMS.

2 Regional and global CAMS forecasting systems

2.1 Regional models

The European Air Quality products are provided from the Copernicus Atmosphere Monitoring Service (http://atmosphere.copernicus.eu/). These data are available in NetCDF or Grib-Edition2 format. The files are available each day through ftp protocol from the Météo-France server (ftp.cnrm-game-meteo.fr). Since the beginning of February 2020, the data are also available through the Atmosphere Data store (https://ads.atmosphere.copernicus.eu), although the reliability of data retrievals from was relatively low during this quarter. The products are available in Near Real Time (NRT) for four forecast days, following the protocol below:

- Each day 96h model forecasts and 24h analyses for the previous day are provided with hourly resolution. Consistent provision of the analysis product started on the 5th of July 2017.
- Products are available at eight vertical height levels: surface, 50, 250, 500, 1000, 2000, 3000, 5000 meters.
- The pollutants are O_3, CO, NO_2, SO_2, PM2.5, PM10, NO, NH_3, NMVOC, PANs, dust aerosols (fraction below 10μm), secondary inorganic aerosols (fraction below 2.5μm), birch, grass, olive, and ragweed pollen. Since the upgrade of February 4th 2020 three new aerosol species (PM from wildfires, EC from fossil fuels and EC from wood burning) are also available.
- The regional datasets cover the longitudes 335.05°E to 44.95°E every 0.1°, and latitudes 71.95°N – 30.05°N also at 0.1° resolution (~10km).
- The forecasts until the 48th hour are available before 7:30 UTC
- The forecasts for the 49-96th hour are available before 9:30 UTC
- The analyses are provided before 12:00 UTC
- Since the June 2019 upgrade, the regional models make use of the CAMS-REG-AP_v2_2_1 emissions (reference year: 2015) and since February 4th (2020), the updated CAMS-REG-AP_v3_1 emissions dataset (reference year: 2016).

The NRT forecast and analysis regional air quality data are available for the nine air quality models and their ensemble median (CAMS-regional or ENSEMBLE):

- MOCAGE model (MFM)
- LOTOS-EUROS model (KNM)
- EMEP MSC-W model (EMP)
- MATCH model (SMH)
- EURAD-IM model (RIU)
- CHIMERE model (CHI)
- SILAM model (FMI)
- DEHM model (DEM)
- GEM-AQ model (GEM)

Before October 16th, 2019 the regional ensemble was calculated based on 7 regional models (except for DEHM and GEM-AQ). This is the first quarterly report which explicitly includes results and analysis from those two new models.

Every evening, a full download of the 96h forecasts and 24h analyses fields to KNMI at full resolution is performed. These fields are co-located to the set of surface stations used, and this largely reduced datasets is shared with all validation partners.

Documentation about the regional models may be found at the address https://atmosphere.copernicus.eu/regional-air-quality-production-systems. For the purposes of this report however, it's useful to indicate what kinds of observations are actively assimilated in each model (Table 2.1).

Table 2.1: Surface and other observations that are actively assimilated in regional models.
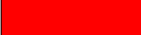
Model	Method	Surface	Other
CHIMERE	Kriging	O_3, NO_2, PM10, PM2.5	
EMEP	Intermittent 3D-var	O_3, SO_2, NO_2, PM10, PM2.5	OMI NO_2
EURAD	Intermittent 3D-var	O_3, SO_2, NO_2, PM10, PM2.5	OMI and MetOp/GOME-2 NO_2 and SO_2, MOPITT and IASI CO
LOTOS-EUROS	Ensemble Kalman filter	O_3, NO_2, PM10, PM2.5	OMI NO_2
MATCH	Intermittent 3D-var	O_3, NO_2, CO, PM10, PM2.5	
MOCAGE	3D-var	O_3, NO_2, PM10	
SILAM	Intermittent 3D-var	O_3, NO_2, CO, PM10, PM2.5	
DEHM	Optimal interpolation	O_3, NO_2	
GEM-AQ	Optimal interpolation	O_3, NO_2, CO, SO_2, PM10, PM2.5	

Validation reports of the CAMS regional products are available in the following portal: https://atmosphere.copernicus.eu/regional-services.

Whenever possible, in this report, models follow the naming and colour scheme of Table 2.2.

Table 2.2: Naming and colour scheme followed throughout this report. Since October 2019 the ensemble is computed using also the two new models, DEHM and GEM-AQ.

Model	Short model name	Colour name	Colour
CAMS-global	CAMSg	Blue	
ENSEMBLE forecast	ENS-fc	Red	
ENSEMBLE analysis	ENS-an	Green	
CHIMERE	CHIM	Yellow	
EMEP	EMEP	Brown	
EURAD	EURAD	Cyan	
LOTOS-EUROS	LOTOS	Purple	
MATCH	MATCH	Grey	
MOCAGE	MOCA	Pink	
SILAM	SILAM	Orange	
DEHM	DEHM	Fuchsia	
GEM-AQ	GEMAQ	Light Green	

2.2   Global CAMS system based on the ECMWF IFS model

The CAMS-global operational assimilation/forecast system consists of the IFS-CB05 chemistry combined with the MACC aerosol model. The chemistry is described in Flemming et al. (2015); aerosol is described by the bulk aerosol scheme (Morcrette et al., 2009). Dissemination of CAMS-global forecasts is twice a day, at about 10:00 and 22:00UTC. The forecast length is 120 h. Users can get access at https://atmosphere.copernicus.eu/catalogue. Table 2.3 provides information on the satellite data used in CAMS-global.

Table 2.3: Satellite retrievals of reactive gases and aerosol optical depth that are actively assimilated in CAMS-global.

Instrument	Satellite	Provider	Version	Type	Status
MLS	AURA	NASA	V4	O_3 Profiles	20130107 -
OMI	AURA	NASA	V883	O_3 Total column	20090901 -
GOME-2	Metop-A	Eumetsat	GDP 4.8	O_3 Total column	20131007 - 20181231
GOME-2	Metop-B	Eumetsat	GDP 4.8	O_3 Total column	20140512 -
GOME-2	Metop-C	Eumetsat	GDP 4.9	O_3 Total column	20200505 -
SBUV-2	NOAA-19	NOAA	V8	O_3 21 layer profiles	20121007 -
OMPS	Suomi-NPP	NOAA / EUMETSAT		O_3 Profiles	20170124 - 20190409
IASI	MetOp-A	LATMOS/ULB EUMETSAT	-	CO Total column	20090901 - 20180621 20180622 - 20191118
IASI	MetOp-B	LATMOS/ULB EUMETSAT	-	CO Total column	20140918 - 20180621 20180622 -
IASI	MetOp-C	EUMETSAT		CO Total column	20191119-
MOPITT	TERRA	NCAR	V5-TIR V7-TIR V7-TIR Lance V8-TIR	CO Total column	20130129 - 20160124 - 20180626 20180626 20190702
OMI	AURA	KNMI	DOMINO V2.0	NO_2 Tropospheric column	20120705 -
GOME-2	METOP A	Eumetsat	GDP 4.8	SO_2 Total column	20150902-20200504
GOME-2	METOP B	Eumetsat	GDP 4.8	SO_2 Total column	20150902-
GOME-2	METOP C	Eumetsat	GDP 4.9	SO_2 Total column	20200505-
MODIS	AQUA / TERRA	NASA	Col. 6 Deep Blue	Aerosol total optical depth, fire radiative power	20090901 - 20150902 -
PMAp	METOP-A METOP-B	EUMETSAT		AOD	20170124 - 20170926 -

GOME-2	METOP A	Eumetsat	GDP 4.8	NO$_2$ Tropospheric column	20180624-20200504
GOME-2	METOP B	Eumetsat	GDP 4.8	NO$_2$ Tropospheric column	20180626-
GOME-2	METOP C	Eumetsat	GDP 4.9	NO$_2$ Tropospheric column	20200505-

The most recent upgrade to cycle 46R1, relevant to this report, took place on 9 July 2019 and involves a change from 60 to 137 vertical levels. A detailed changelog and the corresponding validation reports for this last upgrade can be found on the following page:

https://atmosphere.copernicus.eu/node/472.

Upgrade and version information is available here:
https://atmosphere.copernicus.eu/changes-cams-global-production-system.

Documentation on the global system can be found here:
https://atmosphere.copernicus.eu/global-production-system.

3 Consistency between the global and regional modelling components of CAMS

3.1 Summary

This chapter reports on the consistency between the global and regional modelling components of CAMS, and the impact of global CAMS boundary conditions on regional forecasts. The current evaluation includes ozone (O_3) carbon monoxide (CO) and aerosol (PM10/PM2.5) forecasts covering the period from June to August 2020. Two new regional models have been taken into account in the analysis, namely the DEHM and the GEM-AQ models. The two models have a normal model performance, no particular outliers where noted.

Global and regional ENSEMBLE forecasts: The two forecast products compare well. Some inconsistency is seen in the PBL O3, which is lower over the Scandinavian Peninsula in the regional ENSEMBLE forecast compared to the global forecast, a feature is noted for the first time. Surface CO in the global forecast is also higher over European and North African hot spots (emission sources) which are not as pronounced in the regional forecast. PM10 appears higher over North Africa/Mediterranean in the global forecast than in the regional ENSEMBLE up to 2000 m.

Regional forecast variability: The regional ensemble members exhibit the expected regional variability for O_3, CO and particulate matter. Minor deviations (compared to CAMS-global and the majority of regional models) are generally noted. The implementation of boundaries in CHIMERE and EURAD appears to be different compared to the other models, something that is particularly evident in the O3/CO lateral boundary plots.

Diurnal cycles: Comparison of the diurnal cycles between the global and the regional forecasts over different European subregions indicate a good agreement for surface O3 and more pronounced diurnal variability for the surface CO global forecast than the regional ensemble members. The agreement in PM10 is better over northern Europe.

Regional forecast and analysis: The comparison of regional analyses and forecasts shows strong model, altitude and species dependence.

3.2 Methodology for the comparisons of CAMS-global and CAMS-regional

Operational download

The daily regional CAMS forecasts are retrieved on a daily basis by the AUTH-server. This includes the 3-hourly (0,3,6,9 etc.) regional forecast data (ensemble members and regional ENSEMBLE) for all provided species at all vertical layers for the 5 forecast days extracted from the Météo-France ftp server, and the 3-hourly (0, 3, 6, 9 etc.) global forecast data for 5 forecast days extracted from the ECMWF CAMS ftp server

Methodology of global-regional comparison

The following methodology is used to a) convert CAMS-global species from mass mixing ratio (kg/kg) to concentration ($\mu g/m^3$) and, b) extract CAMS-global species concentrations from the vertical levels that lie closest to the regional height levels.

The following parameters are used from the CAMS-global model: hybrid layer coefficients; temperature, surface pressure; "GEMS" ozone; carbon monoxide; "aermr01-11" and "aermr16-18" (aerosol species, kg/kg). Data from the first 78 vertical layers (from the surface) are used.

The thickness of each vertical layer Δz (m) is calculated:

$$\Delta z_k = \frac{R * T_k}{M_{air} * g} * \ln\left(\frac{p_{i_{k+1}}}{p_{i_k}}\right) \qquad (E.1)$$

where R=8.314 J/mol·K the gas constant, T the temperature at vertical layer midpoint, M_{air}=28.97·10^{-3} kg/mol the molecular weight of air and g=9.8 m/s^2 the gravity acceleration.

The mass-mixing ratio (kg/kg) for ozone (go3) and carbon monoxide (CO) is initially provided. Conversion from mass mixing ratio (kg/kg) to concentration ($\mu g/m^3$) is performed using the following approach:

$$\rho_{O3} = mmr_{O3} * \left(\frac{p_m * M_{air}}{R * T}\right) \qquad (E.2)$$

where ρ_{O3} is the ozone concentration (kg/m^3) and mmr_{O3} the ozone mass mixing ratio (kg/kg). The expression inside the parentheses in E.2 corresponds to the air concentration (kg/m^3). The same approach is also used for the CO unit conversion from kg/kg to $\mu g/m^3$.

Figure 3.1. Left: Mean global and regional ENSEMBLE forecast ozone fields for four different vertical layers (0, 250, 2000, 5000 m) for JJA2020. Right: Cross sections for the same period of the global and regional ENSEMBLE ozone boundaries (south, west, north, east).

The mass mixing ratio for all aerosol species (see Table 3.1) is initially provided. The PM10 and PM2P5 species are converted to µg/m³ as follows:

$$\rho_{PM10} = \left[\frac{aermr01}{4.3} + \frac{aermr02}{4.3} + aermr04 + aermr05 + 0.4 * aermr06 + aermr07 \right.$$
$$+ aemr08 + aermr11 + aerm09 + aermr10 + aermr16 + aemr17$$
$$\left. + aermr18\right] * \left(\frac{p_m}{R_{spec} * T}\right) \qquad (E.3)$$

$$\rho_{PM2P5} = \left[\frac{aermr01}{4.3} + \frac{aermr02 * 0.5}{4.3} + aermr04 + aermr05 + 0.7 * aermr07 + 0.7\right.$$
$$* aermr08 + 0.7 * aermr11 + aermr09 + aermr10 + 0.7 * aermr16 + 0.25$$
$$\left. * aermr17 + 0.7 * aermr18\right] * \left(\frac{p_m}{R_{spec} * T}\right) \qquad (E.4)$$

where R_{spec}=287.058 J/(kg·K) is the specific gas constant for dry air. The expression inside the parentheses in E.3 and E.4 corresponds to the dry air concentration (kg/m³).
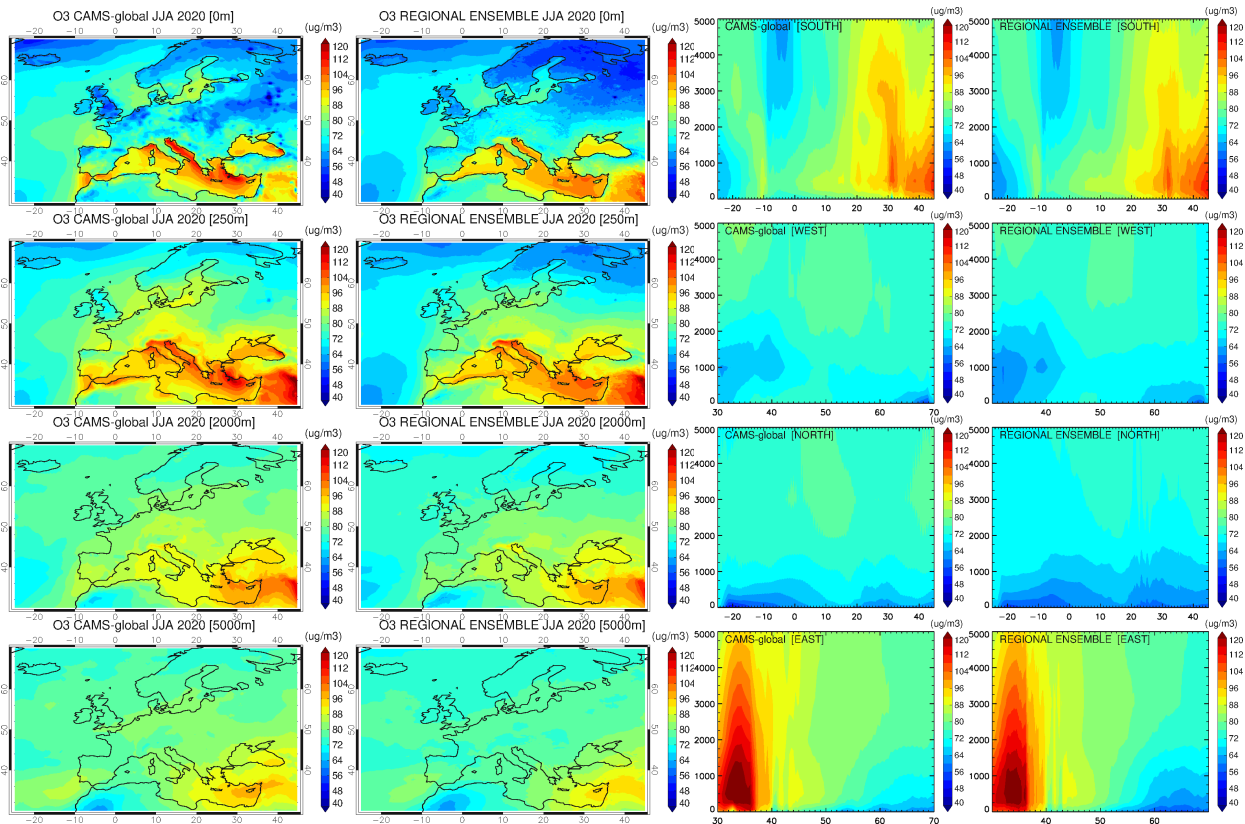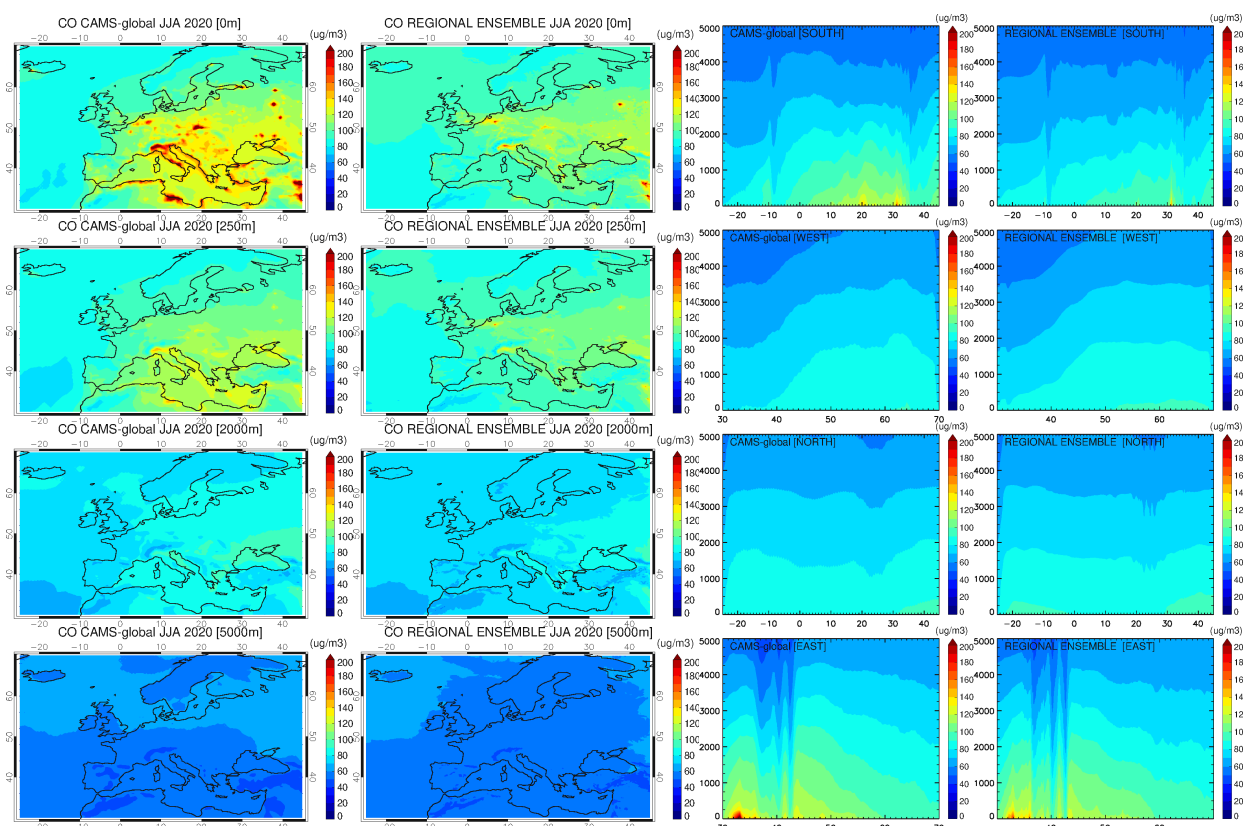
Figure 3.2. Left: Mean global and regional ENSEMBLE forecast CO fields for four different vertical layers (0, 250, 2000, 5000 m) for JJA2020. Right: Cross sections for the same period of the global and regional ENSEMBLE ozone boundaries (south, west, north, east).

Regional model products are provided at the height levels of 0, 50, 250, 500, 1000, 2000, 3000 and 5000m. For every grid point and time step of the CAMS-global model, the differences between the height of each vertical layer midpoint z_m and the regional model height (e.g. 5000m) is calculated. The layer midpoint that exhibits the minimum height difference is the one that lies closest to the regional height level and is therefore selected for extraction of both chemical and aerosol species concentrations. The above procedure is performed for every regional height level. The final global product contains the O_3, CO, PM10 and PM2.5 concentrations in eight height levels that correspond to the CAMS-global vertical levels that lie closest to the regional height levels.
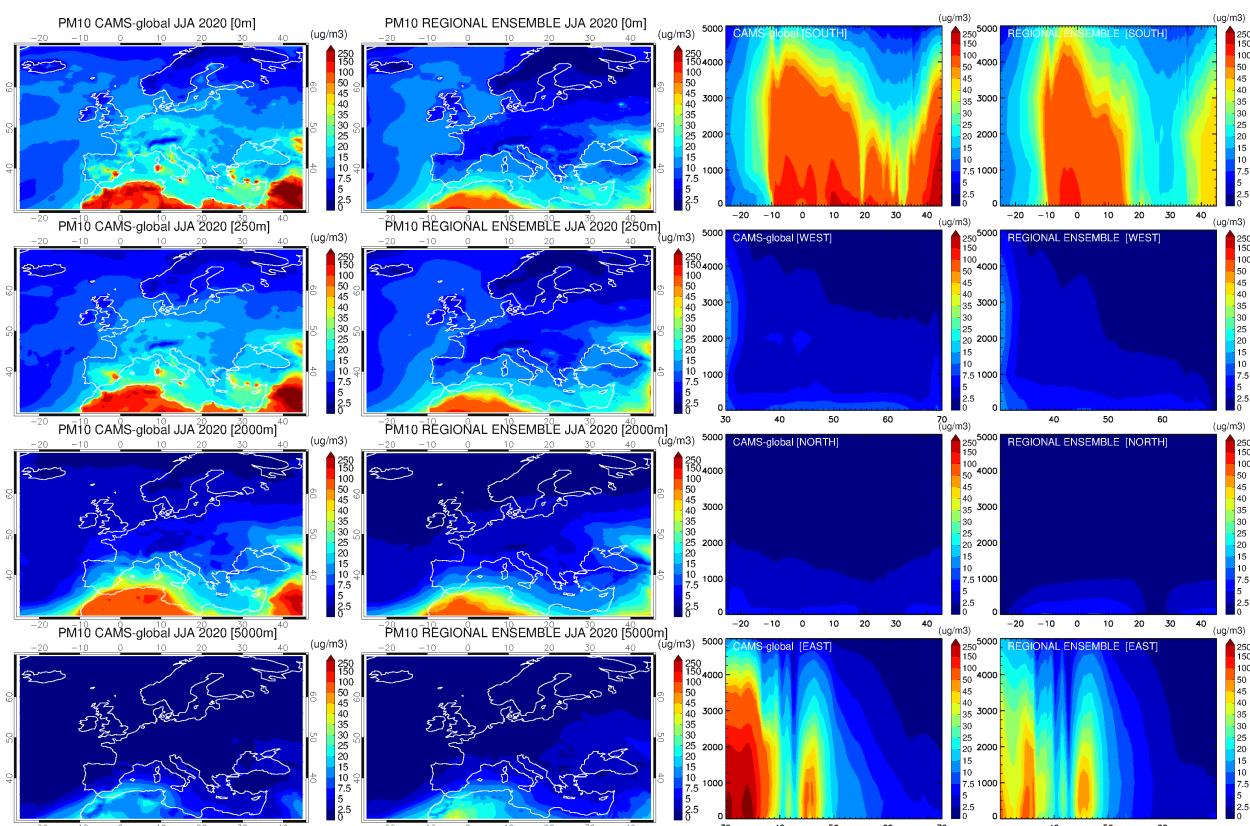
Figure 3.3. Left: Mean global and regional ENSEMBLE forecast PM10 fields for four different vertical layers (0, 250, 2000, 5000 m) for JJA2020. Right: Cross sections for the same period of the global and regional ENSEMBLE ozone boundaries (south, west, north, east).

Table 3.1: Aerosol species description.

| Label		Name	Size (μm)
aermr01	SS1	Sea Salt Aerosol	0.03-0.5
aermr02	SS2	Sea Salt Aerosol	0.5-5
aermr03	SS3	Sea Salt Aerosol	5-20
aermr04	DD1	Dust Aerosol	0.03-0.55
aermr05	DD2	Dust Aerosol	0.55-0.9
aermr06	DD3	Dust Aerosol	0.9-20
aermr07	OM1	Hydrophobic Organic Matter Aerosol	
aermr08	OM2	Hydrophilic Organic Matter Aerosol	
aermr09	BC1	Hydrophobic Black Carbon Aerosol	
aermr10	BC2	Hydrophilic Black Carbon Aerosol	
aermr11	SU1	Sulphate Aerosol	
aermr16	NI1	Nitrate fine mode aerosol	
aermr17	NI2	Nitrate coarse mode aerosol	
aermr18	NH3	Ammonium aerosol	
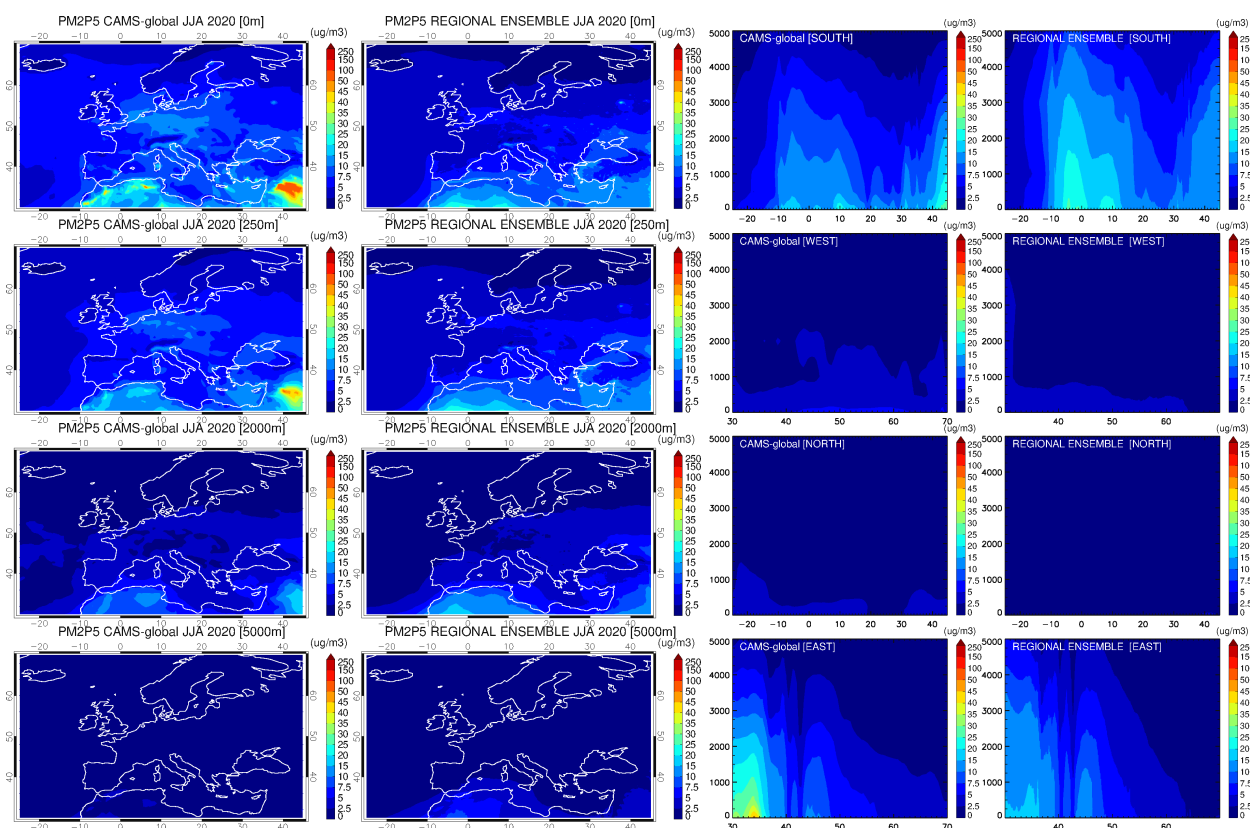
Figure 3.4 Left: Mean global and regional ensemble forecast PM2.5 fields for four different vertical layers (0, 250, 2000, 5000 m) for JJA2020. Right: Cross sections for the same period of the global and regional ensemble ozone boundaries (south, west, north, east).

3.3 Consistency between the global and regional forecasts

This section reports on the consistency of the global and regional ensemble forecast. The analysis is performed for O_3, CO, PM10 and PM2.5 at four levels (0, 250, 2000 and 5000 m) for the time period from June to August 2020 (JJA).

Ozone (O_3)

Figure 3.1 shows the average JJA2020 spatial distribution of O_3 for different vertical layers (left) and the cross sections of the lateral boundaries (right) for the CAMS-global forecast and the regional ENSEMBLE. Overall, CAMS-global and regional ENSEMBLE agree well for O3, with CAMS-global exhibiting somehow lower ozone near the surface and in the boundary layer over continental Europe, while the opposite stands for Scandinavia. The agreement in the four boundaries is very good.

Carbon monoxide (CO)

Figure 3.2 illustrates the seasonal mean fields of CO and the cross sections of lateral boundaries (right) for ENSEMBLE and CAMS-global. There are more local surface CO maxima in the global forecast product, over continental Europe and the coast of North Africa. Probably, this is due to the different emissions in the two products. The comparison in the boundaries is very good.

Aerosols (PM10 and PM2.5)

Figures 3.3 and 3.4 illustrate the CAMS-global and ENSEMBLE spatial distributions (left) and lateral boundary cross sections (right) of PM10 and PM2.5 mean fields, respectively. The main inconsistencies for particulate matter are summarized below:

- Higher PM10 in CAMS-global compared to The regional ENSEMBLE over the southern and south-eastern boundaries of the domain up to 2000 m altitude.
- Some hotspots of PM10 concentrations in CAMS-global over the Mediterranean not seen in The regional ENSEMBLE.
- High PM10/PM2.5 surface concentrations over Middle East seen in CAMS-global are not found in The regional ENSEMBLE.

3.4   Regional variability

Ozone (O$_3$)

Figure 3.5 illustrates ozone mean fields for JJA2020 of the individual regional ensemble members and CAMS-global (bottom panel) for selected altitudes (0, 250, 2000, 5000 m). The two new regional ensemble members (DEHM and GEM-AQ) have relatively lower surface O3 over Scandinavia and GEM-AQ has higher surface O3 over the Mediterranean, in comparison to CAMS-global. The EURAD model has considerably lower O3 over the Atlantic (western boundary). MOCAGE exhibits considerably higher concentrations over central Europe at 250m compared to CAMS-global.

Carbon monoxide (CO)

The JJA2020 mean fields of carbon monoxide for the regional ensemble members as well as for CAMS-global are illustrated in Figure 3.6 for different vertical layers. No particular outliers are found, as CO in the regional models is within the expect variability. The new regional model GEM-AQ has higher CO than all other regional models within the PBL in better agreement with CAMS-global.

Aerosols (PM10 and PM2.5)
The mean PM10 and PM2.5 fields for JJA2020 are illustrated in Figures 3.7 and 3.8, respectively. The inconsistences noted are:

- MOCAGE appears as an outlier in the boundary layer for PM10.
- EURAD, MATCH and DEHM exhibit relatively lower PM10 levels at the south boundary.
- EMEP, SILAM and GEMAQ have higher PM2.5 levels in the PBL near the south boundaries compared to CAMS-global, while all other models have lower PM2.5. This may be related to the fact that these models use their own desert dust sources in the areas near the southern boundary of the domain.

Figure 3.5 Mean regional ozone forecasts for JJA2020 for four different vertical layers (0, 250, 2000, 5000 m) from the nine ensemble members and CAMS-global (top to bottom: CHIMERE, EMEP, SILAM, LOTOS-EUROS, MOCAGE, EURAD-IM, MATCH, DEHM, GEM-AQ and CAMS-global).
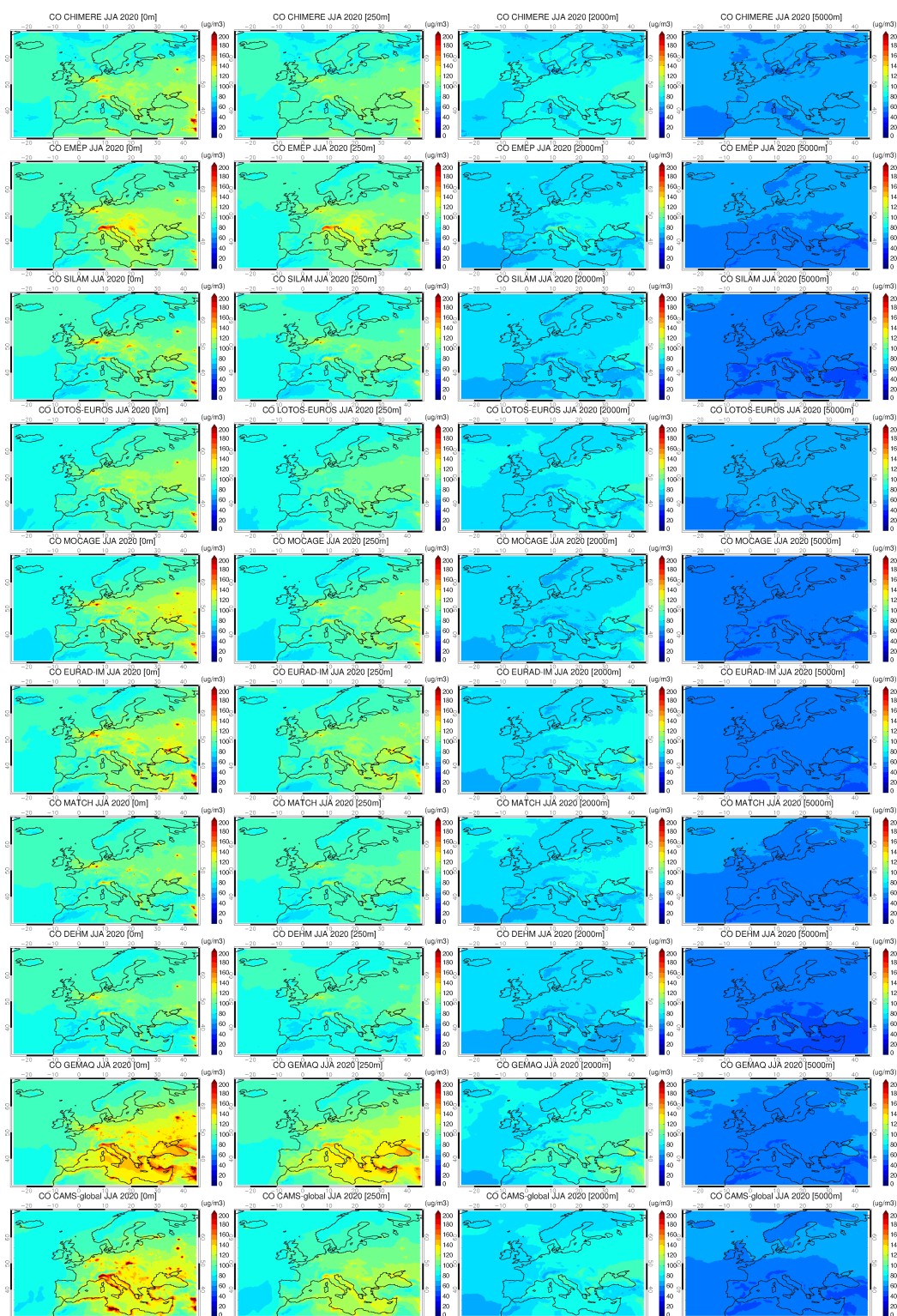
Figure 3.6. Mean regional CO forecasts for JJA2020 for four different vertical layers (0, 250, 2000, 5000 m) from the nine ensemble members and CAMS-global (top to bottom: CHIMERE, EMEP, SILAM, LOTOS-EUROS, MOCAGE, EURAD-IM, MATCH, DEHM, GEM-AQ and CAMS-global).
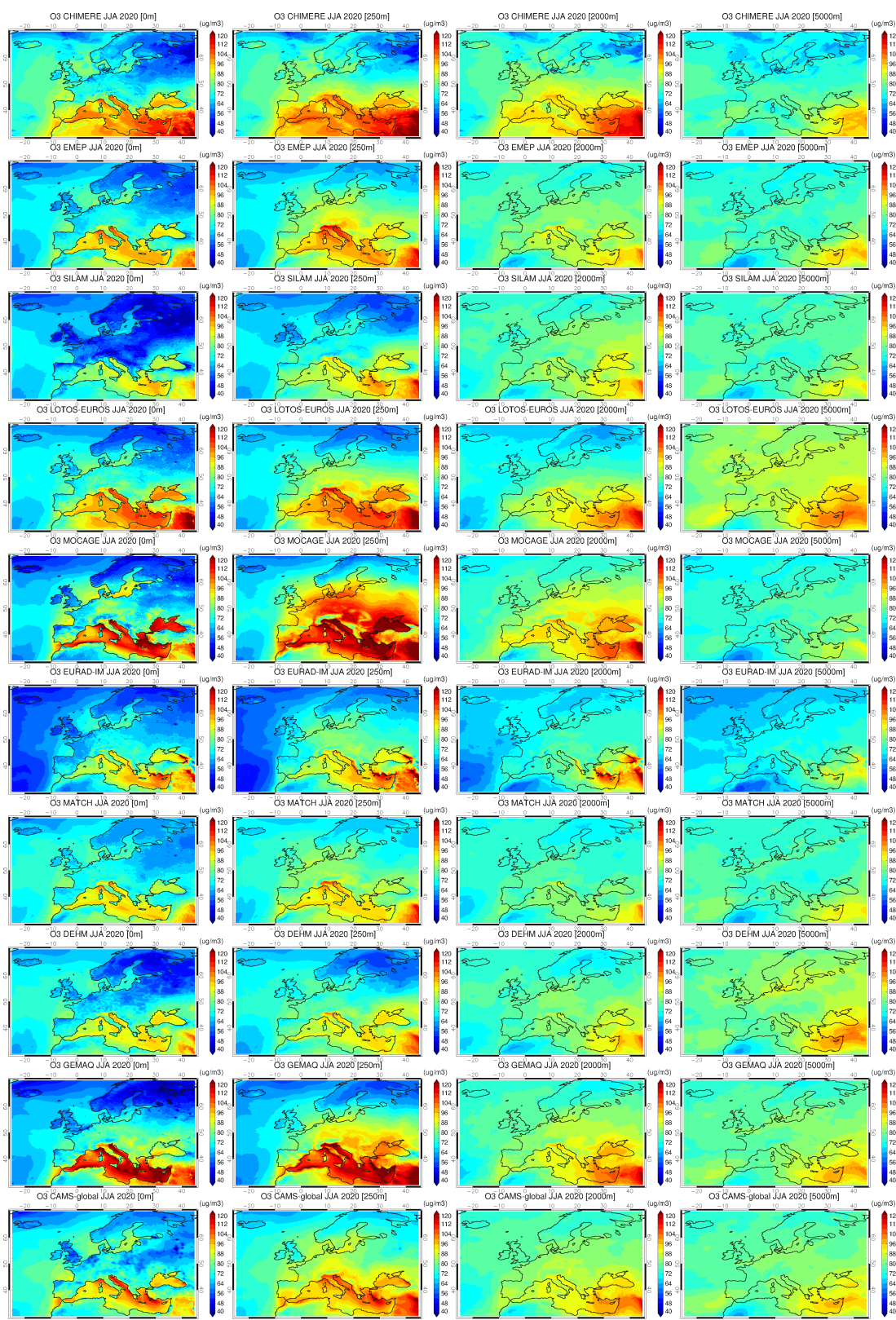
Figure 3.7. Mean regional PM10 forecasts for JJA2020 for four different vertical layers (0, 250, 2000, 5000 m) from the nine ensemble members and CAMS-global (top to bottom: CHIMERE, EMEP, SILAM, LOTOS-EUROS, MOCAGE, EURAD-IM, MATCH, DEHM, GEM-AQv and CAMS-global).

Figure 3.8. Mean regional PM2.5 forecasts for JJA2020 for four different vertical layers (0, 250, 2000, 5000 m) from the nine ensemble members and CAMS-global (top to bottom: CHIMERE, EMEP, SILAM, LOTOS-EUROS, MOCAGE, EURAD-IM, MATCH, DEHM, GEM-AQ and CAMS-global).
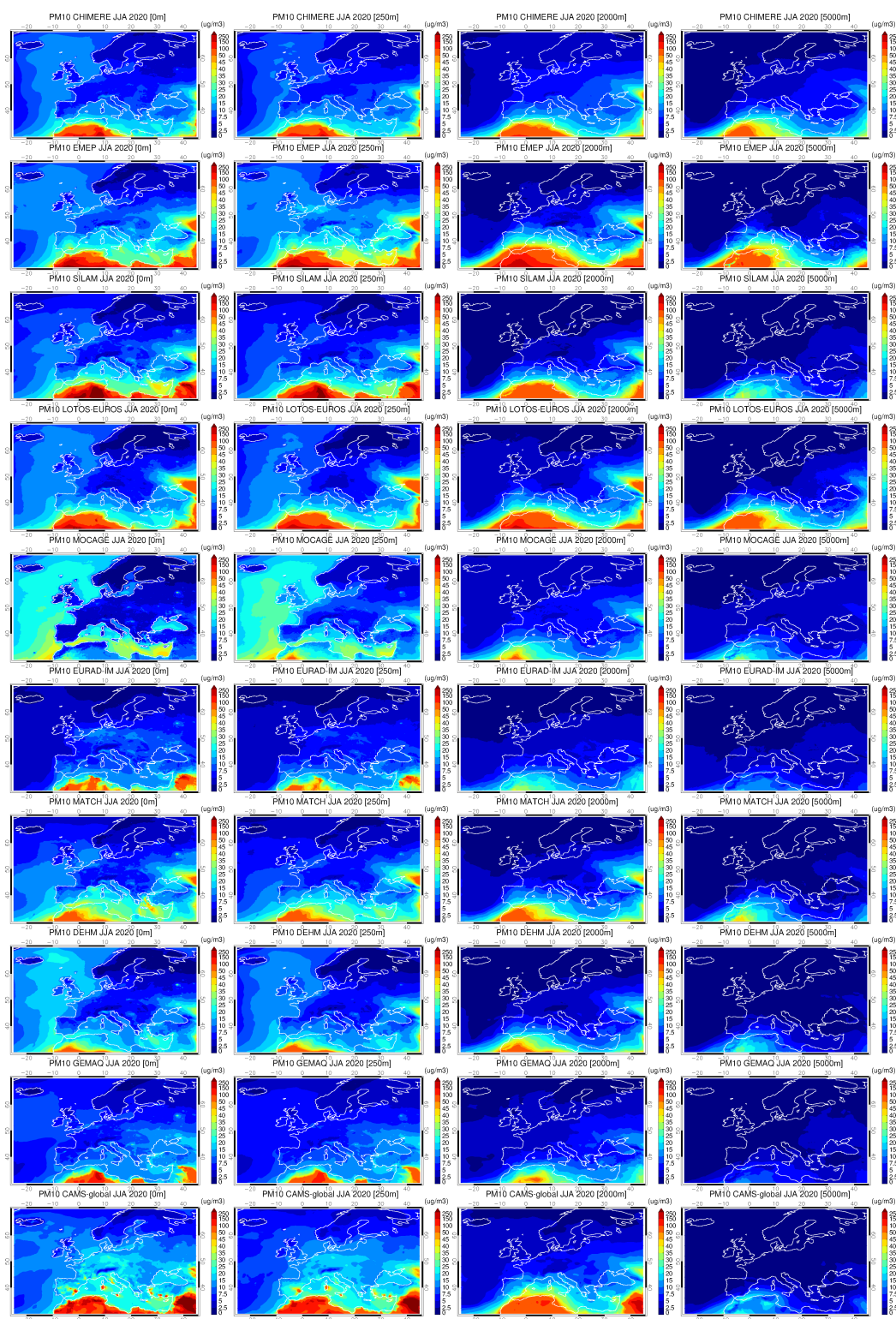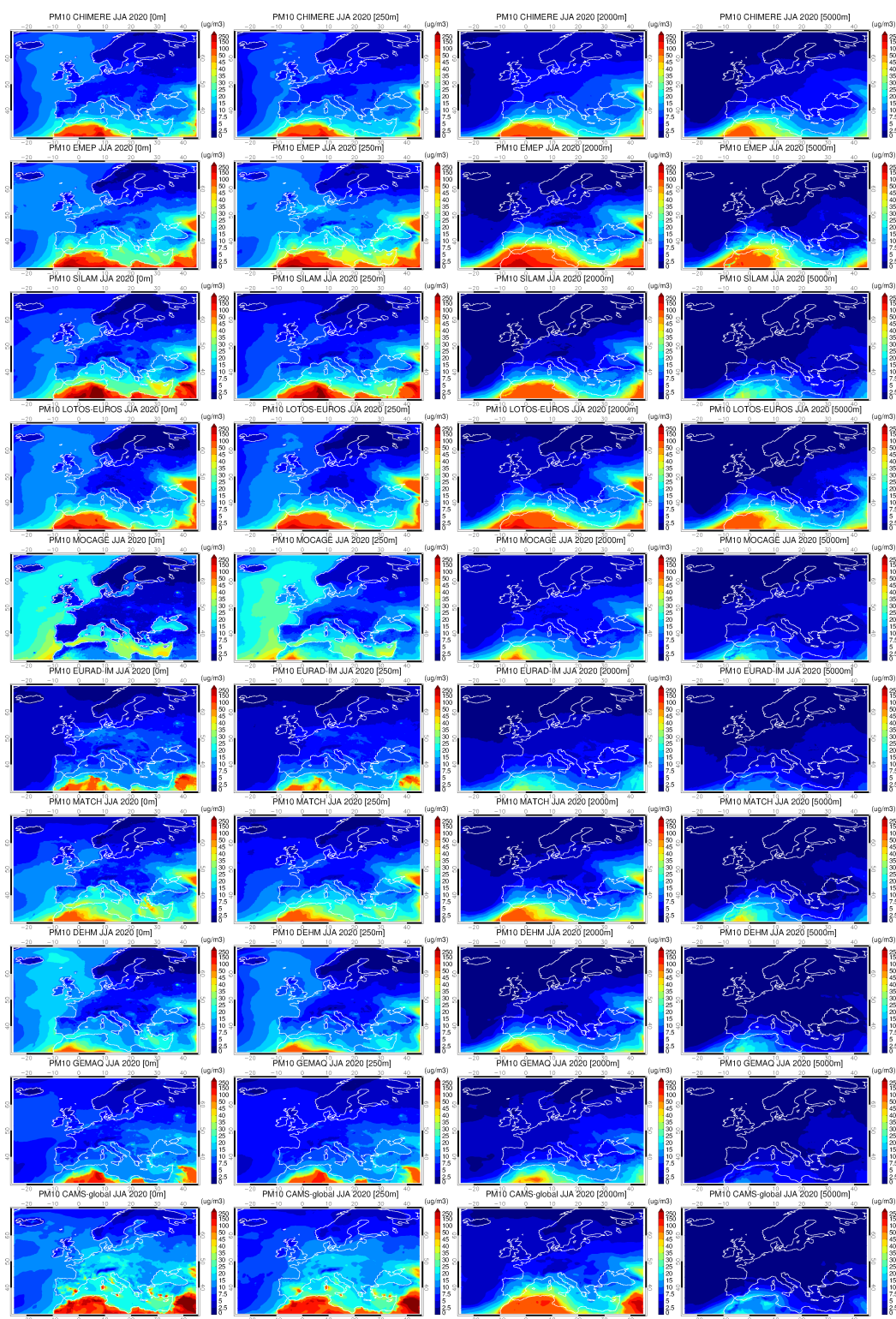
3.5   Time series

Figure 3.9 shows the mean daily time series from June to August 2020 for the four species, namely O_3, CO, PM10, PM2.5 (from left to right) for different European sub-regions (from top to bottom): Alps (AL), British Isles (BI), East Europe (EA), France (FR), Iberian Peninsula (IP), Mediterranean (MD), Mid-Europe (ME), Scandinavia (SC).

Each subregion is defined with the following latitude/longitude boundaries:

Name = (BI, IP, FR, ME, SC, AL, MD, EA)

West = (-10, -10, -5, 2, 5, 5, 3, 16)

East = (2, 3, 5, 16, 30, 15, 25, 30)

South = (50, 36, 44, 48, 55, 44, 36, 44)

North = (59, 44, 50, 55, 70, 48, 44, 55)

The letter R denotes the temporal correlation between the two products. Only concentrations over land are used.

For O_3 the temporal variability between CAMS-global and The regional ENSEMBLE is quite good (0.73 to 0.94) with the exception of the MD area (0.49).

The temporal correlation for CO is particularly good (0.85 to 0.95) with the exception of the IP (0.66).

The agreement in the PM10 correlation in the global and regional forecasts ranges is good in the majority of the regions (>0.7) with the exception of EA (0.67), MD (0.59), ME (0.65) and SC (0.66).

The magnitude of PM10 in CAMS-global is close to the regional ENSEMBLE, however CAMS-global PM10 is higher in all regions but the BI.

The temporal agreement in surface PM2.5 is good (0.71 to 0.93) with the exception of EA (0.53).

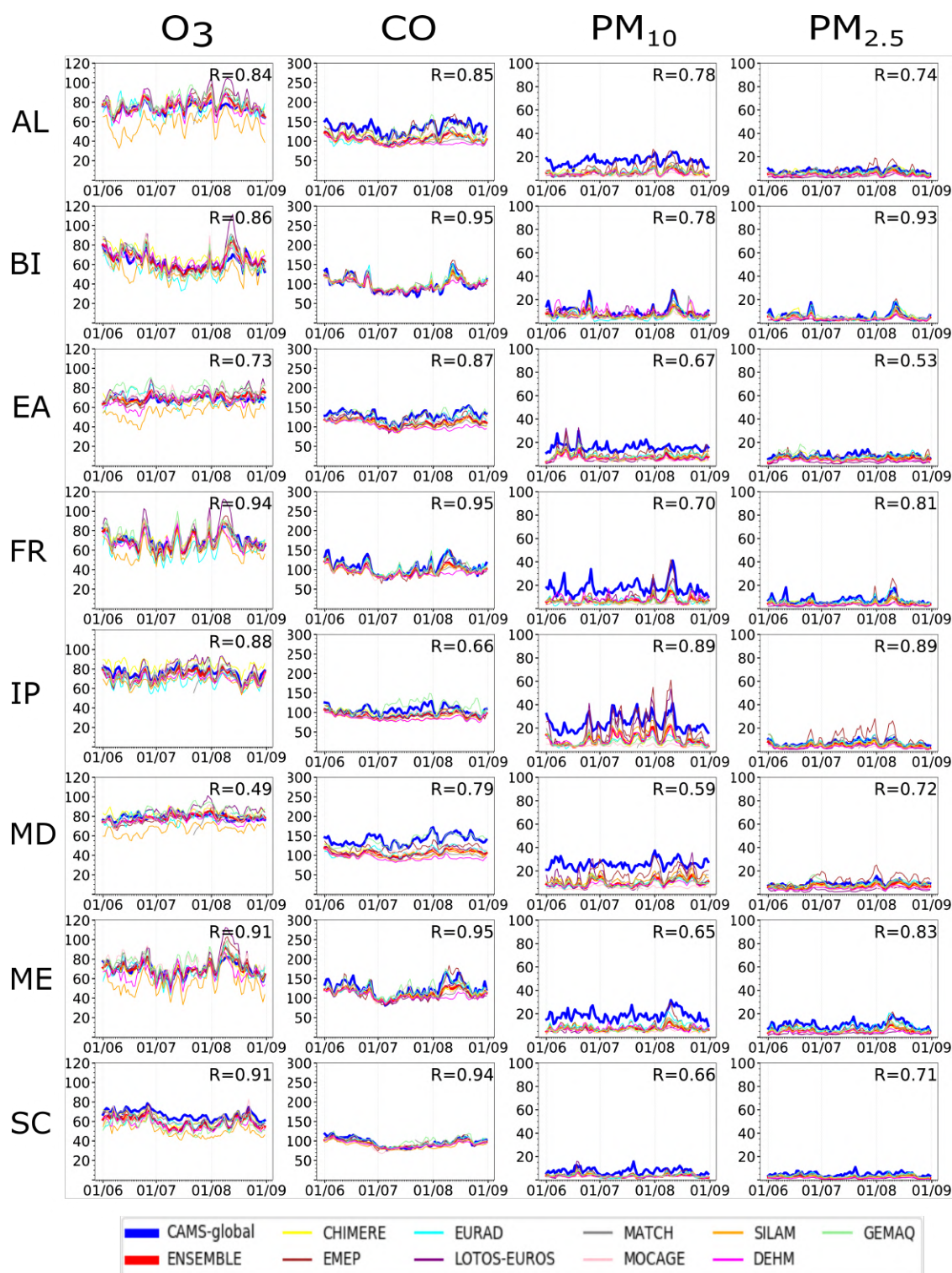Figure 3.9. Mean daily time series for surface O_3, CO, PM10 and PM2.5 for JJA2020 (unit µg/m^3). The blue line is CAMS-global and the red line the ENSEMBLE forecasts. Each line in the composite plot denotes a different European subregion.

3.6   Diurnal cycles

Figure 3.10 shows the diurnal cycles for surface O_3, CO, PM10 and PM2.5 averaged over the period JJA2020 for different European sub-regions (from top to bottom): Alps (AL), British Isles (BI), East Europe (EA), France (FR), Iberian Peninsula (IP), Mediterranean (MD), Mid-Europe (ME), Scandinavia (SC). The red colour is used for the regional ensemble and the blue for CAMS-global.

There is a good agreement between the O_3 CAMS-global and the regional diurnal cycles, both in timing and amplitude of the diurnal cycle.

The diurnal range for surface CO negligible for the regional products over some regions, mostly over the Alps and to a lesser extent over the southern Europe and central Europe.

The diurnal cycle in the PM10 global product is more pronounced, especially over southern Europe, in comparison to the regional products. The agreement is better for PM2.5, which generally exhibit less pronounced diurnal cycles than PM10.

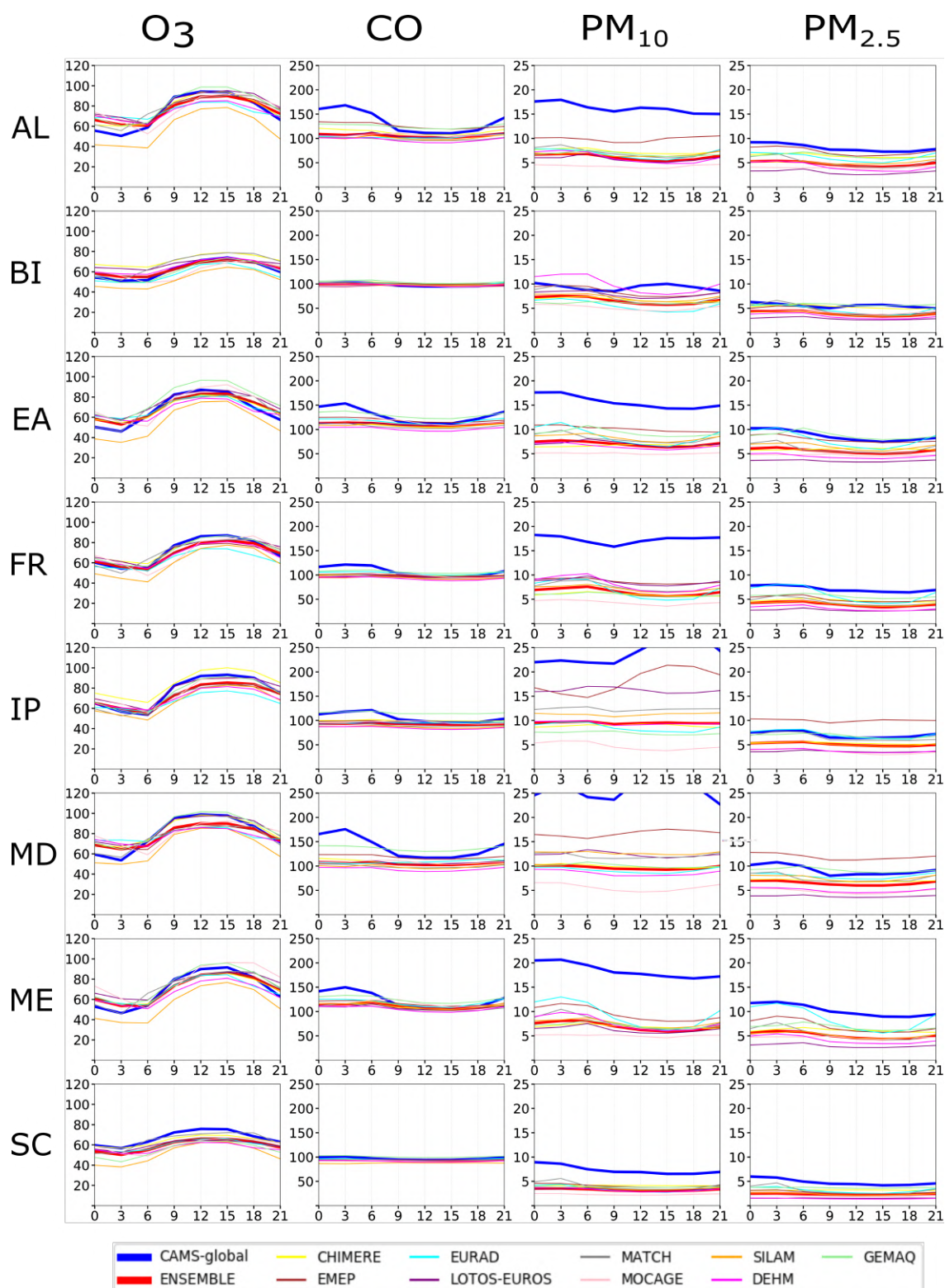Figure 3.10. Diurnal cycles for surface O₃, CO, PM10 and PM2.5 for the period JJA2020 (unit µg/m³). The blue line is CAMS-global and the red line the ENSEMBLE forecast. Each line in the composite plot denotes a different European sub-region.

3.7   Regional domain boundary cross sections

Ozone (O$_3$)

Figure 3.11 shows the regional variability in the lateral domain boundary cross sections of O$_3$ (from left to right): south, west, north, east and the different ensemble members and CAMS-global (from top to bottom) averaged over the period JJA2020. The implementation of O$_3$ boundaries in CHIMERE and EURAD appears to be different than in other models.

Carbon monoxide (CO)

Figure 3.12 shows the regional variability in the lateral cross sections of CO (from left to right): south, west, north, east and the different ensemble members and CAMS-global (from top to bottom) averaged over the period JJA2020. CHIMERE has an outstanding eastern boundary implementation.

Aerosols (PM10 and PM2.5)

Figure 3.13/3.14 shows the regional variability in the lateral cross sections of PM10/PM2.5 respectively (left to right): south, west, north, east and the different ensemble members and CAMS-global (top to bottom). We note the following inconsistencies:

- MOCAGE, EURAD, and to a lesser extent GEM-AQ exhibit lower PM10 concentrations over the south boundaries than CAMS-global.

- MOCAGE has higher PM10 than CAMS-global at the west and north BCs, probably associated to sea salt

- MOCAGE, EURAD and GEM-AQ have lower PM10 at the eastern boundaries.

- CHIMERE, EMEP and GEM-AQ have higher southern PM2.5 boundaries that CAMS-global.

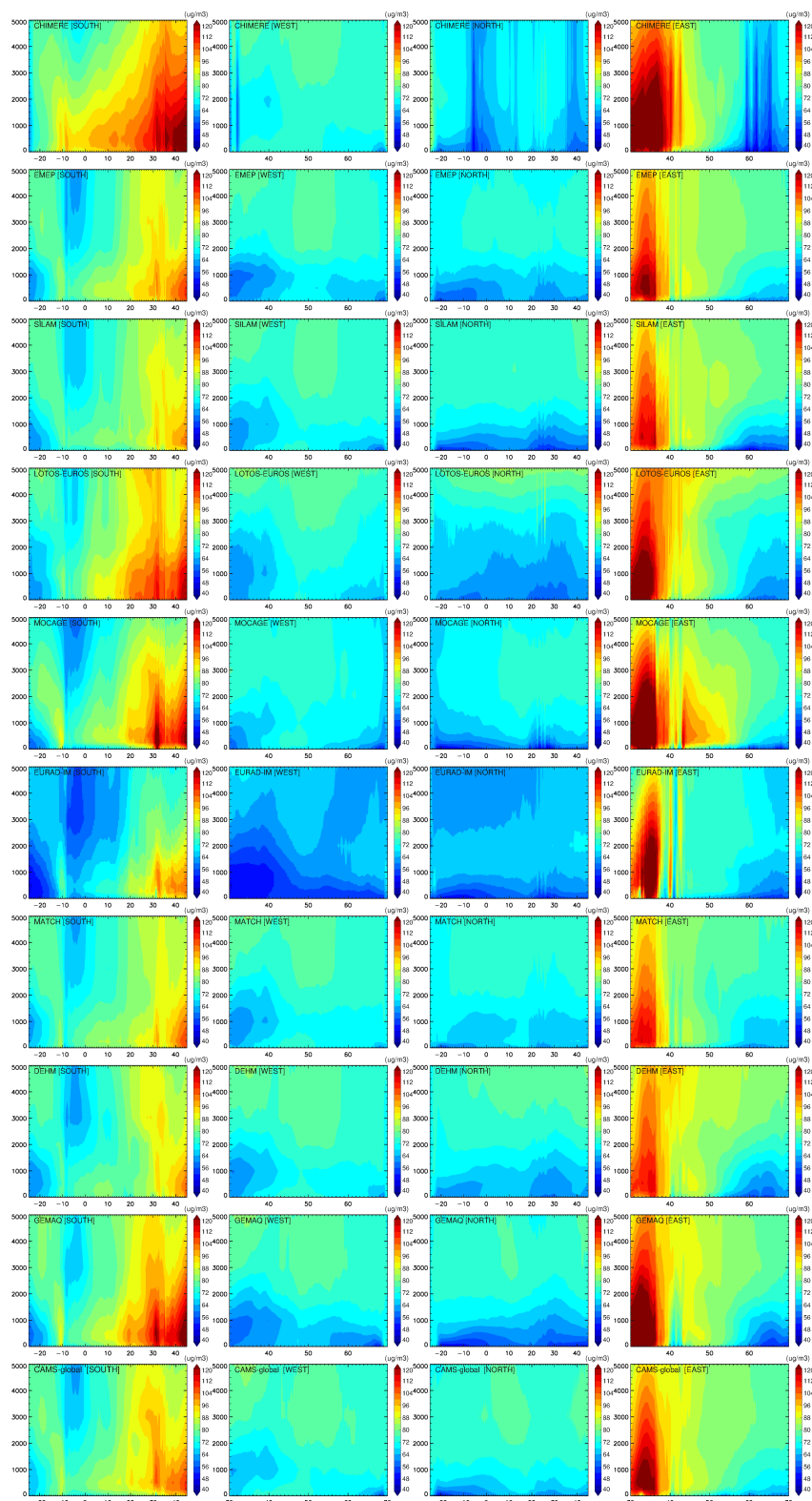- EMEP has additionally higher eastern PM2.5 boundaries than CAMS-global.

Figure 3.11. Ozone cross sections for JJA2020 for the nine ensemble members and CAMS-global (top to bottom: CHIMERE, EMEP, SILAM, LOTOS-EUROS, MOCAGE, EURAD-IM, MATCH, DEHM, GEM-AQ and CAMS-global) and the lateral boundaries (left to right: south, west, north, east).
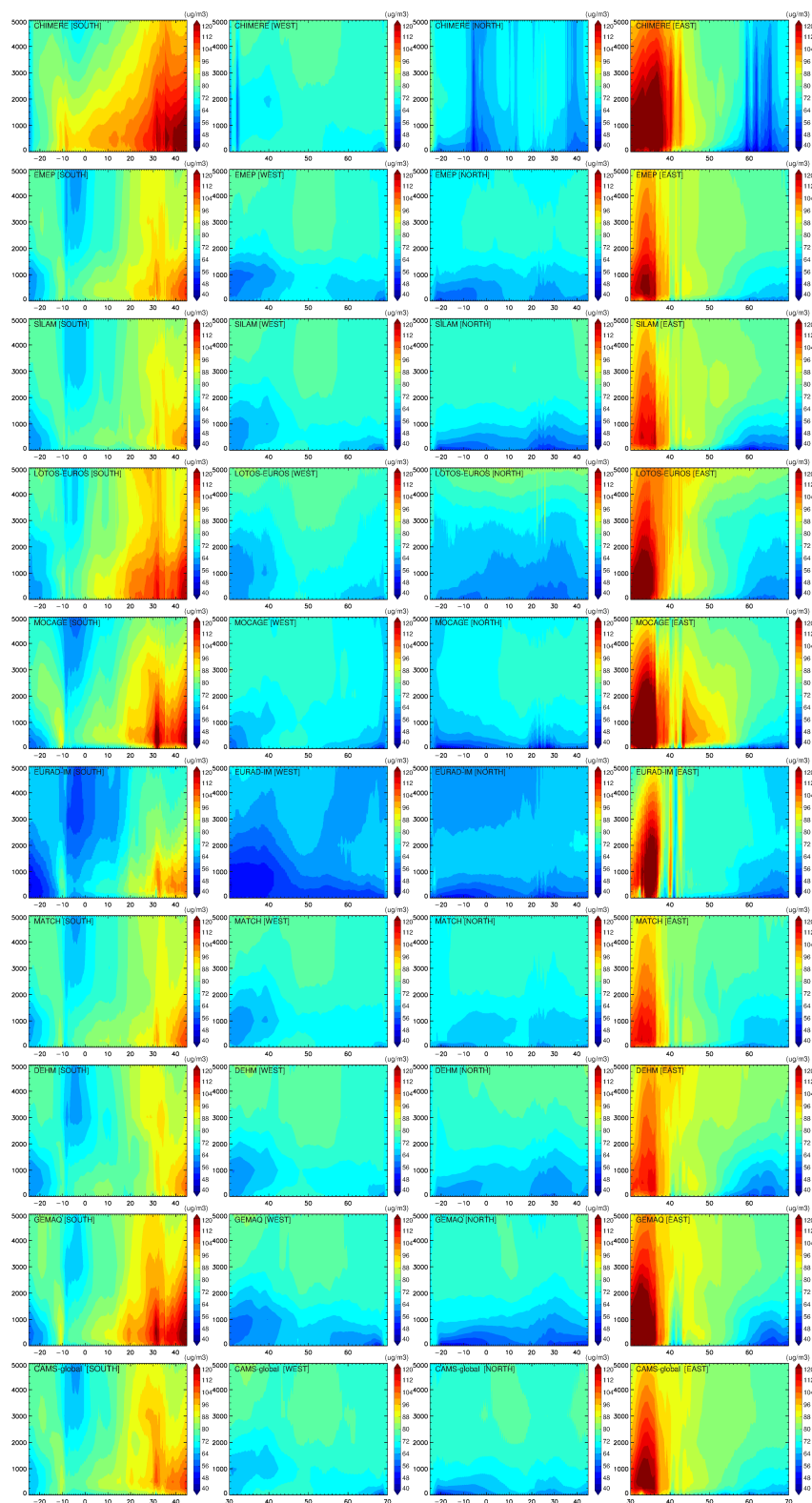
Figure 3.12. Carbon Monoxide cross sections for JJA2020 for the nine ensemble members and CAMS-global (top to bottom: CHIMERE, EMEP, SILAM, LOTOS-EUROS, MOCAGE, EURAD-IM, MATCH, DEHM, GEM-AQ and CAMS-global) and the lateral boundaries (left to right: south, west, north, east).

Figure 3.13. Aerosol PM10 cross sections for JJA2020 for the nine ensemble members and CAMS-global (top to bottom: CHIMERE, EMEP, SILAM, LOTOS-EUROS, MOCAGE, EURAD-IM, MATCH, DEHM, GEM-AQ and CAMS-global) and the lateral boundaries (left to right: south, west, north, east).

Figure 3.14. Aerosol PM2.5 cross sections for JJA2020 for the nine ensemble members and CAMS-global (top to bottom: CHIMERE, EMEP, SILAM, LOTOS-EUROS, MOCAGE, EURAD-IM, MATCH, DEHM, GEM-AQ and CAMS-global) and the lateral boundaries (left to right: south, west, north, east).
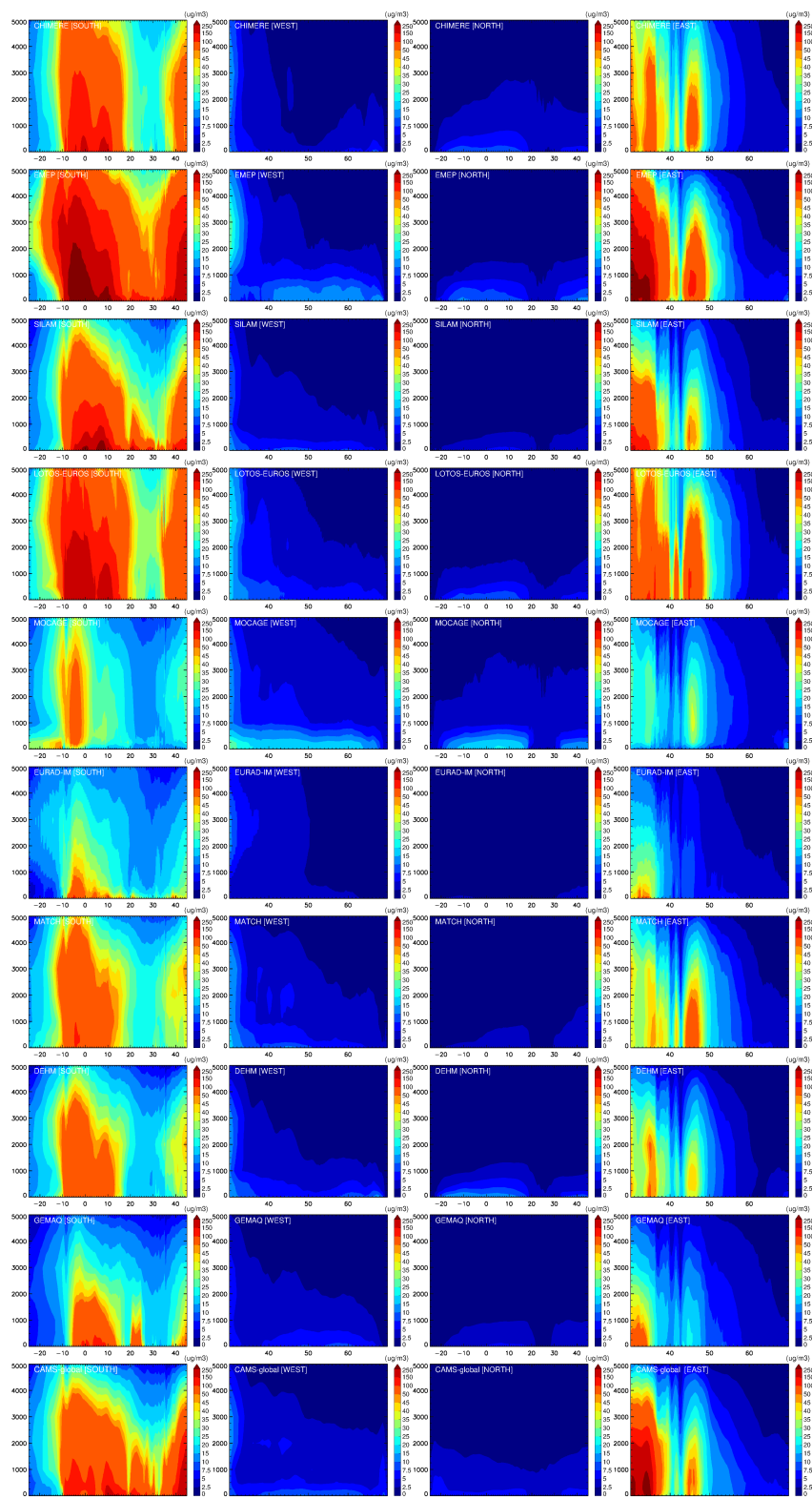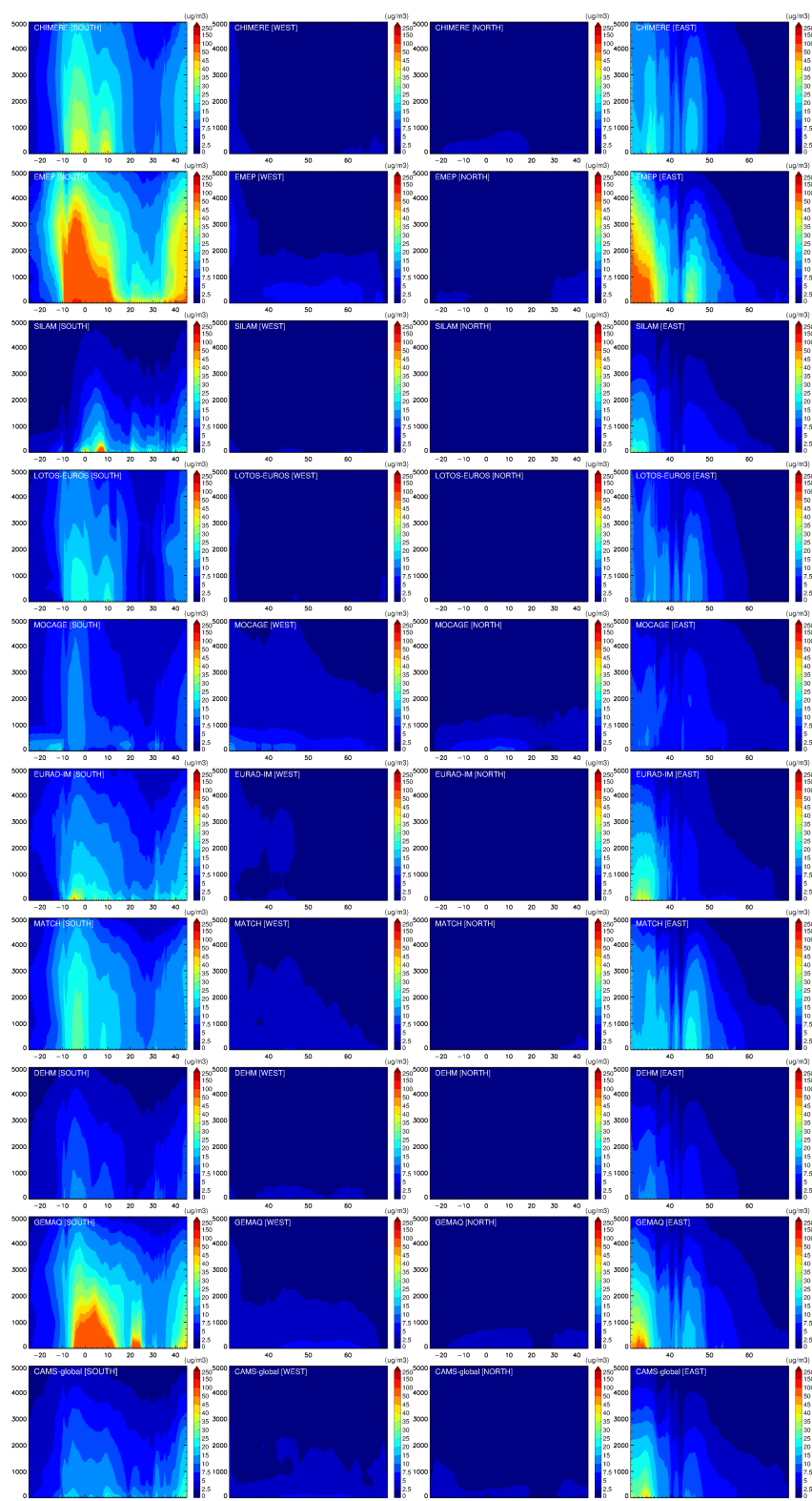
3.8   Regional analysis vs. regional forecasts

In the following sections we compare the regional analysis products with the regional forecasts (Day1). The four following figures (3.15-3.18) show the mean regional differences between analysis and forecasts for the time period JJA2020 at four different vertical layers (0, 250, 2000, 5000 m, left to right) including O_3, CO, PM10 and PM2.5, respectively.

Regional models with the largest differences between the analyses and the 1st day forecasts are for:

Ozone

- CHIMERE in the surface layer.
- EMEP and SILAM mostly within the PBL (patchy patterns)
- DEHM up to 2000 m.
- LOTOS-EUROS, MOCAGE, EURAD, MATCH in all vertical layers.
- EURAD exhibits strange features

Carbon monoxide
Differences are seen in

- CHIMERE and GEM-AQ at the surface
- EMEP up to 250 m
- SILAM and MATCH up to 2000 m altitude.

PM10 and PM2.5

- Minor differences in some models, most pronounced in MATCH for PM10 and PM2.5.

It is important to note that the differences observed between the analysis and the forecast do not only reflect the impact of the assimilation but may also result from differences in the model setup between the analysis systems and the forecast systems.

Figure 3.15. Mean regional O₃ differences between analysis and forecast for JJA2020 for four different vertical layers (0, 250, 2000, 5000 m) from the regional ENSEMBLE and individual ensemble members (top to bottom: CHIMERE, EMEP, SILAM, LOTOS-EUROS, MOCAGE, EURAD-IM, MATCH, DEHM, GEM-AQ and ENSEMBLE).
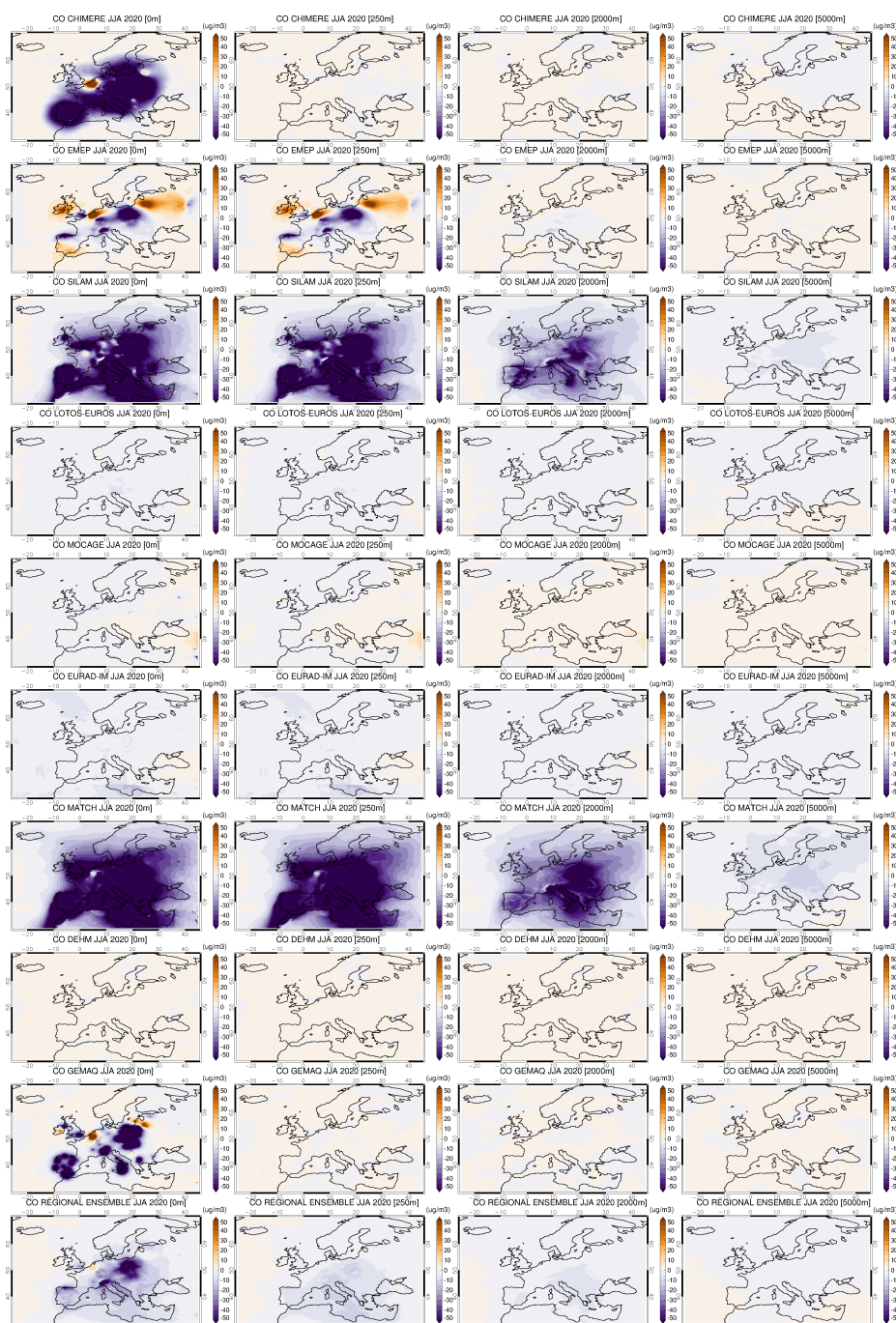
Figure 3.16. Mean regional CO differences between analysis and forecast for JJA2020 for four different vertical layers (0, 250, 2000, 5000 m) from the regional ENSEMBLE and individual ensemble members (top to bottom: CHIMERE, EMEP, SILAM, LOTOS-EUROS, MOCAGE, EURAD-IM, MATCH, DEHM, GEM-AQ and ENSEMBLE).
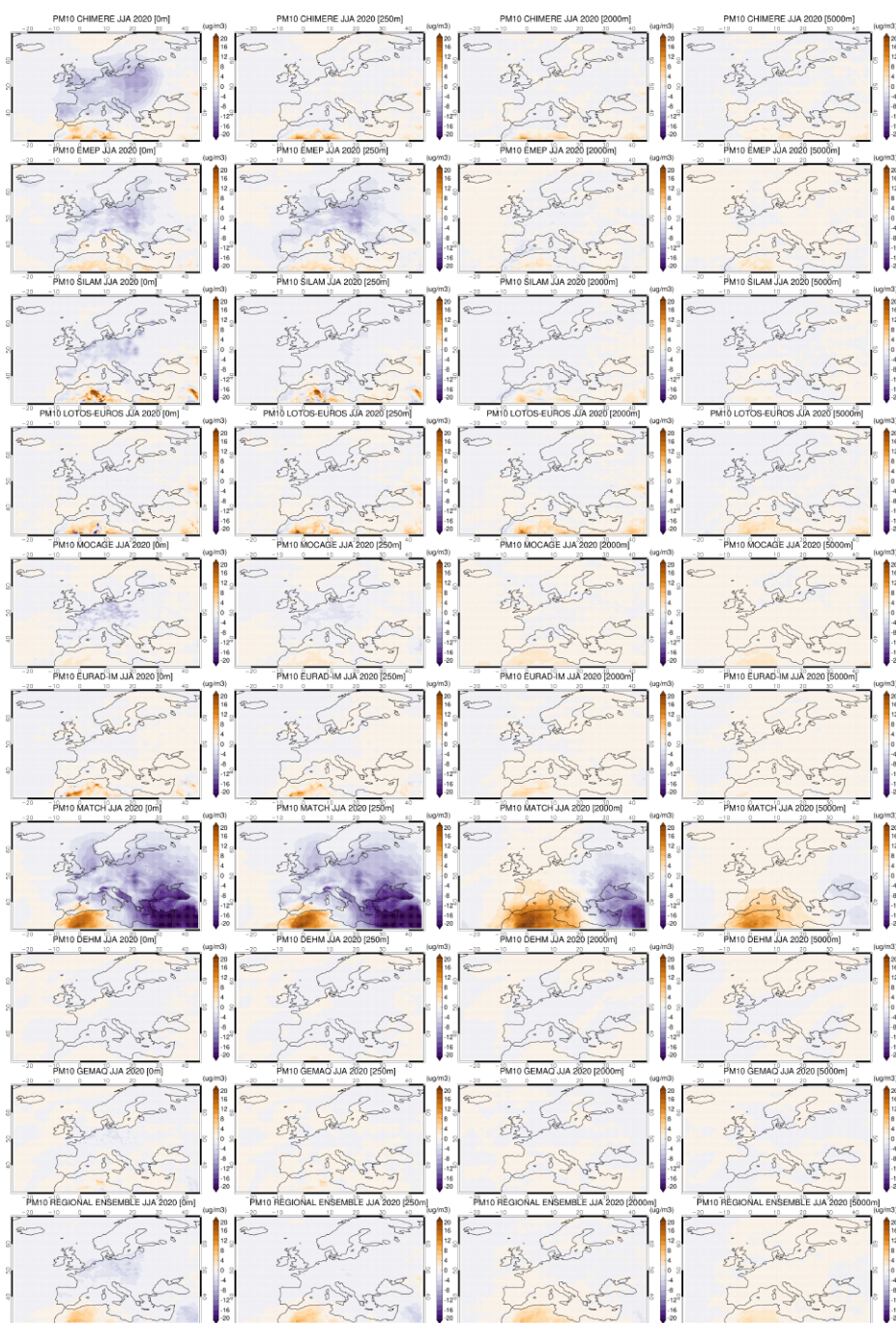
Figure 3.17. Mean regional PM10 differences between analysis and forecast for JJA2020 for four different vertical layers (0, 250, 2000, 5000 m) from the regional ENSEMBLE and individual ensemble members (top to bottom: CHIMERE, EMEP, SILAM, LOTOS-EUROS, MOCAGE, EURAD-IM, MATCH, DEHM, GEM-AQ and ENSEMBLE).
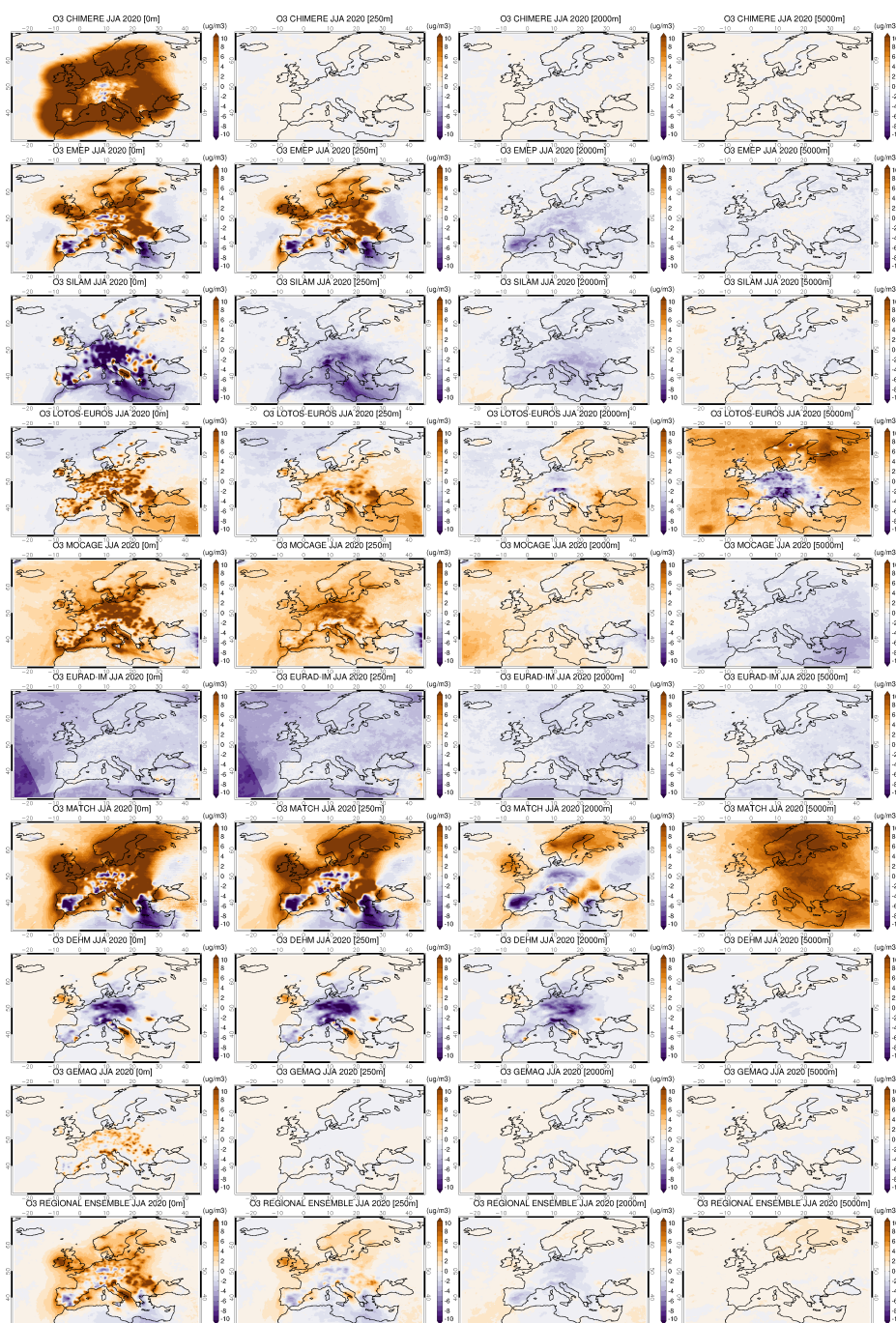
Figure 3.18. Mean regional PM2.5 differences between analysis and forecast for JJA2020 for four different vertical layers (0, 250, 2000, 5000 m) from the regional ENSEMBLE and individual ensemble members (top to bottom: CHIMERE, EMEP, SILAM, LOTOS-EUROS, MOCAGE, EURAD-IM, MATCH, DEHM, GEM-AQ and ENSEMBLE).
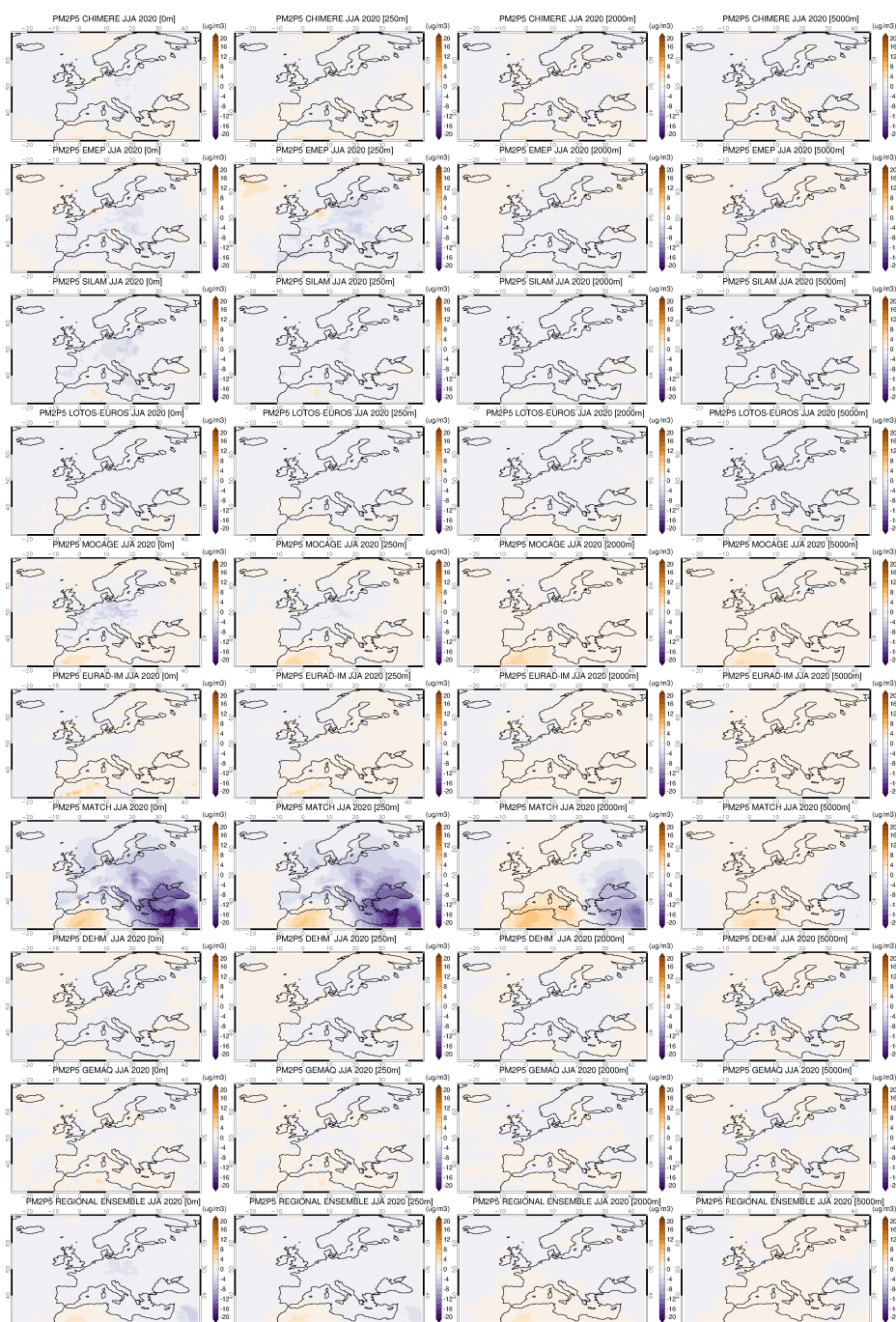
4 Vertical profile and column aerosol comparisons

4.1 Summary for the EARLINET lidar and Aeronet comparisons

The regional models are compared with climatological lidar profiles for each season (EARLINET/ACTRIS data from 2006-2018). Missing information on composition, size and humidity growth of the aerosol in the models introduces considerable uncertainty to the PM derived extinction, which conservatively spans up to a factor 10 for absolute extinction values. Aeronet data are used to calibrate the conversion from modelled mass to optical property aerosol extinction. This way the order of magnitude in extinction is similar between the models and the lidar profiles, but also significant differences appear at some stations in the lowest layers (Granada, Athens). Relative differences in the form of extinction profiles are more certain. We choose the most representative five stations to compare in retrospective the seasonal average aerosol profiles since 2016. The retrospective of the seasonal comparisons since 2016 shows very similar profiles during this season. The respective overestimation or underestimation of the extinction found in 2016 are usually also found in following years with the ENSEMBLE.

Introduction

The vertical distribution of aerosol reflects processes like atmospheric mixing, removal, and aerosol transport from outside of the domain or formation of secondary aerosol. The vertical mixing processes determine ground concentrations in polluted areas. Long-range transported aerosol, often carried aloft, may contribute to pollution in clean regions. Evaluation of the simulated aerosol column and vertical profiles are thus valuable for the performance characterisation of air quality models.

The 9 regional models provide mass concentration vertical profiles (PM2.5 and PM10) over Europe and may thus be evaluated for their aerosol vertical distribution. However, only very few aircraft campaigns and mountain sites are available to validate aerosol mass at altitude. In contrast frequent measurements of vertical profiles of aerosol backscatter, extinction or its integral, aerosol optical depth, exist. Deriving aerosol optical properties from the mass concentrations is thus needed, assuming lidar ratios and mass extinction coefficients, at least until the models provide more specific output on aerosol composition and optical properties.

In order to assess mass extinction coefficients chosen, the aerosol optical depth derived from the model mass profiles is first compared to AERONET AOD measurements. Secondly, we document a comparison of the extinction profiles derived from modelled mass concentration with climatological extinction profiles derived from European EARLINET/ACTRIS lidars.
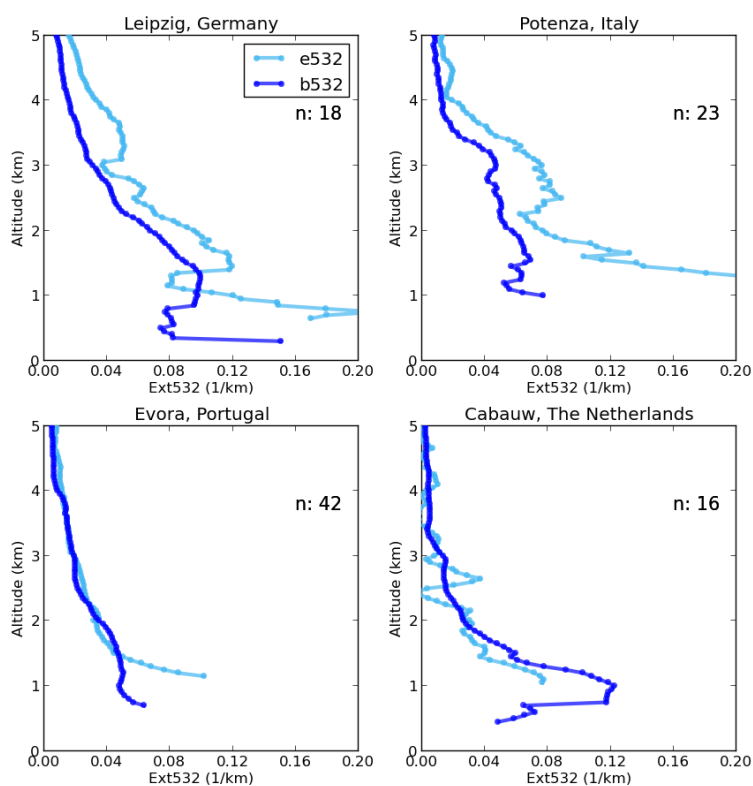
Figure 4.1. Comparison of mean profiles of two aerosol optical properties derived from the same Earlinet Raman lidar at the same time: a) aerosol extinction @532nm (light blue line) and b) extinction @532nm derived from the aerosol backscatter coefficient (dark blue) using a lidar ratio of 50 sr. The profiles use data taken at simultaneous times at each station; number of profiles given as n: x. Based on measurements in the period 2006-2014.

4.2   Methodology

AERONET data

The AERONET sun-photometers measure in non-cloudy conditions the aerosol optical depth at several wavelength and in near real time. The spatial distribution of the instruments allows a good coverage of aerosol observation over Europe. The version 3 level 1.5 data has been used in the reporting period presented here. This version and quality level ensures an efficient filtering of the residual clouds (mainly cirrus) for data in near real time. Daily AERONET aerosol optical depth, measured at 550 nm, has been averaged over summer for the European sites available in CAMS model output.

Lidar data

The EARLINET/ACTRIS Lidars are distributed over several locations in Europe and allow comparison across different climates in Europe (Pappalardo et al., 2014). Yet, up to now, there are no near-real-time data available. Regular measurements in EARLINET are sparse and often acquired once per week, with gaps due to maintenance or funding restrictions. A climatology has been computed per station and per season with all measurements available between 2006 and 2018.
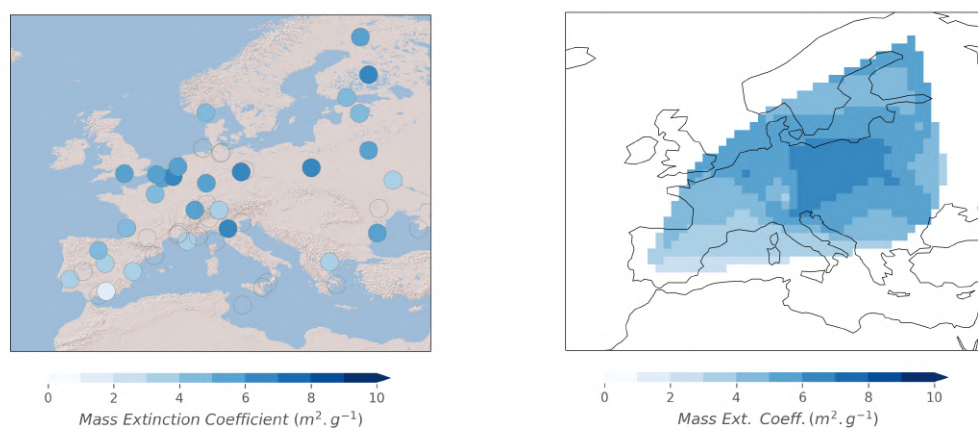
Figure 4.2. . Mass extinction coefficient estimated at AERONET location sites (blue dots) and EARLINET stations (triangles) for JJA2020 (left). A European map of the Mass Extinction Coefficient has been constructed with cubic interpolation in the inner part of the region covered by AERONET on the model grid (right), and with the nearest neighbours in the outer part of this region (not shown).

The backscatter coefficient and extinction profiles at 532 nm have been extracted from the EARLINET database. The more frequently measured backscatter profiles are considered here with priority. An aerosol extinction coefficient profile is computed from the backscatter coefficient using a range of plausible lidar ratios. This latter parameter depends on the aerosol type and is more likely decreasing with the size of the aerosol. Minimum values are observed for sea salt aerosol (below 30 sr at 550 nm, [Ackermann et al., 1998, Omar et al., 2009]), while larger values are related to urban particles (55 sr in [Muller et al., 2007], 70 sr in [Cattrall et al., 2005]). Desert dust is associated with intermediate lidar ratios, ranging from 30 sr to 60 sr depending on the sources and the transport regime. A climatology of aerosols in West Africa published in Mortier et al. [2016] revealed an average lidar ratio (over 9 years) of about 30±15 sr. Due to the location of the stations involved in this study, both dust and urban aerosols and any a mix of them might occur. In order to represent the uncertainty on the nature of aerosols, we show the range in likely mean extinction using a lidar ratio extending from 30 to 70 sr.

As a test, the conversion of lidar backscatter to extinction coefficient is performed assuming a constant average lidar ratio of 50 sr at locations where both backscatter and extinction coefficients are measured in EARLINET. This allows consistent comparison and visualization of some of the error associated to our simplified constant lidar ratio assumption. We have excluded from this comparison cases where local extinction coefficient was above 0.5 km^{-1} in order to avoid outliers. The profiles shown in figure 4.1 from 4 stations reveal an error in the mean profile of 0-30% in extinction, which is small compared to the model spread documented below. The extinction profiles derived from the backscatter coefficients look vertically smoother. We have excluded from this comparison cases where local extinction coefficient was above 0.5 km^{-1} in order to avoid outliers.

In addition to this aerosol typing uncertainty, a sampling error should be accounted for. The observations are sporadic, while the models predict the aerosol concentration continuously. Therefore, seasonal averages are not computed with the same coverage in model and observation. Our earlier model-based bootstrap studies revealed, that, depending on the station, a set of ca. 30 daily observations allows reproducing the seasonal average with an error of about 10% [ACTRIS Deliverable WP6/D6.21]. In our case, this error might be larger since the synoptic situation is very different between our EARLINET climatological dataset, covering 2006-2018, and the season covered in this report. An overall uncertainty of about 20% has been chosen to represent the sampling error.

Model data

The ENSEMBLE and the underlying 9 regional CAMS models are investigated. For each of these models, the hourly PM10 and PM2.5 vertical profiles are extracted at the EARLINET station locations from the first day of each daily forecast at levels 0, 50, 250, 500, 1000, 2000, 3000, 5000 m.

The conversion of PM10 mass concentration to extinction requires a mass extinction coefficient (MEC). MEC depends on the size distribution, refractive index and density of the particles. This information is not yet available from the models. For different kind of aerosols, MEC values can vary from about 0.5 m^2g^{-1} in the case of desert dust aerosols up to 8 m^2g^{-1} for urban particles [Chin et al., 2002]. No variation with height or aerosol type is taken into account, mainly because the models provide no further info on aerosol speciation. We derive the MEC value to convert the model profile data to extinction profiles from a combination of the modelled mass column load and consistent Aeronet AOD data.

For the Aeronet based computation of the MEC, the model data are picked at the location and on the day when sun photometer observations were available. The CAMS-regional mass concentrations have been averaged for coincident days (with the measurements) and averages are converted, with a seasonal and site dependent mass extinction coefficient estimated with AERONET retrievals, into extinction profiles. A seasonal and site dependent mass extinction coefficient is obtained when combining it with AERONET AOD retrievals. Values of MEC are ranging from 1 m^2g^{-1} in South-West of Europe to more than 10 m^2g^{-1} in the North-East. Since some of the EARLINET stations are not co-located with Sun photometers, a European map of MEC has been constructed, for each season, by interpolating (cubic interpolation) and extrapolating (nearest neighbour) the available AERONET based MEC calculations on the grid of the model (figure 4.2). One can notice a gradient with the longitude with lower values found in the Western part while the highest values are observed in the Eastern part of Europe. Also, the values are generally lower as compared to last year, which might reveal a higher concentration in coarse particles (dust). The seasonal AERONET-based MEC is then used at each EARLINET station to calculate the model extinction from concentration profiles. The uncertainty on the MEC being removed allows more accurate comparisons with the observed vertical profiles than using an average MEC over whole Europe.
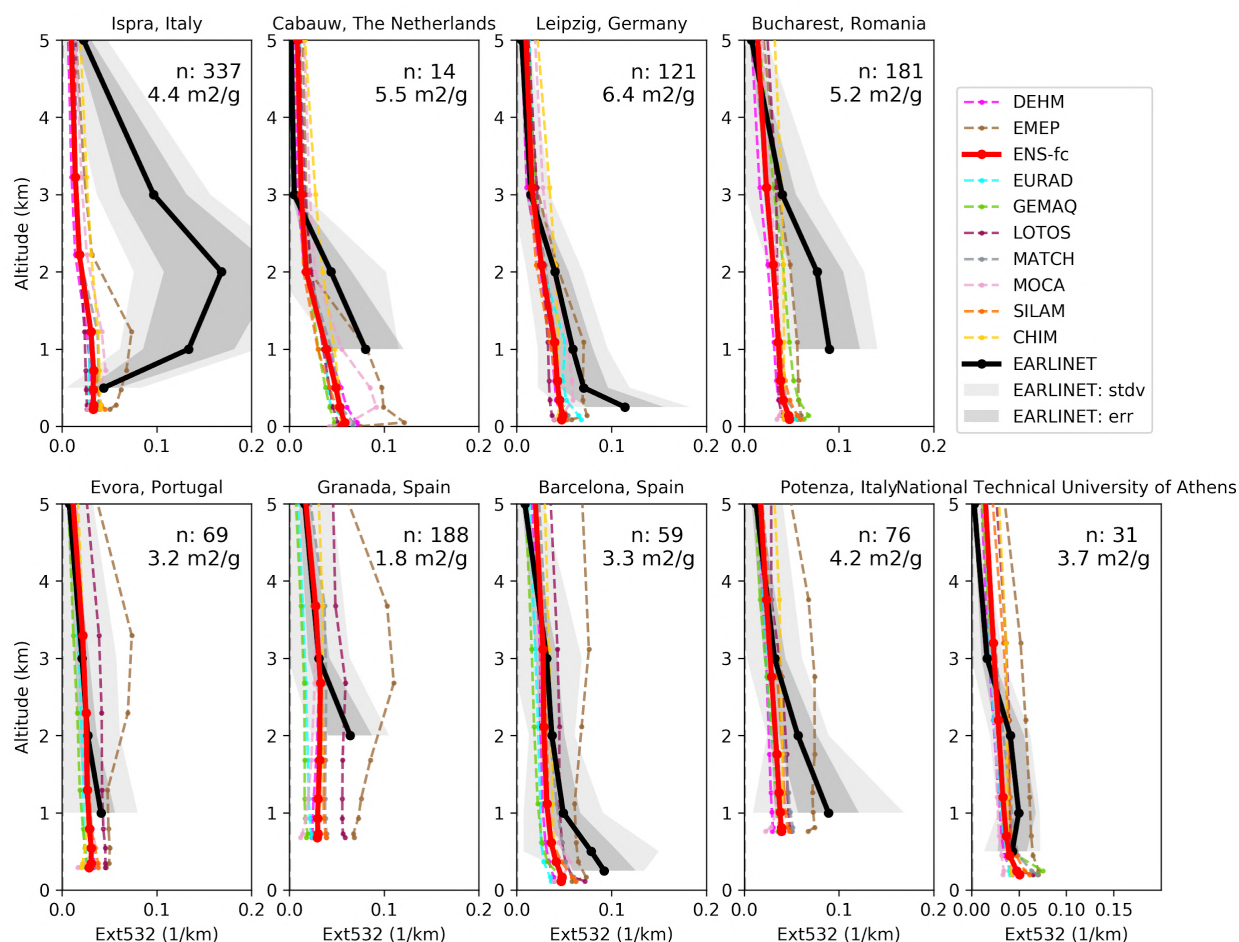
Fig 4.3. Extinction profiles June - August 2020 derived from the ENSEMBLE forecast mass concentration profiles (red envelope) and from EARLINET (climatology) backscatter profiles (grey envelope: lidar ratio uncertainty, light grey: including sampling error). "n: XX means number of individual EARLINET profiles assembled (June - August 2006-2018). The EMC used for the calculation of the extinction from the concentration profiles is indicated for each station below the number of EARLINET profiles "n" used for the calculation of the climatology.

4.3   Results

4.3.1 Comparison of extinction profiles

The extinction profiles estimated from the 9 CAMS models and the ENSEMBLE and EARLINET measurements are compared for JJA 2020 (figure 4.3).

Two new regional models (DEHM and GEMAQ) are plotted together with the seven usual CAMS regional models, in addition to the ENSEMBLE. This summer, the models are showing lower extinction values than the climatological observed values at most of the locations. The underestimation is the highest in the Northern stations, while a quite good agreement is found in Evora and in Athens where the profiles exhibit a flat vertical structure.
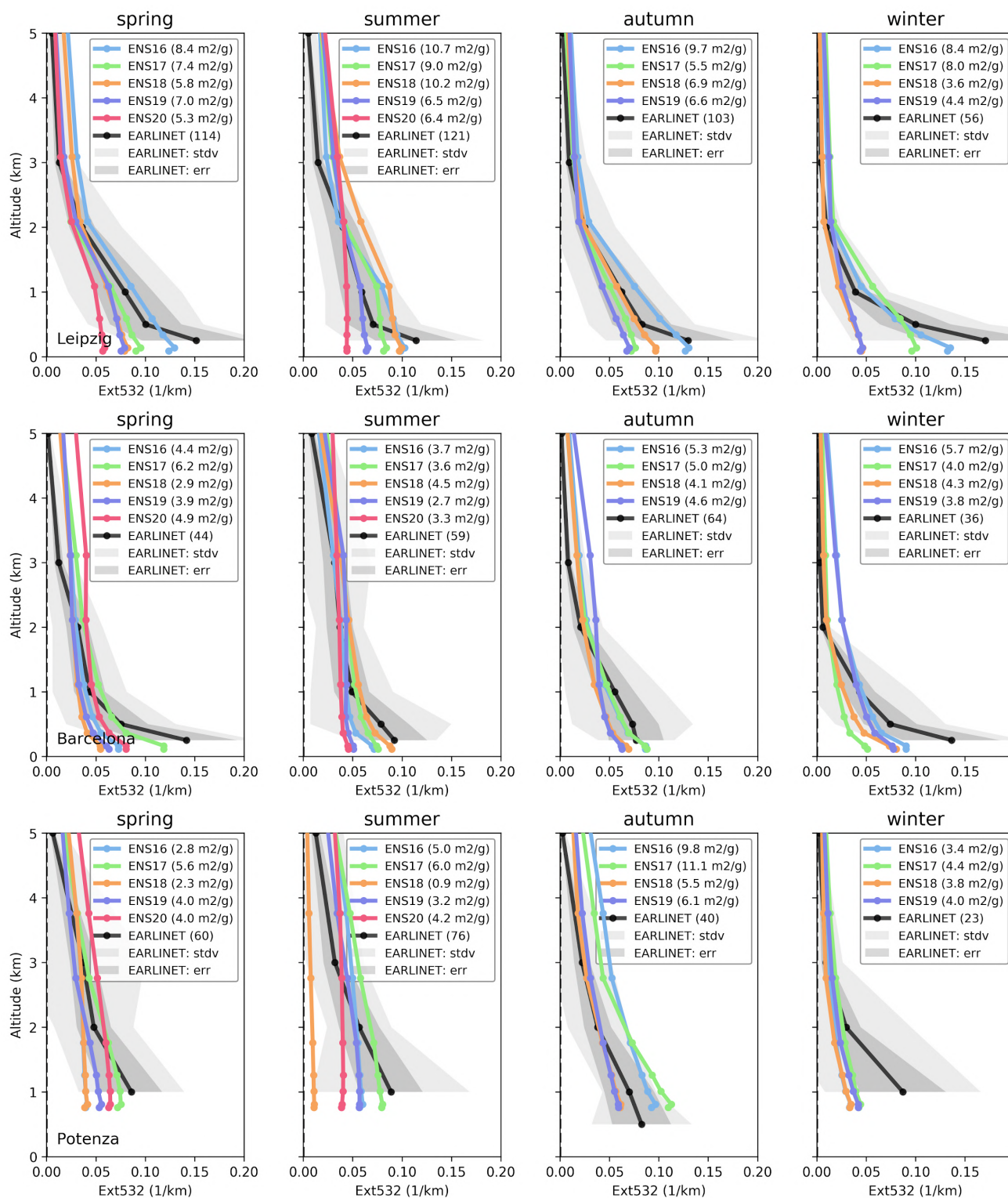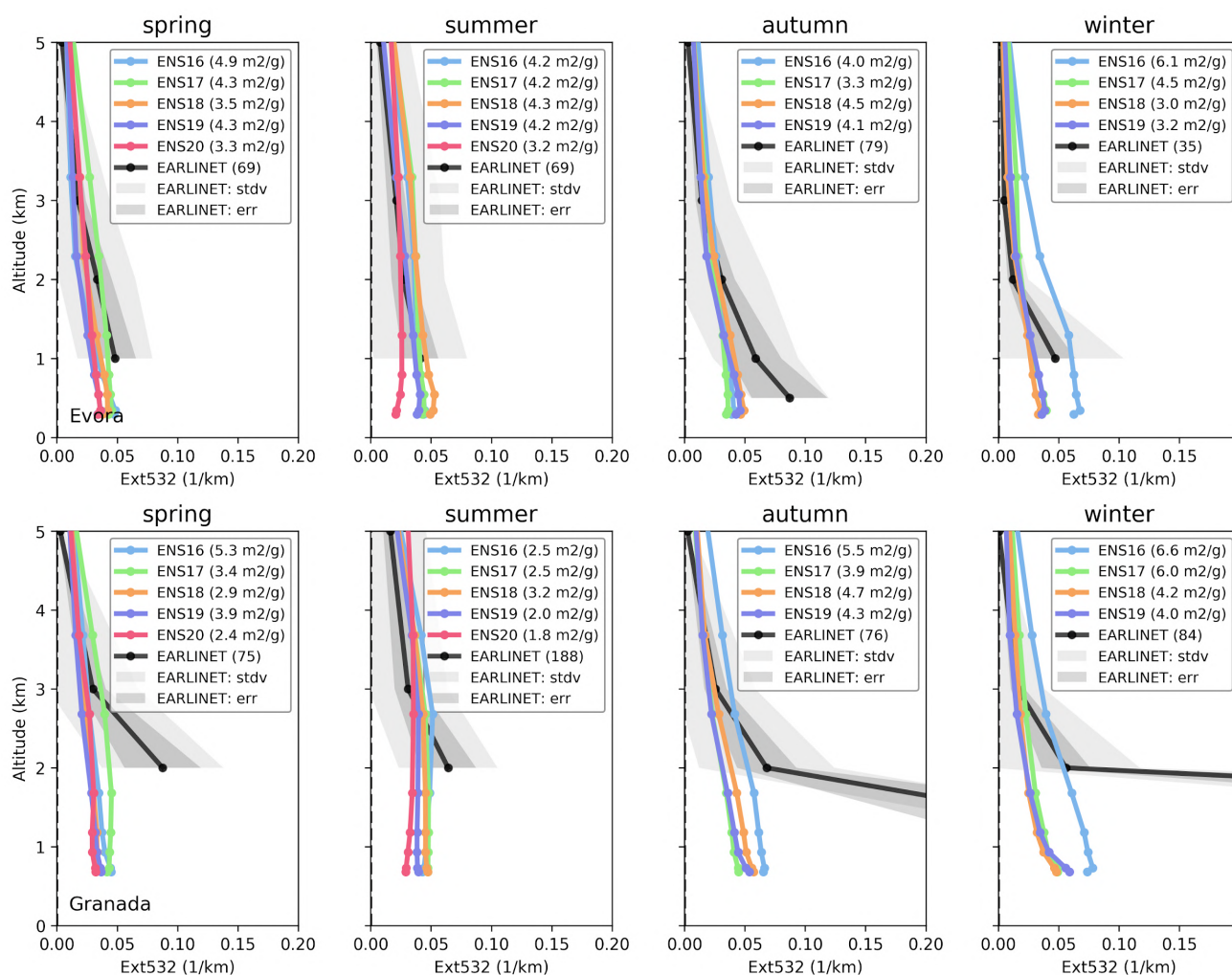
Fig 4.4: Seasonal extinction profiles derived from ENSEMBLE forecast mass concentration profiles for 2016 (ENS16), 2017 (ENS17), 2018 (ENS18), 2019 (ENS19), 2020 (ENS20) and EARLINET climatology. The parenthesis indicates for the CAMS profiles, the MEC used for the extinction estimation, and for the EARLINET, the number of profiles used for constructing the climatology.

4.3.2 Seasonal variability

In order to investigate the performance of the model in reproducing the vertical profiles, it is interesting to observe the inter-annual variability for the different seasons. This will be of use for the development of a score providing an assessment of the models skills, and is also useful to investigate the models synoptic variability. The seasonal profiles have been reported since 2016 at 5 stations (Leipzig, Barcelona, Potenza, Evora and Bucharest) in Figure 4.4, using the new EARLINET climatology. The ENSEMBLE model is showing lower extinction values below the altitude of 2 km this summer as compared to summer 2019. This usually implies a greater underestimate of the extinction in this vertical range.

5 IAGOS aircraft CO and O₃ profile comparisons

5.1 Summary

Routine observations of ozone and CO over European airports are available from the IAGOS fleet. Take-off and landing profiles were sampled from the hourly model 3D forecasts along the flight tracks.

Ozone in-situ: The regional ENSEMBLE and CAMS-global present very similar behaviour in all layers. In the lowest layers, ozone is well represented by the models with sometimes a better performance from the regional ENSEMBLE especially when ozone values are high with respect to climatology. The bias in the free troposphere is larger with overestimations from all models although a better performance is found with high ozone values.

Carbon monoxide: In the lowest layers, CO is mostly underestimated by the regional ENSEMBLE and CAMS global which present similar behaviour. In the free troposphere the agreement is better and CAMS-global performs better than the regional ENSEMBLE which underestimates.

IAGOS Validation Method

Validation is possible at the European airports visited by the IAGOS fleet. For the European-based carriers, there are regular profiles at the home airports. There are two aircraft operated by Lufthansa, one operated by Air France. Thus, when the fleet is fully operational, there are daily profiles Frankfurt and Paris (CDG). IAGOS is also installed on two aircraft operated by the Asian-based carrier China Airlines. Aircraft fly regularly from Taipei to Amsterdam or Vienna and sometimes to Rome. Other airports may be visited depending on the operational schedules of the airlines. Despite the end of the first COVID-19 pandemic lockdown over Europe, air traffic remains considerably reduced during this JJA period as compared to other years which obviously also affects IAGOS operations. Over Europe only the airport of Frankfurt has been visited as shown in Figure 5.1.

We download the daily latitude-longitude datasets for the 9 regional models and the ENSEMBLE for two species (carbon monoxide and ozone) on 8 vertical levels (surface, 50m, 250m, 500m, 1000m, 2000m, 3000m, 5000m). The aircraft takes about 10 minutes to climb or descend the 5000m vertical extent covered by the regional models. During this time and travelling at up to 166 m s^{-1}, it covers about 120km and therefore traverses many grid-boxes of resolution 10km. We perform a spatial interpolation from the grid of the regional models to the aircraft's trajectory. The IAGOS measurements in ppbv are converted to μg m^{-3} using the temperatures measured by IAGOS. The data are validated by the PI but are not yet calibrated. Calibration takes place after an operational period of about 6 months.
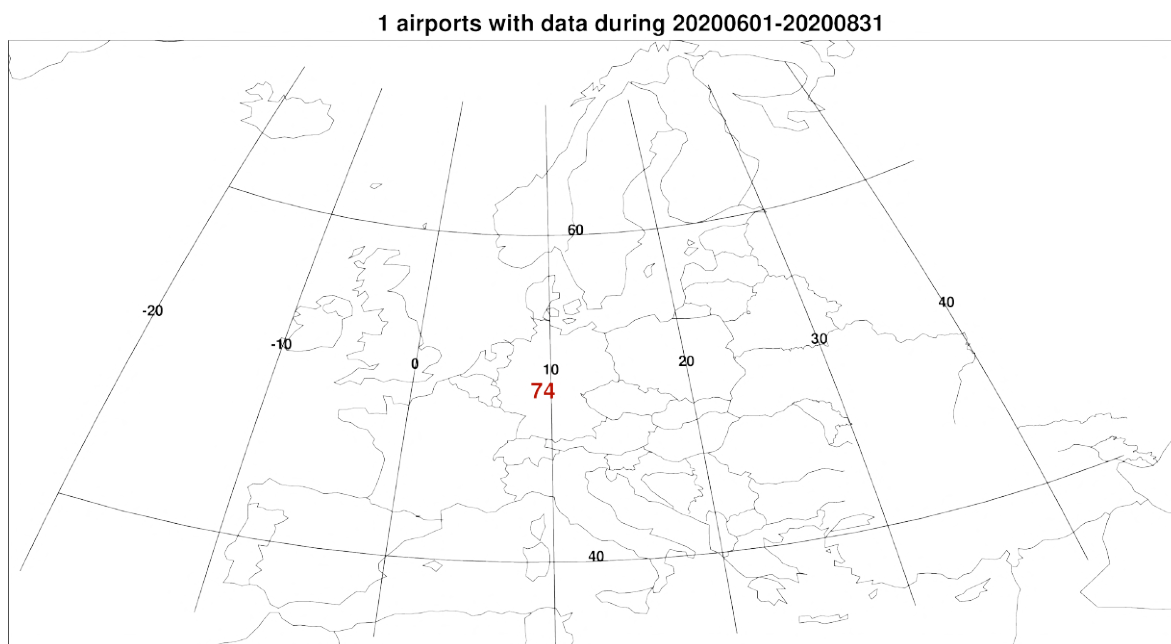
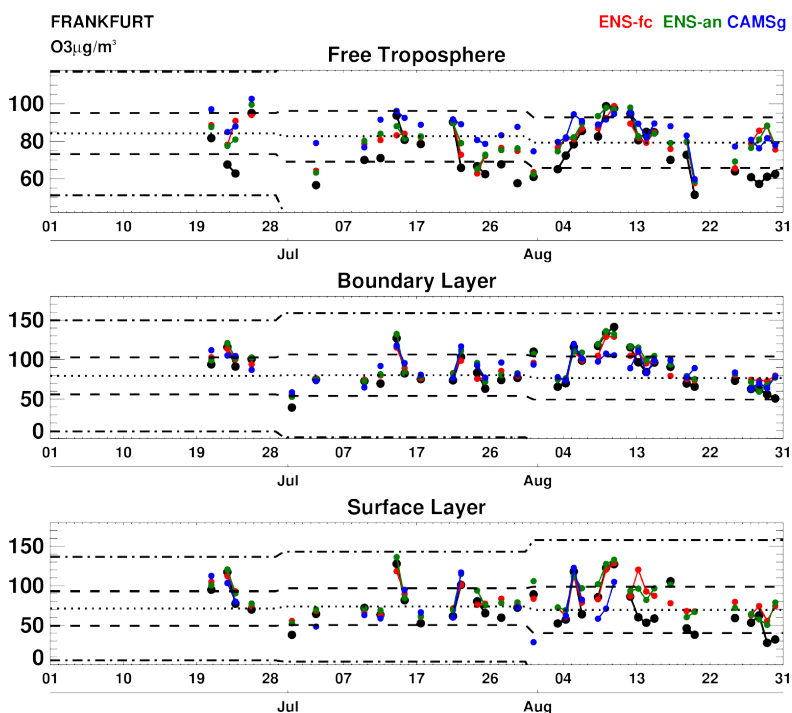Figure 5.1. Map showing the number of profiles available in the period June - August 2020.

Figure 5.2: : Daily time series of ozone at Frankfurt for the period JJA2020. IAGOS observations are shown in black, the regional ENSEMBLE and associated analysis are shown in red and green respectively, and CAMS-global (o-suite) is shown in blue. The black dotted line is the monthly mean of the observations over the period 2003-2016 (IAGOS/MOZAIC, Level 2 data), the black dashed line shows 1 standard deviation from the monthly mean and the black doted-dashed line shows 3 standard deviations from the monthly mean. (Units: µg.m^{-3}).
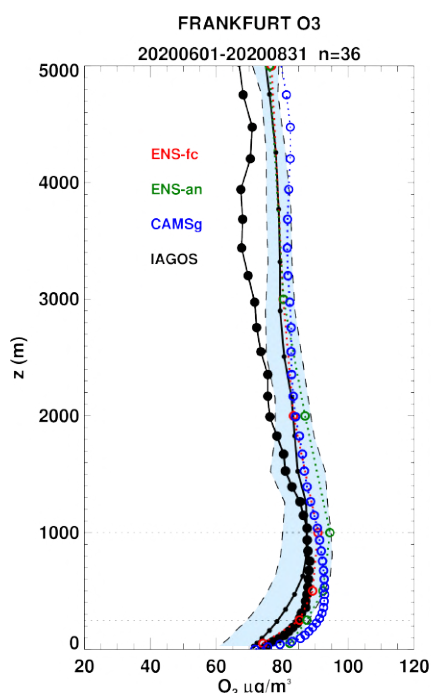
Figure 5.3: Mean profile of ozone at Frankfurt for the period JJA2020. IAGOS observations are shown in black, the regional ENSEMBLE and associated analysis are shown in red and green respectively, and the global o-suite is shown in blue. The shaded area indicates the range of the mean climatology of the observations plus/minus one standard deviation during the same period for all years between 2003 and 2016 (IAGOS/MOZAIC, level 2 data). (Units: $\mu g.m^{-3}$).

5.2   IAGOS Ozone

For the period June - August 2020, observations are available only at Frankfurt with discontinuities: sampling is very poor during the month of June, and many gaps are found in both July and August. The daily time series are presented in Figure 5.2 and associated averaged profiles over the whole period are presented in Fig. 5.3. In all plots, the forecast and analysis of the regional ENSEMBLE is shown in red and green respectively, and CAMS-global analysis is shown in blue. Fig. 5.2 and 5.3 show that the regional and global models differ mostly at around 1 km. On average, in the lowest layers the ENSEMBLE performs slightly better than CAMS global presenting small overestimations (Fig. 5.3). In the upper layers, the regional and global models behave more similarly with mostly large overestimations (Fig. 5.3) and a better agreement with high ozone values (Fig.5.2).

During this JJA period, IAGOS ozone observations in the surface and boundary layer are mostly within one standard deviation from the climatological mean. The comparison with the climatological variability shows that the positive anomaly of ozone in the low troposphere that was found in the previous quarter period is not present for JJA 2020 (Fig. 5.3, here and in MAM 2020 report). However four peaks of ozone are found in the surface and boundary layer which are above one standard deviation from the mean around days of 23 June, of 14 July and of 5 and 10 August. The maximum value in the surface layer reaches nearly 3 standard deviation from the climatology around 15 July. The individual profiles related to these increases in ozone are presented in Fig. 5.4.a-c. These individual profiles show that surface ozone reaches 120 $\mu g.m^{-3}$ on 24 June, 130 $\mu g.m^{-3}$ on 14 July, 120 $\mu g.m^{-3}$ on 5 August and 160 $\mu g.m^{-3}$ on 10 August, with boundary values of the same magnitude.
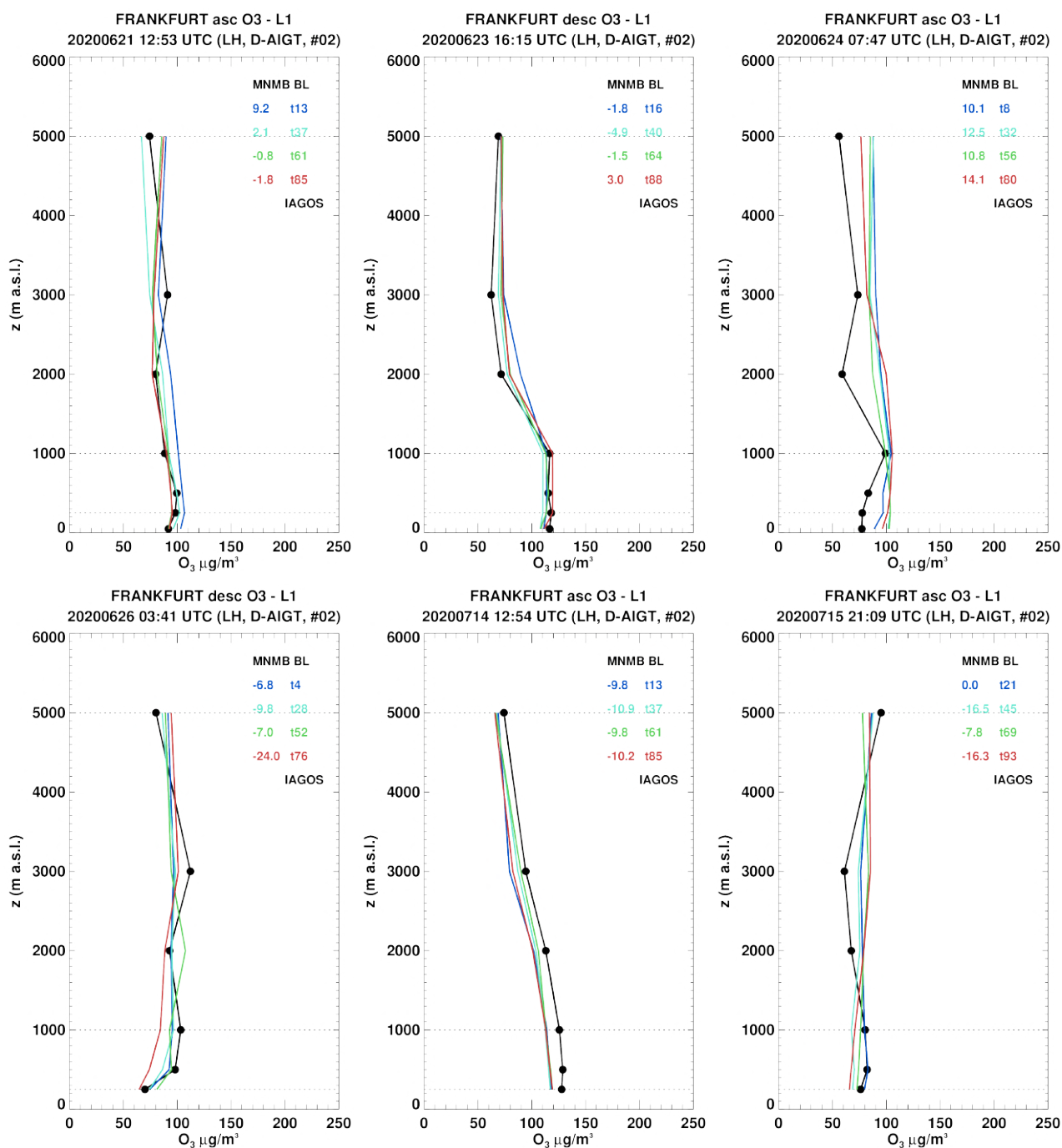
Figure 5.4.a. Selection of ozone profiles at Frankfurt during the period JJA2020. IAGOS is shown in black and the ENSEMBLE is shown at 4 forecast times (blue: 1-day; cyan: 2-day; green: 3-day; red: 4-day). (Units: µg m^{-3}).
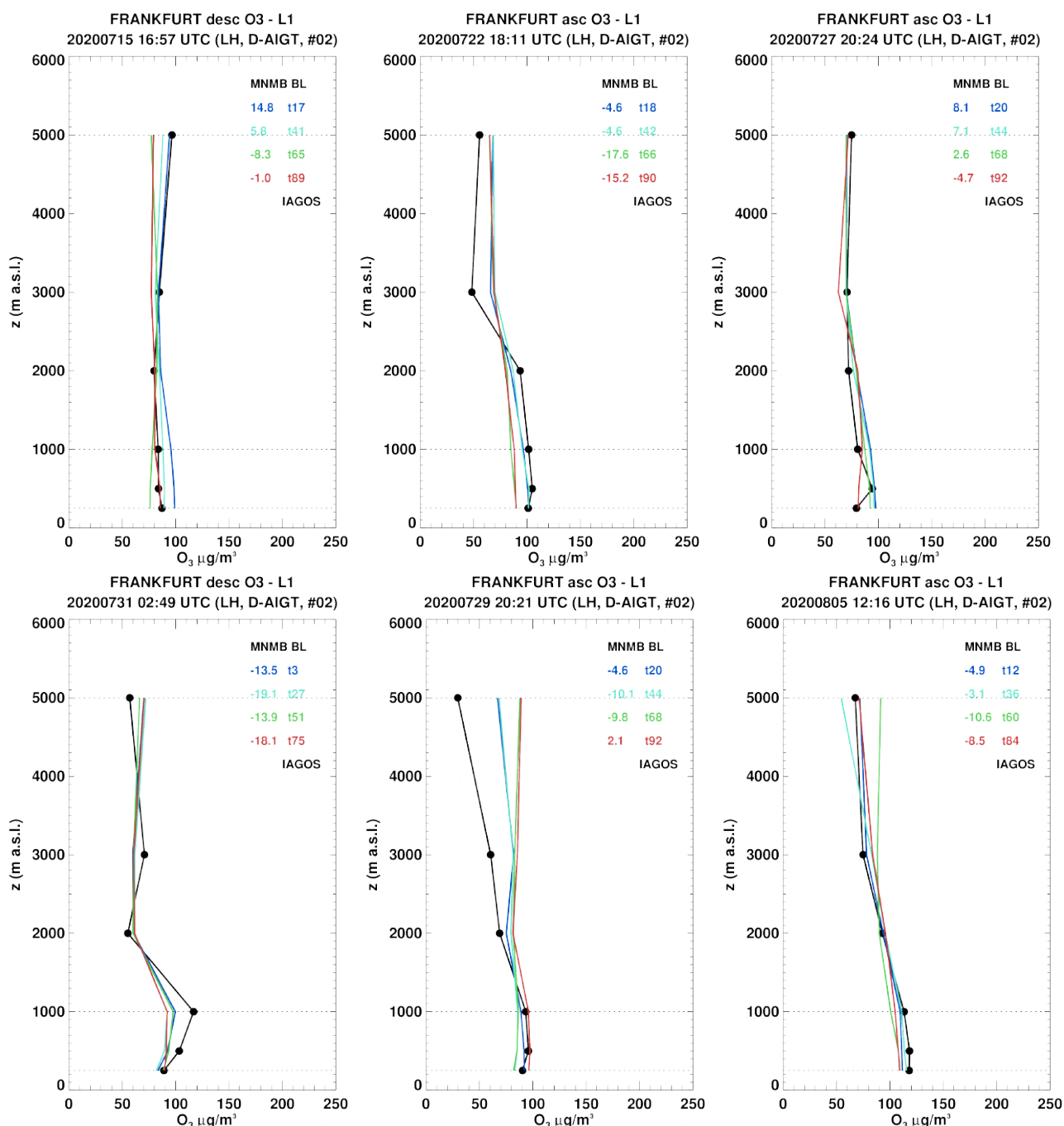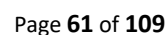
Figure 5.4.b. Selection of ozone profiles at Frankfurt during the period JJA2020. IAGOS is shown in black and the ENSEMBLE is shown at 4 forecast times (blue: 1-day; cyan: 2-day; green: 3-day; red: 4-day). (Units: µg m^{-3}).

These increases in ozone are well represented by the models (Fig. 5.3 and 5.4.a-c) and for all profiles the results for the different forecast time are similar. Although and as mentioned before CAMS global and regional often present similar results in the lowest layers for this period (Fig. 5.3), the regional ENSEMBLE provides a better performance in some cases of high ozone values. This is the case for instance on 10 August as shown in Fig. 5.3 (see also 5.4.c).

Figure 5.4.c. Selection of ozone profiles at Frankfurt during the period JJA 2020. IAGOS is shown in black and the ENSEMBLE is shown at 4 forecast times (blue: 1-day; cyan: 2-day; green: 3-day; red: 4-day). (Units: µg m⁻³).

For most profiles of the period the results from all models are similar. The profiles in which the results of the individual models differ are presented in Fig. 5.5.a-c. The MOCAGE and GEM-AQ sometimes present much larger bias in the low troposphere as compared to other models (0621 12:53, 0624 07:47, 0714 12:54, 0715 21:09, 0722 18:11, 0725 11:54, 0727 08:31, 0729 20:21, 0819 09:03). As observed in the previous quarter period, the EURAD model also presents here some large biases in

the free troposphere on two profiles (0714 12:54, 0721 03:12). For several profiles DEHM provides large values in the free troposphere, while SILAM underestimation in the lowest layers in comparison to other model results (0717 02:54, 0729 03:20, 0815 18:24, 0819 03:03).
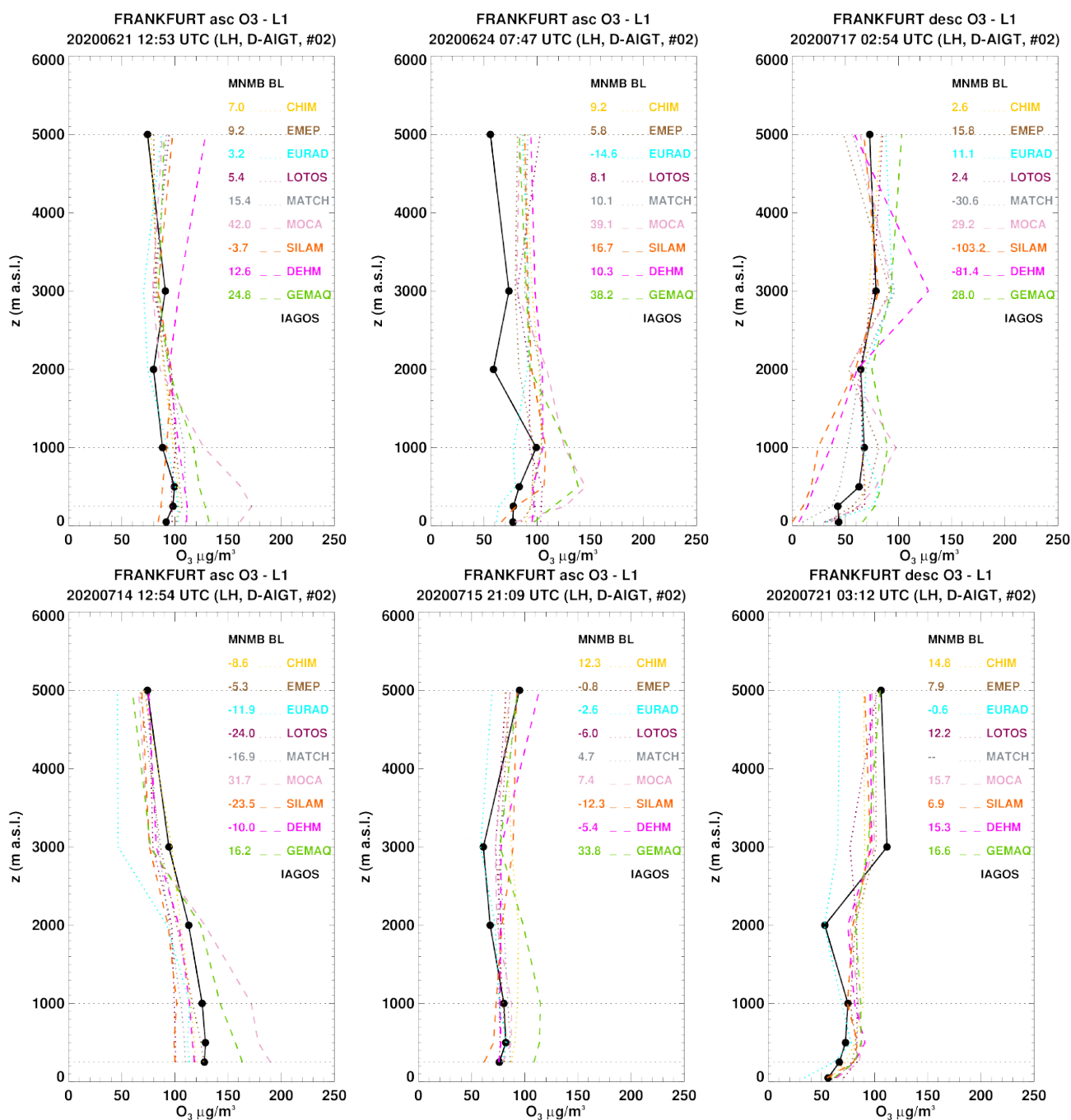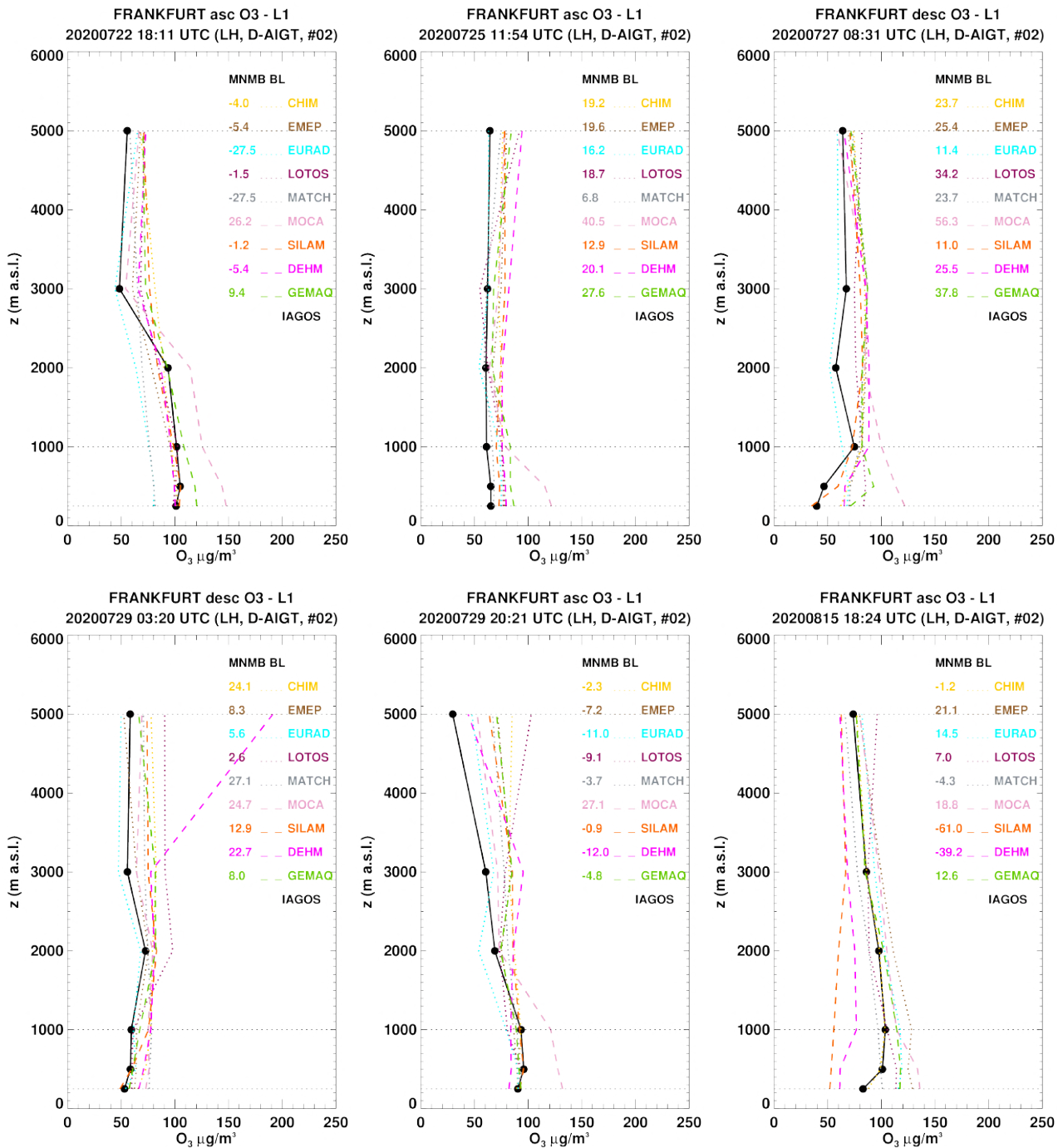
Figure 5.5.a: Ozone profiles at Frankfurt during the period JJA 2020. IAGOS is shown in black and each colour/Line style corresponds to one of the 9 models for the ENSEMBLE. In the legend, the models are CHIM=CHIMERE, EMEP=EMEP, SILAM=SILAM, LOTOS=LOTOS-EUROS, MOCA=MOCAGE, EURAD=EURAD, MATCH=MATCH, DEHM=DEHM, GEMAQ=GEM-AQ. (Units: µg m^{-3}).
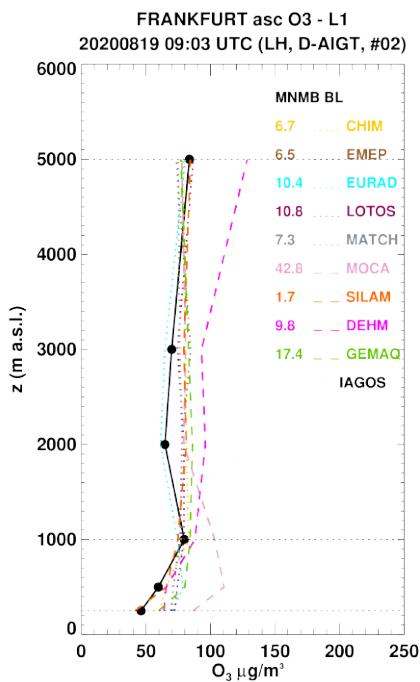
Figure 5.5.b: Ozone profiles at Frankfurt during the period JJA 2020. IAGOS is shown in black and each colour/Line style corresponds to one of the 9 models for the ENSEMBLE. In the legend, the models are CHIM=CHIMERE, EMEP=EMEP, SILAM=SILAM, LOTOS=LOTOS-EUROS, MOCA=MOCAGE, EURAD=EURAD, MATCH=MATCH, DEHM=DEHM, GEMAQ=GEM-AQ. (Units: μg m^{-3}).

Figure 5.5.c: Ozone profiles at Frankfurt during the period JJA 2020. IAGOS is shown in black and each colour/Line style corresponds to one of the 9 models for the ENSEMBLE. In the legend, the models are CHIM=CHIMERE, EMEP=EMEP, SILAM=SILAM, LOTOS=LOTOS-EUROS, MOCA=MOCAGE, EURAD=EURAD, MATCH=MATCH, DEHM=DEHM, GEMAQ=GEM-AQ. (Units: μg m^{-3}).
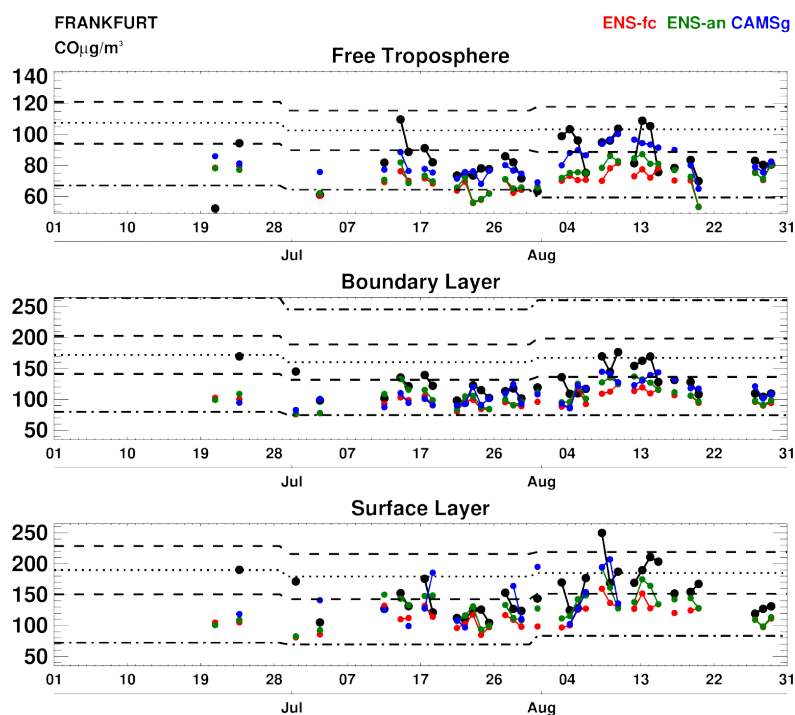
Figure 5.6: Daily time series of CO at Frankfurt for the period JJA 2020. IAGOS observations are shown in black, the regional ENSEMBLE and associated analysis are shown in red and green respectively, and CAMS-global (o-suite) is shown in blue. The black dotted line is the monthly mean of the observations over the period 2003-2016 (IAGOS/MOZAIC, Level 2 data), the black dashed line shows 1 standard deviation from the monthly mean and the black doted-dashed line shows 3 standard deviations from the monthly mean. (Units: $\mu g.m^{-3}$)

5.3   IAGOS Carbon Monoxide

Like for the ozone section, the daily time series of CO and associated averaged profile at Frankfurt are presented in Fig. 5.6 and 5.7. The forecast and analysis of the regional ENSEMBLE is shown in red and green respectively, and CAMS-global is shown in blue. On the time series, available observations of CO in the lowest layers are mostly below or close to the climatological values and often reaching values below one standard deviation from the mean. Two exceptions are found in the surface layer where values are close to one standard deviation above the climatology during the months of August (Fig. 5.6 and Fig. 5.8).

These low values of CO in the lowest layers lead to a negative anomaly for this JJA period as it was already the case for the previous quarter MAM 2020. This can be clearly seen on the average profile in Fig. 5.7 as compared to the climatological profile. This effect is likely related to the persistence of an emission reduction due to COVID-19 crisis after the lockdown period although to a lesser extent.
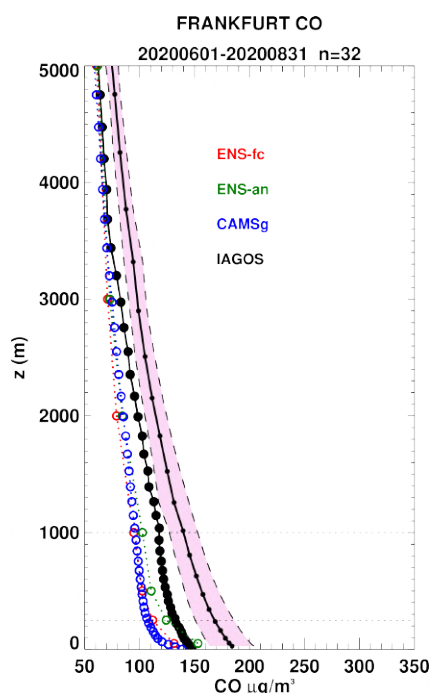
Figure 5.7: Mean profile of ozone at Frankfurt for the period JJA 2020. IAGOS observations are shown in black, the regional ENSEMBLE and associated analysis are shown in red and green respectively, and the global o-suite is shown in blue. The shaded area indicates the range of the mean climatology of the observations plus/minus one standard deviation during the same period for all years between 2003 and 2016 (IAGOS/MOZAIC, level 2 data). (Units: $\mu g.m^{-3}$).

In the surface and boundary layers CO is always underestimated by both regional and global models (Fig 5.6) with similar performance for both models (Fig. 5.7). However, in some cases, the analysis of the regional model clearly presents the best agreement (Fig. 5.6). In the free troposphere, the agreement is better than in the lowest layers with smaller underestimations (Fig. 5.6 and 5.7), and from the time series CAMS-global is often performing better than the regional ENSEMBLE (Fig. 5.6).

Some of the individual profiles at Frankfurt of the ENSEMBLE are presented in Fig. 5.8. These profiles correspond to the highest CO values observed in the low troposphere during this JJA period. Highest observed values of CO in the surface and boundary layers are roughly between 200 and 300 $\mu g.m^{-3}$, reaching a maximum of 300 $\mu g.m^{-3}$ in the surface layer on 8 August at 04:19 (Fig. 5.8.). As shown in Fig. 5.8, these CO values are often largely underestimated by the regional model in the lowest layers. No notable difference is found between the different forecast times.

Mean profile of ozone at Frankfurt for the period JJA 2020. IAGOS observations are shown in black, the regional ENSEMBLE and associated analysis are shown in red and green respectively, and the global o-suite is shown in blue. The shaded area indicates the range of the mean climatology of the observations plus/minus one standard deviation during the same period for all years between 2003 and 2016 (IAGOS/MOZAIC, level 2 data). (Units: $\mu g.m^{-3}$).
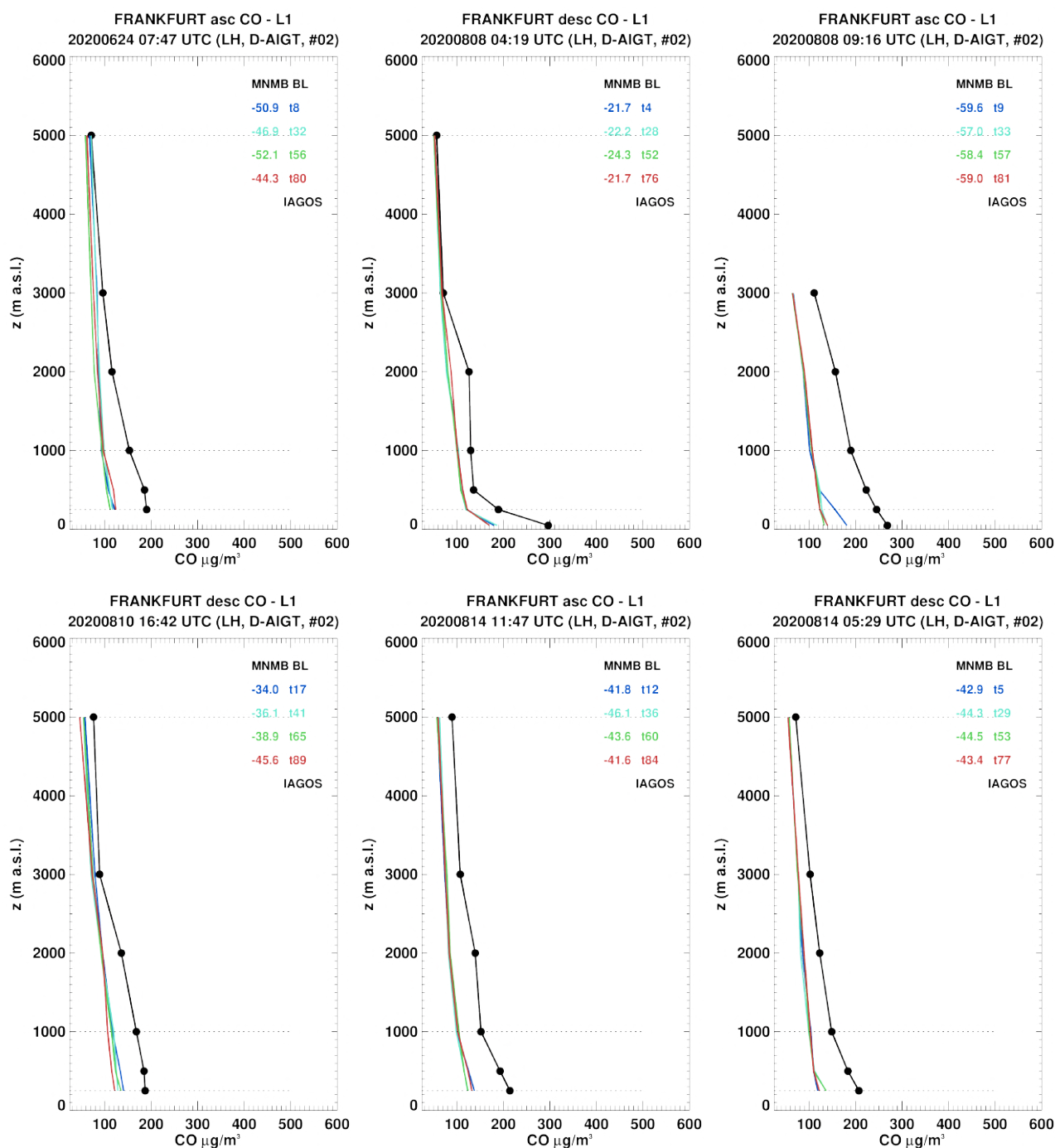
Figure 5.8: Selection of CO profiles at Frankfurt during the period JJA 2020. IAGOS is shown in black and the ENSEMBLE is shown at 4 forecast times (blue: 1-day; cyan: 2-day; green: 3-day; red: 4-day). (Units: $\mu g.m^3$).
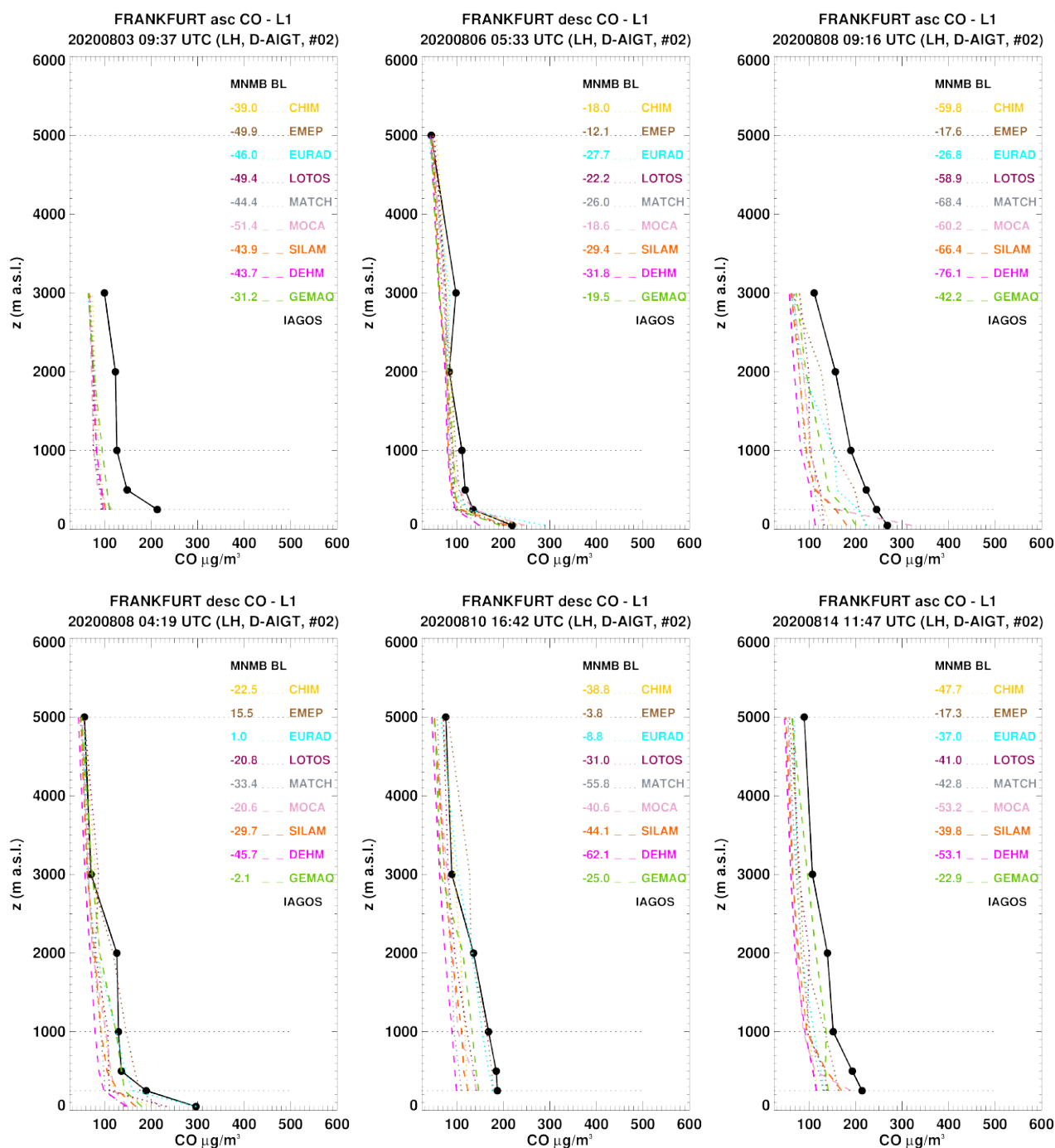
Figure 5.9: CO profiles at Frankfurt during the period JJA 2020. IAGOS is shown in black and each color/ line style corresponds to one of the 9 models for the ENSEMBLE. In the legend, the models are CHIM=CHIMERE, EMEP=EMEP, SILAM=SILAM, LOTOS=LOTOS-EUROS, MOCA=MOCAGE, EURAD=EURAD, MATCH=MATCH, DEHM=DEHM, GEMAQ=GEM-AQ. (Units: $\mu g.m^3$).

As regards the results of the individual models (Fig. 5.9), there are in general very similar for all profiles of JJA 2020. Although rare, some differences are found in the surface and boundary layer for profiles with the highest CO values in the lowest layers (here between about 200 and 300 µg.m^3). The spread of the bias from the different models is often much larger in these cases.

6 Validation of regional model tropospheric NO₂ using MAX-DOAS

6.1 Summary

MAX-DOAS surface remote sensing observations provide tropospheric columns of NO_2, with the largest sensitivity in the boundary layer. While the magnitude of VCDs derived from the measurements for the urban stations De Bilt and Bremen are reproduced well by the model ENSEMBLE, the ENSEMBLE underestimates the values in Athens (bias ~-2 x 10^{15} molec. cm^{-2}). The spread between individual models is rather large depending on the station and large differences are therefore found depending on the station and model. Although many of the model simulated values probably fall within the uncertainty range of MAX-DOAS retrievals, the latter alone cannot explain differences between retrievals and simulations, especially those found for variations in time. For the regional ENSEMBLE, moderate correlations on the order of 15-55 % are found for each station. The regional ENSEMBLE performs significantly better than the global model in terms of correlation at all stations except for De Bilt. However, some of the larger NO_2 values inside individual pollution plumes are underestimated by the models.

6.2 Introduction

MAX-DOAS observations of atmospheric composition are performed by taking measurements of the scattered sunlight at different elevation (and sometimes also azimuthal) angles. Depending on the viewing angle and solar position, the light path through the atmosphere is different, with the observation in the zenith direction usually providing the shortest light path through the lower troposphere. Therefore, using the zenith measurement as intensity of incident radiation and the observations in other angles as intensity of transmitted radiation, the total amount of molecules of a certain species along the light path difference, the so-called slant column densities, can be determined using Lambert Beer's law. Using radiative transfer modelling and Optimal Estimation techniques, this can be inverted to tropospheric columns and even lower altitude tropospheric profiles.

The advantage of MAX-DOAS measurements is their ability to observe several pollution related species at the same time (e.g. NO_2, HCHO, CHOCHO, SO_2, aerosols, potentially also O_3) and to provide data which is virtually free of interferences from other species such as PAN or NO_y for NO_2. Also, the fact that the observations integrate over a comparatively large volume can be an advantage for satellite and model validation as the observed quantity is relatively close to the modelled one. On the other hand, the uncertainty of the retrievals is considerable (on the order of 30% for NO_2 tropospheric columns and larger for individual layers) and depends on cloud occurrence and aerosol loading.

In this report, regional air quality model forecasts of tropospheric NO_2 columns are compared to MAX-DOAS retrievals from 3 urban stations (De Bilt– KNMI, Bremen – IUP-UB, Athens – IUP-UB). The reader is referred to previous reports for comparisons from the rural station OHP (BIRA-IASB) (which showed in general an underestimation by the model ENSEMBLE and an overall better performance for CAMS-global here) as the instrument at this site stopped working in March 2017. The MAX-DOAS instrument for the urban site in Uccle (BIRA-IASB) was dismantled in March 2020, and comparisons for this station are therefore not continued since this time. Since the MAM 2020 report, the MAX-DOAS data used in

the comparisons is based on tropospheric column retrievals, meaning that no profile information is incorporated in the comparisons since this time. A simple block profile is assumed in the MAX-DOAS retrievals and column averaging kernels are estimated based on the box air mass factor for each observation layer. An overview of the station data is given by Table 6.1. The MAX-DOAS instrument in Athens was not working correctly in August 2020 and measurements are therefore not available for that month.

Note that in contrast to previous reports, a shorter time period (i.e. JJA2020) was investigated in this report for consistency as the two new models were introduced in our analysis.

6.3   Inter-comparison method

Model VCDs (vertical column densities) have been calculated based on regional model data interpolated to MAX-DOAS output altitudes. Column averaging kernels (AVKs) from the measurements were applied to model NO_2 partial columns before summing up NO_2 values in the vertical:

$$VCD_{method2}^{model} = \sum_{i=1}^{Nobs} AVK_i \cdot VCD_i^{model}$$

The averaging kernels are part of the profiling output and represent the sensitivity of the retrieved column to the NO_2 amount at different altitudes. As the sensitivity of MAX-DOAS retrievals is largest in the boundary layer, the application of averaging kernels from the measurements to model simulations can have a crucial influence on validation results.

Only those model values closest to the measurement time are used below. As the model output is given in hourly time steps, the maximum possible time difference between measurements and simulations shown here is 30 minutes.

6.4   Results

Figure 6.1 shows time series of tropospheric NO_2 VCDs derived from MAX-DOAS for the model ENSEMBLE. The magnitude of VCDs from the measurements for De Bilt and Bremen is reproduced by the models, an understimation is found for Athens. The underestimation for Athens may be related to problems in simulating vertical transport of pollution within the boundary layer, as the instrument is placed on a small hillside above the city centre and the comparisons for Athens are therefore representative for altitudes larger than approximately 500 m absl. The retrievals show a larger variability of values compared to the models. Measurements and simulations don't agree very well for some of the time steps investigated. The models underestimate some of the larger NO_2 values inside individual pollution plumes. Models may fail to reproduce these peaks due to errors in NOx emissions, transport of NO_2 towards the stations and chemistry.
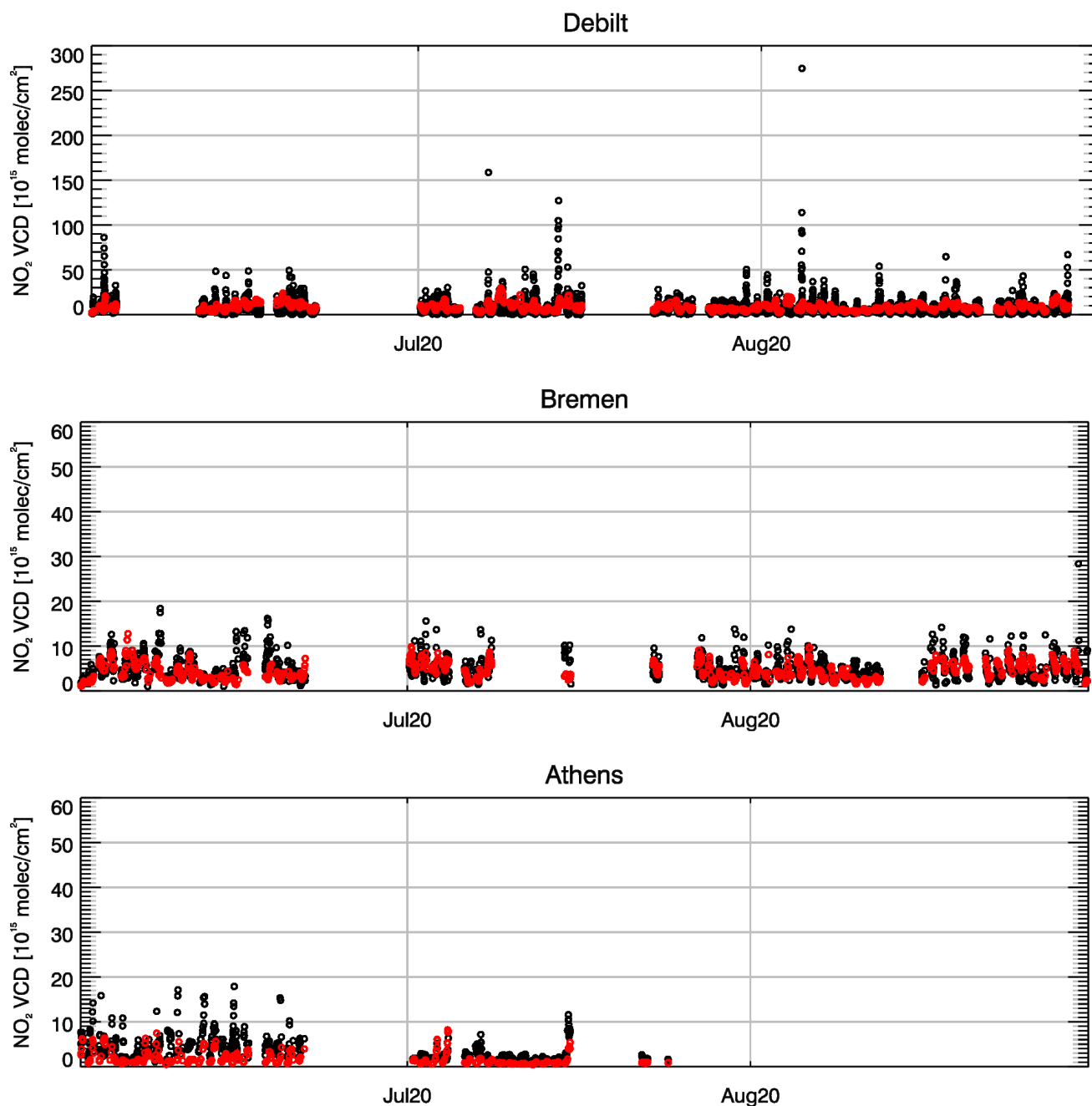
Figure 6.1. Time series of tropospheric NO_2 VCDs [10^{15} molec. cm^{-2}] from (black circles) MAX-DOAS and (red circles) the ENSEMBLE forecasts for (from top to bottom) De Bilt, Bremen, Athens.). Model results were calculated by multiplying NO_2 partial columns with averaging kernels for each observation layer followed by summing up resulting values in the vertical. Model data was interpolated to the averaging kernel layer altitudes prior to calculation of VCDs. Time period: June - August 2020. The MAX-DOAS instrument in Athens was not working correctly in August 2020 and measurements are therefore not available for this month.

Table 6.1. Overview of MAX-DOAS station data used for validation of regional air quality model simulations. The time period covered in this report is July 2019 to August 2020.

Station	Latitude, longitude	Altitude above sea level	Institution	Quantity	Character
Bremen (Germany)	53.106°N, 8.86°E	21 m	IUP-UB	column	urban
De Bilt (Netherlands)	52.1° N, 5.18° E	23 m	KNMI	column	urban
Athens (Greece)	38° N, 23.7° E	527 m	IUP-UB	column	urban

Figure 6.2 shows comparisons of diurnal cycles. Again, the mean column amounts of the ENSEMBLE are comparable for De Bilt and Uccle, but values are generally underestimated in Athens. Although larger differences are found depending on the regional model, the model ENSEMBLE performs much better than CAMS-global for all stations, CAMS-global is negatively biased. Some regional models show different variations from one hour to another. The afternoon peak observed by MAX-DOAS for De Bilt is not reproduced by all models. This may be related to inadequate photochemistry, scaling of emissions in time or vertical distribution of NO_2 and errors in simulating pollution transport towards the station.

Comparisons of weekly cycles are shown in Figure 6.3. Weekly cycles are underestimated by all regional models, with a stronger decrease of NO_2 columns from workdays towards the weekend retrieved by MAX-DOAS for all urban stations. CAMS-global mainly fails to reproduce weekly cycles, with no reduction in the NO_2 columns visible towards the weekend. Note that some variations of values from one day to another may just be coincidence due to data sampling. Large variability is found between the Monday averages by the individual regional models for Athens and De Bilt.
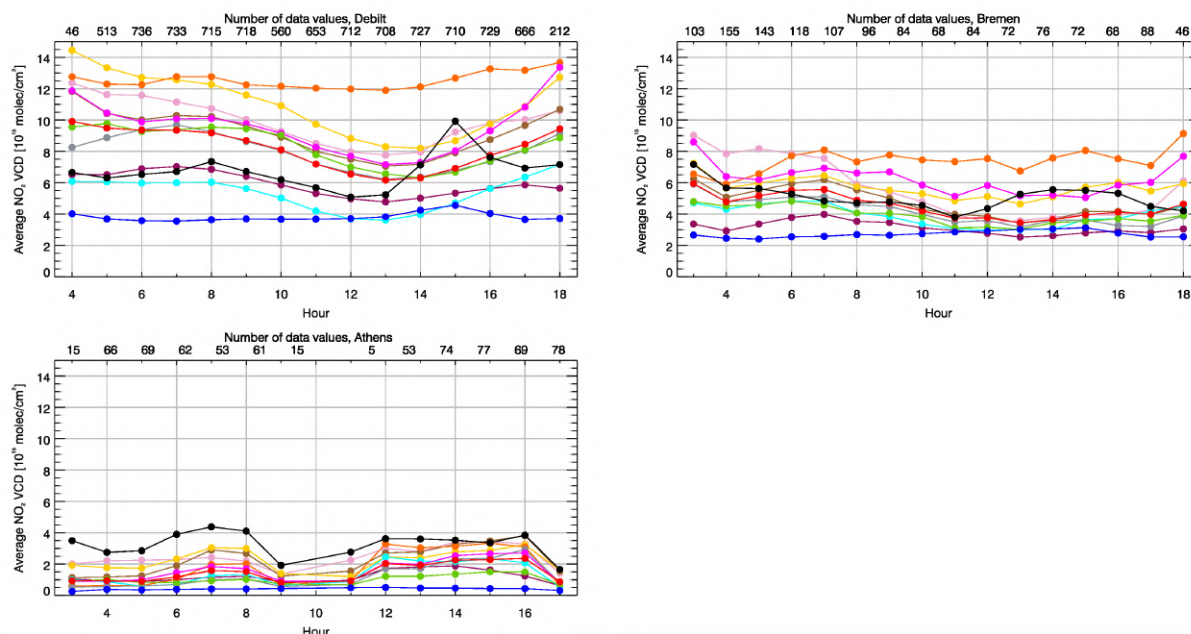
Figure 6.2. Diurnal cycles (averages over hourly bins) of tropospheric NO_2 VCDs [10^{15} molec. cm^{-2}] for (top left) De Bilt, (top right) Bremen and (lower left) Athens. The black lines show the MAX-DOAS retrievals. All other lines refer to model data: (red) ENSEMBLE, (yellow) CHIMERE, (brown) EMEP, (orange) SILAM, (purple-blue) LOTOS-EUROS, (cyan) EURAD-IM, (pink) MOCAGE, (grey) MATCH, (fuchsia) DEHM, (light green) GEM-AQ and (blue) CAMS-global. Time period: June – August 2020.

Scatter density plots or heat maps of tropospheric NO_2 VCDs from MAX-DOAS against model ENSEMBLE values corresponding to the time series displayed in Figure 6.1 as well as statistical values (root mean squared error, bias, correlation) are given in Figure 6.4. Corresponding statistical values for all individual models are given in Table 6.2. Correlations on the order of 15-55 % are found for each station for all models, with the ENSEMBLE reaching the highest correlation of about 51 % at Athens. Models tend to overestimate lower and underestimate higher NO_2 VCDs for the three urban stations. While the spread of values is quite large for individual data points, there is a fair agreement between models and retrievals for the majority of measurements for urban stations (as shown by the high percentage of values close to the reference line). The ENSEMBLE performs significantly better than CAMS-global in terms of correlation for all stations except for De Bilt. CAMS-global has a rather strong negative bias at all stations, while all regional models are negatively biased in Athens.
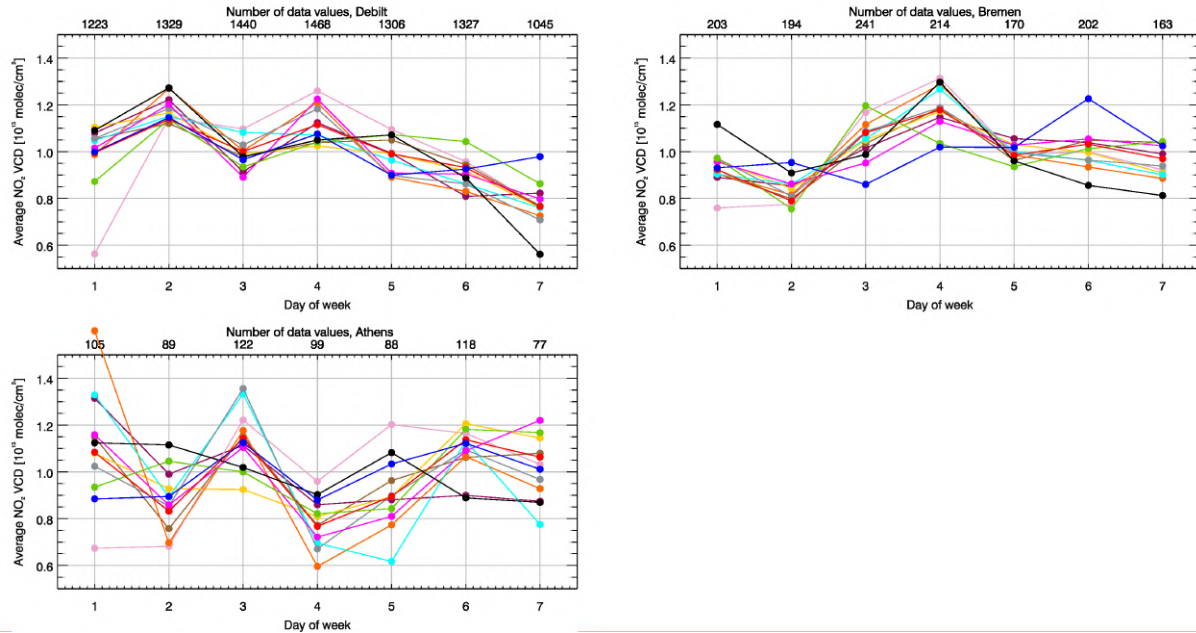
Figure 6.3. Weekly cycles (averages over daily bins divided by mean over whole week) of tropospheric NO$_2$ VCDs [10^{15} molec cm^{-2}] for (top left) De Bilt, (top right) Bremen and (lower left) Athens. The black lines show the MAX-DOAS retrievals. All other lines refer to model data: (red) ENSEMBLE, (yellow) CHIMERE, (brown) EMEP, (orange) SILAM, (purple) LOTOS-EUROS, (cyan) EURAD-IM, (pink) MOCAGE, (grey) MATCH, (fuchsia) DEHM, (light green) GEM-AQ and (blue) CAMS-global. Time period: June-August 2020.

	De Bilt	Bremen	Athens
ENS	6.932/1.248/0.269	2.640/-0.446/0.368	2.965/-1.848/0.511
CHIMERE	8.751/3.880/0.200	2.878/0.661/0.406	2.490/-1.011/0.559
EMEP	7.470/2.188/0.283	2.821/-0.184/0.329	2.859/-0.971/0.485
SILAM	12.144/5.741/0.209	5.109/2.357/0.331	3.739/-1.358/0.388
LOTOS-EUROS	6.988/-0.854/0.167	3.227/-1.813/0.272	3.374/-2.188/0.309
EURAD-IM	8.257/2.947/0.244	3.719/0.775/0.336	3.068/-0.816/0.322
MOCAGE	6.544/-1.546/0.284	3.047/-1.020/0.261	3.237/-1.944/0.421
MATCH	7.640/1.172/0.249	3.046/-0.780/0.294	3.343/-2.023/0.372
DEHM	8.499/2.431/0.174	3.603/1.148/0.297	2.908/-1.662/0.477
GEMAQ	7.857/1.428/0.146	3.007/-1.015/0.245	3.506/-2.383/0.315
OSUITE	6.806/-2.917/0.410	3.524/-2.317/0.066	4.089/-3.102/0.216

Table 6.2: Statistics on how tropospheric NO$_2$ VCDs [10^{15} molec. cm^{-2}] from models compare to MAX-DOAS retrievals at the three stations. Each column entry shows from left to right: root mean squared error [10^{15} molec. cm^{-2}], bias [10^{15} molec. cm^{-2}] and correlation coefficient (cor). Time period: June-August 2020.
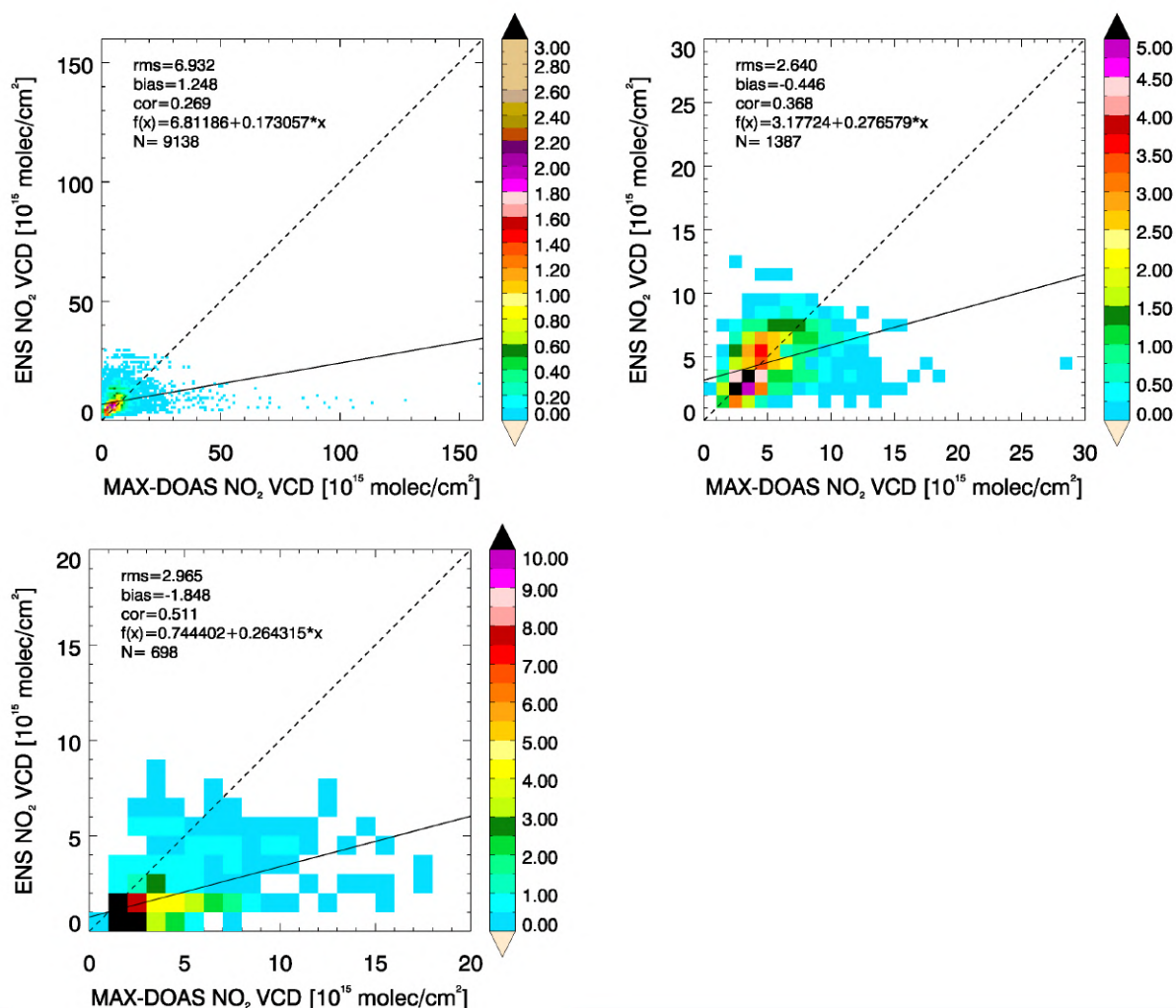
Figure 6.4. Scatter density plots of tropospheric NO_2 VCDs [1015 molec. cm−2] from MAX-DOAS against model ENSEMBLE hourly data for (top left) De Bilt, (top right) Bremen and (lower left) Athens. The data is shown with a bin size of 1 x 10^{15} molec. cm^{-2} and colour according to the number of data points per bin [%]. The dashed line is the reference line (f (x)=x). The solid line is the regression line (see top left of each plot for f(x) of this line). The root mean squared error (rms) [10^{15} molec. cm^{-2}], bias [10^{15} molec. cm^{-2}], pearson correlation coefficient (cor) as well as the number of data points N are given at the top left of each plot. Time period: June - August 2020.

7 Validation of tropospheric NO$_2$ columns against satellite retrievals

7.1 Summary

Regional air quality model columns of tropospheric NO$_2$, derived from the output provided on 8 levels with a top at 5km, are compared to 9:30 local time GOME-2/MetOp-A NO$_2$ satellite retrievals (IUP-UB v1.0 product). The overall spatial distribution of tropospheric NO$_2$ is reproduced by the ENSEMBLE, but values over central European emission hotspots are significantly underestimated by the majority of the models during winter, which results in a strong underestimation over these regions and of the seasonal cycle for the ENSEMBLE. There are stronger shipping signals compared to the satellite data. As a result of a major regional model upgrade in June (2019), which includes the use of an updated European emissions inventory with improved estimates for North African and Middle Eastern anthropogenic emissions, enhanced tropospheric columns of NO$_2$ are reproduced over these regions by all models. Differences between models and satellite observations may result from errors in anthropogenic emissions, photochemistry during winter months and from chemical processing inside ship plumes. In contrast to the analysis, the regional ENSEMBLE forecast shows a negative bias compared to the retrievals which is most pronounced for winter (~2-3 x 10^{15} molec/cm^2) but smaller during the rest of the year (overall ~0.5 x 10^{15} molec/cm^2). The negative bias is even larger for CAMS-global, which is in agreement with the stronger underestimation of values for European emission hotspots compared to regional models, demonstrating the benefit of running models at higher horizontal resolution. A systematic negative bias is however not present in the analysis for seasons other than winter. With respect to emission reductions during COVID-19 lockdown periods, regional ENSEMBLE analyses compare better to the changes observed by GOME-2 than regional ENSEMBLE forecasts and CAMS-global. However, other effects (changes in meteorological conditions, model versions) need to be considered.

7.2 Comparison with GOME-2 NO$_2$

In this section, regional air quality model columns of tropospheric NO$_2$ are compared to GOME-2/MetOp-A NO$_2$ satellite retrievals (IUP-UB v1.0) [Richter et al., 2011]. 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 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. Uncertainties in NO$_2$ satellite retrievals are large and depend on the region and season. Winter values in mid and high latitudes are usually associated with larger error margins. As a rough estimate, systematic uncertainties in regions with significant pollution are on the order of 20% – 30%. Conclusions may differ for comparisons to other satellite NO$_2$ products (e.g. TEMIS GOME-2, http://www.temis.nl shows lower retrieved NO$_2$ values for January). It should be noted here that model data is only available for altitudes up to 5000 m, meaning that (depending on tropopause height) tropospheric model columns may not be representative of the total amount of NO$_2$ in the troposphere. Note that since the CAMS-global upgrade of 26 June 2018, GOME-2 observations are assimilated by the global system. This is, however, a different retrieval product than what is used in the validation reported here (University of Bremen retrieval).
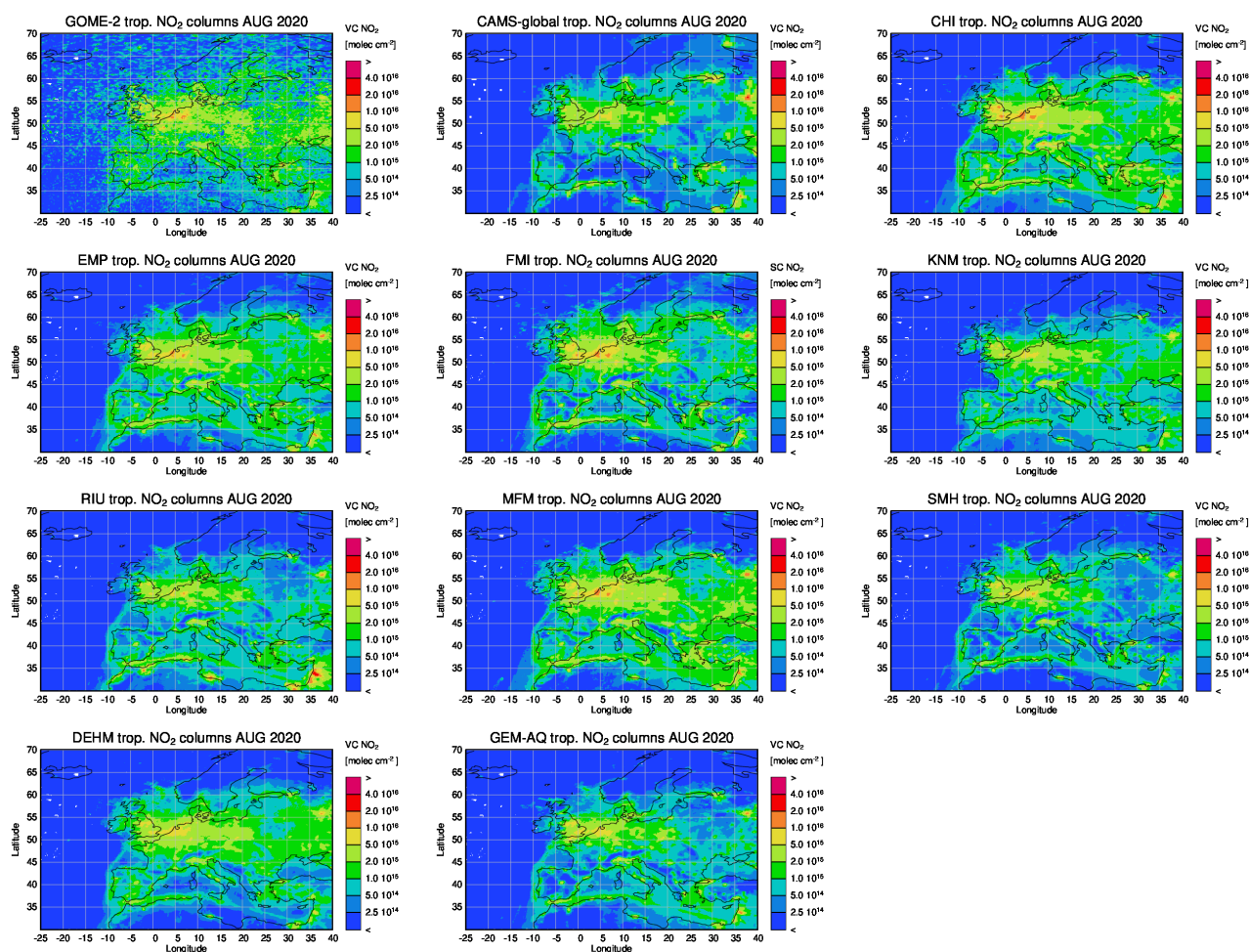
Figure 7.1. Maps of satellite-retrieved and model-simulated tropospheric NO$_2$ columns [molec cm^{-2}] for August 2020. From left to right: (first row) GOME-2, CAMS-global and CHIMERE; (second row) EMEP, SILAM and LOTOS-EUROS; (third row) EURAD-IM, MOCAGE and MATCH; (fourth row) DEHM and GEM-AQ. GOME-2 data were gridded to regional model resolution (i.e. 0.1° x 0.1°). Model data were treated with the same reference sector (25°W - 20°E) subtraction approach as the satellite data and linearly interpolated to the satellite overpass time (9:30 LT).

Figure 7.1 shows maps of monthly mean tropospheric NO$_2$ columns from GOME-2, regional models and CAMS-global for August 2020. The overall spatial distribution and magnitude of tropospheric NO$_2$ is reproduced by the regional models in principle. There are generally stronger shipping signals in the models compared to the satellite data, which may result from errors in anthropogenic emissions or from chemical processing inside the ship exhaust plumes (see e.g. Vinken et al., 2014).

Compared to CAMS-global, most regional models perform better for Central European emission hotspots, showing the benefit of higher horizontal resolution runs. As a result of a major regional model upgrade in June (2019), which includes the use of an updated European emissions inventory with improved estimates for North African and Middle Eastern anthropogenic emissions, enhanced tropospheric columns of NO$_2$ are now reproduced over these regions (e.g. Lebanon, Israel) by all models.
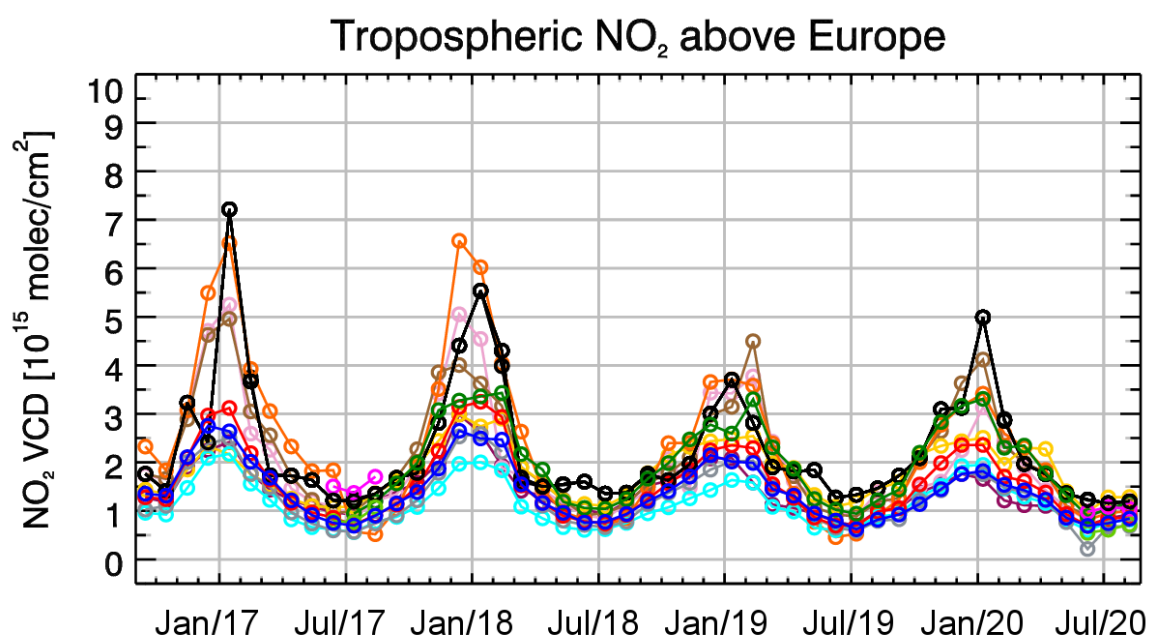
Figure 7.2. Time series of monthly averaged tropospheric NO_2 columns [10^{15} molec cm^{-2}] retrieved by (black) GOME-2 and simulated by (red) ENSEMBLE forecast, (green) ENSEMBLE analyses, (yellow) CHIMERE, (brown) EMEP, (orange) SILAM, (purple) LOTOS-EUROS, (cyan) EURAD-IM, (pink) MOCAGE, (grey) MATCH, (fuchsia) DEHM, (light green) GEM-AQ and (blue) CAMS-global. GOME-2 data were gridded to regional model resolution (i.e. 0.1° x 0.1°). Model data were treated with the same reference sector (25°W - 20°E) subtraction approach as the satellite data and linearly interpolated to the satellite overpass time (9:30 LT). Time period: September 2016 – August 2020.

Figure 7.2 shows time series of monthly mean tropospheric NO_2 columns for GOME-2 and the models. The seasonal variation is better reproduced by SILAM, MOCAGE and EMEP than by the other models. The latter clearly underestimate the seasonal cycle over Europe due to the strong underestimation of values in winter described above. The regional ENSEMBLE forecast shows a negative bias compared to the retrievals which is most pronounced during winter (~2-3 x 10^{15} molec/cm^2) but smaller during the rest of the year (overall ~0.5 x 10^{15} molec/cm^2). One of the reasons for this may be that the regional model output is limited to 5 km altitude. Compared to the ENSEMBLE forecast, the negative bias of CAMS-global is a bit larger, which is in agreement with the stronger underestimation over European emission hotspots for CAMS-global especially during winter. A systematic negative bias is not present in the regional model ENSEMBLE analysis for seasons other than winter. The decrease in retrieved wintertime maxima from 2017 to 2019, and increase in 2020 is not reproduced by the majority of regional models and CAMS-global.
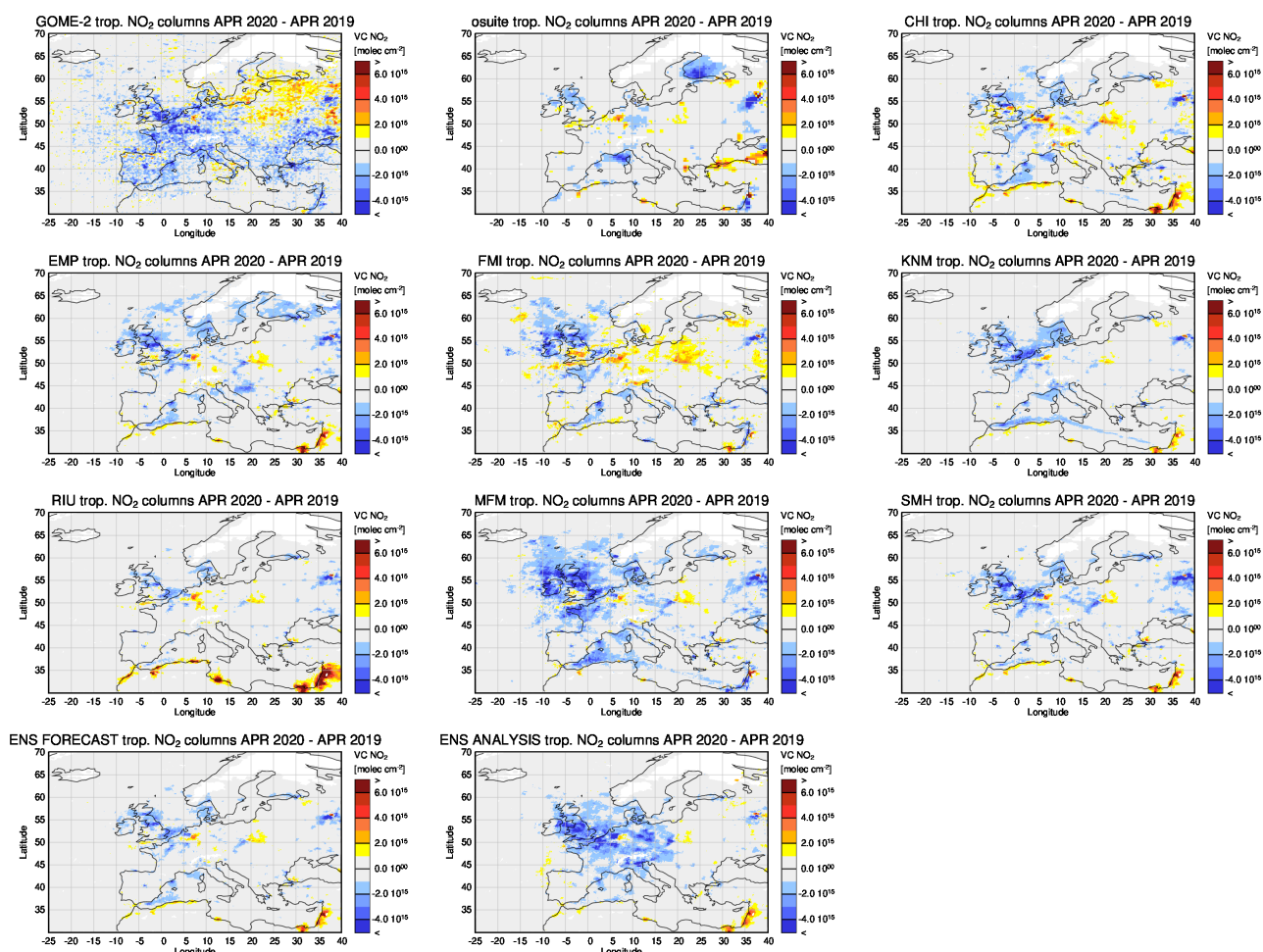
Figure 7.3. Maps of satellite-retrieved, and model-simulated tropospheric NO₂ columns [molecules cm⁻²] for April 2020 minus April 2019. From left to right: (first row) GOME-2, CAMS-global and CHIMERE; (second row) EMEP, SILAM and LOTOS-EUROS; (third row) EURAD-IM, MOCAGE and MATCH; (fourth row) regional ENSEMBLE forecasts and analyses. GOME-2 data were gridded to regional model resolution (i.e. 0.1° x 0.1°). Model data were modified with the same reference sector (25°W - 20°E) subtraction as the satellite data, and linearly interpolated to the satellite overpass time (9:30 LT).

With respect to emission reductions during COVID-19 lockdown periods (see Fig. 7.3), the forecasts from the individual regional models do not show significantly better results compared to GOME-2 than CAMS-global. On the other hand, the regional ENSEMBLE analyses compare considerably better to the changes observed by GOME-2 than regional ENSEMBLE forecasts and CAMS-global over many countries in Central Europe (e.g. Germany, France, the Benelux countries), likely due to more data being assimilated in the analyses. However, the interpretation of the results is complicated by the impact of possible changes in meteorological conditions and by changes of model versions (e.g. update of emission inventory in the regional model ENSEMBLE in June 2019, update of CAMS-global in July 2019).

8 Comparison with high-altitude EEA Air Quality e-reporting surface stations

8.1 Summary

European ozone EA Air Quality e-reporting measurements from high-altitude stations (above 1km) have been used to evaluate the regional models. Differences between the regional model orography and the true altitude of the station were used to select the model altitude level to compare with. The ensemble median mostly overestimates ozone levels during the period June 2020 - August 2020. More specific, depending on the station the observed ozone levels are reproduced to within 0% and 30% by the ensemble median D+0 forecast (1h-24h). Correlations observed were between 0.61 and 0.93 and the ensemble median D+0 forecast has a performance better than any of the individual nine models. MOCAGE appears to be deviating significantly from the ensemble median in terms of MNMBs and MOCAGE and EURAD are deviating significantly from the ensemble median in terms of correlation with observations. Validation metrics are also given for the ENSEMBLE analysis. The ENSEMBLE analysis has almost equivalent performance with the ENSEMBLE D+0 forecast in terms of MNMBs but performs better than ENSEMBLE D+0 forecast in terms of correlations (significantly higher correlations).

8.2 Introduction

The nine models and their ENSEMBLE median (D+0 forecast as well as the analysis) have been compared against Background-Rural EA Air Quality e-reporting measurements for surface stations at elevation greater than 1000 m above mean sea level (http://www.eea.europa.eu/data-and-maps/data/airbase-the-european-air-quality-database-7). Elevated stations were selected to fall within classes 1-2 in the O_3 Joly-Peuch (2012) classification for EA Air Quality e-reporting NRT stations. Table 8.1 shows the stations altitude above mean sea level together with the LOTOS-EUROS model altitude (i.e. from model's topography) pertaining to the nearest to the station grid point. Modelled gas mixing ratios were extracted at the model level, which is closest to the stations altitude as defined from the orography (see column 7 in Table 8.1).
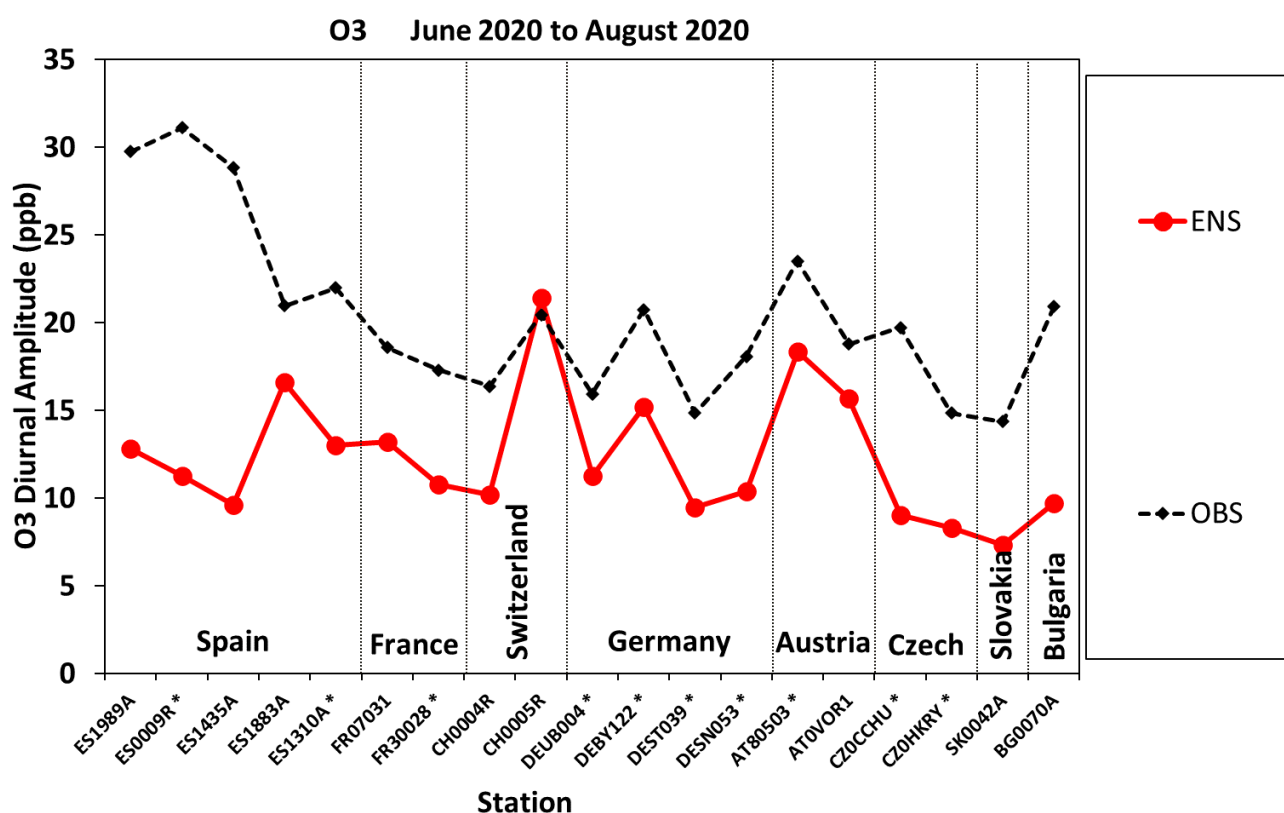
Figure 8.1. Amplitude of the diurnal cycle as measured (ppb) from stations observations (black diamonds) and as calculated from the ENSEMBLE forecast D+0 (red circles). With asterisks are denoted stations used in the assimilation process.

Table.8.1: Background-Rural EEA Air Quality e-reporting Stations (with NRT data) with Elevation higher than 1000 m.

Station Name	Stat_id	Longitude	Latitude	real Altitude	model Altitude	nearest Level	use in CAMS50
Lario	ES1989A	-5.09	43.04	1140	1199	0	validation
Capmisabalos	ES0009R	-3.14	41.27	1360	1124	2	assimilation
Vilafranca	ES1435A	-0.25	40.42	1125	907	2	validation
Torrelisa	ES1883A	0.18	42.46	1005	1282	0	validation
Ak- Pardines	ES1310A	2.21	42.31	1226	1117	1	assimilation
Chaumont	CH0004R	6.98	47.05	1136	727	3	-
Rageade	FR07031	3.28	45.11	1040	944	2	validation
Schlucht	FR30028	7.01	48.05	1200	520	3	assimilation
Schauinsland	DEUB004	7.91	47.91	1205	554	3	assimilation
Rigi-Seebodenalp	CH0005R	8.46	47.07	1031	997	1	validation
Sulzberg im Bregenzerwald	AT80503	9.93	47.53	1020	961	1	assimilation
Bad Hindelang/Oberjoch	DEBY122	10.40	47.52	1169	1150	0	assimilation
Brocken	DEST039	10.62	51.80	1130	302	4	assimilation
Fichtelberg	DESN053	12.95	50.43	1214	555	3	assimilation
Vorhegg bei Kötschach-Mauthen	AT0VOR1	12.97	46.68	1020	1427	0	validation
Churanov	CZ0CCHU	13.62	49.07	1118	739	4	assimilation
Krkonose-Rychory	CZ0HKRY	15.85	50.66	1001	530	4	assimilation
Bratislava - Jeséniova	SK0042A	20.99	48.78	1244	445	5	validation
Vitosha mountain	BG0070A	23.24	42.64	1321	863	4	-
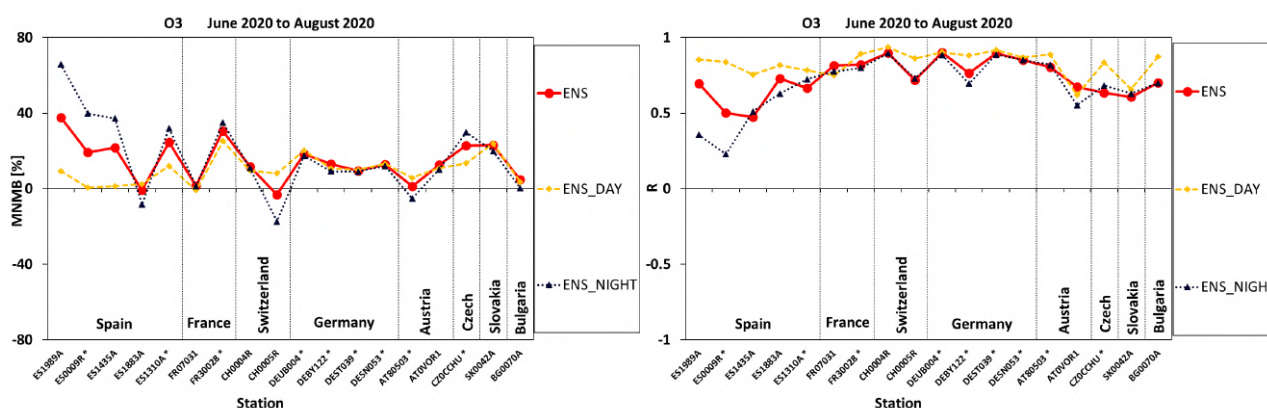
Figure 8.2. MNMBs [%] (left) and temporal correlations (right) calculated for the ENSEMBLE model for during daytime (orange) night-time (dark blue) as well as for the whole day (red) for the JJA 2020 period. With asterisks are denoted stations used in the assimilation process.

For the validation, hourly O_3 concentration values ($\mu g/m^3$) are extracted from the nine models as well as for the regional ENSEMBLE. It should be noted that, in the EEA Air Quality e-reporting network the O_3 measurements that were made by the instrument in ppb were converted from ppb to $\mu g/m^3$, following the EU directive 2008/50, i.e. by multiplying by 2. This conversion is approximately correct for low altitude stations. However, at high altitude stations pressure and temperature should be taken into account when converting from ppb to $\mu g/m^3$ and vice versa. As hourly pressure and temperature data were not available for all EEA Air Quality stations the comparison between observed and modelled ozone was done by re-converting both modelled and observed hourly O_3 concentration in ppb. For modelled ozone values the conversion was done by applying the following ideal gas equation with the model's estimates of temperature (T) and pressure (P) (from CAMS-global):

$$O3 \text{ (in ppb)} = O3 \text{ (in mg/m3)} * \left(\frac{R * T}{p_m * M_{o3}} \right)$$

8.3   Regional ENSEMBLE results

In the previous report it was shown that comparing the observed and modelled amplitude of the diurnal variation of ozone at each high-altitude station could provide a criterion concerning the exposure suitability of the stations. We found out in this report that an additional criterion is needed to differentiate stations as to their suitability in exposure. The additional criterion is the correlation coefficients between the amplitude of the diurnal cycle as observed and modelled to be statistically significant roughly higher than 0.3. Figure 8.1 shows the observed and modelled diurnal amplitude of ozone at each station, moving from Spain to Bulgaria. Figure 8.2 shows the MNMBs and the correlation coefficients calculated for the ENSEMBLE model during daytime, at night-time as well as for the whole day. We can see that the 2 criteria of diurnal amplitude and day and night MNMBs and correlation coefficients differentiate the 3 stations in Spain from all other stations. For the above-mentioned findings these 3 high altitude stations will be excluded from our analysis. It should also be noted that daytime correlations have very small variations from stations to station and are higher than those calculated for night-time as well as for the 24-h daily means.
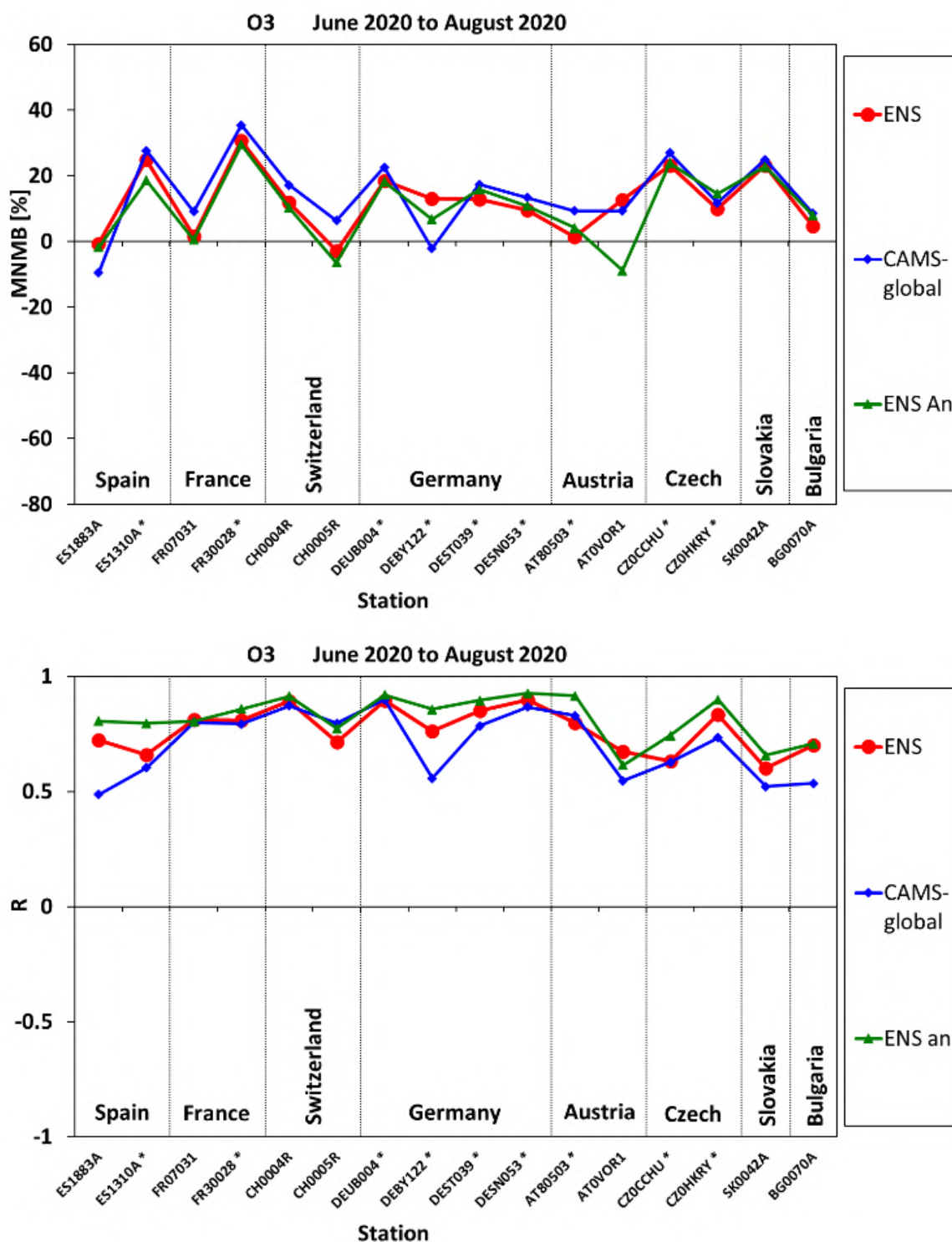
Figure 8.3. O_3 MNMBs [%] (top) and correlation coefficient (bottom) for the regional ENSEMBLE (forecast D+0; red circles and analysis; green triangles) as well as for CAMS-global (forecast D+0; blue diamonds) for the period JJA2020. With asterisks are denoted stations used in the assimilation process.
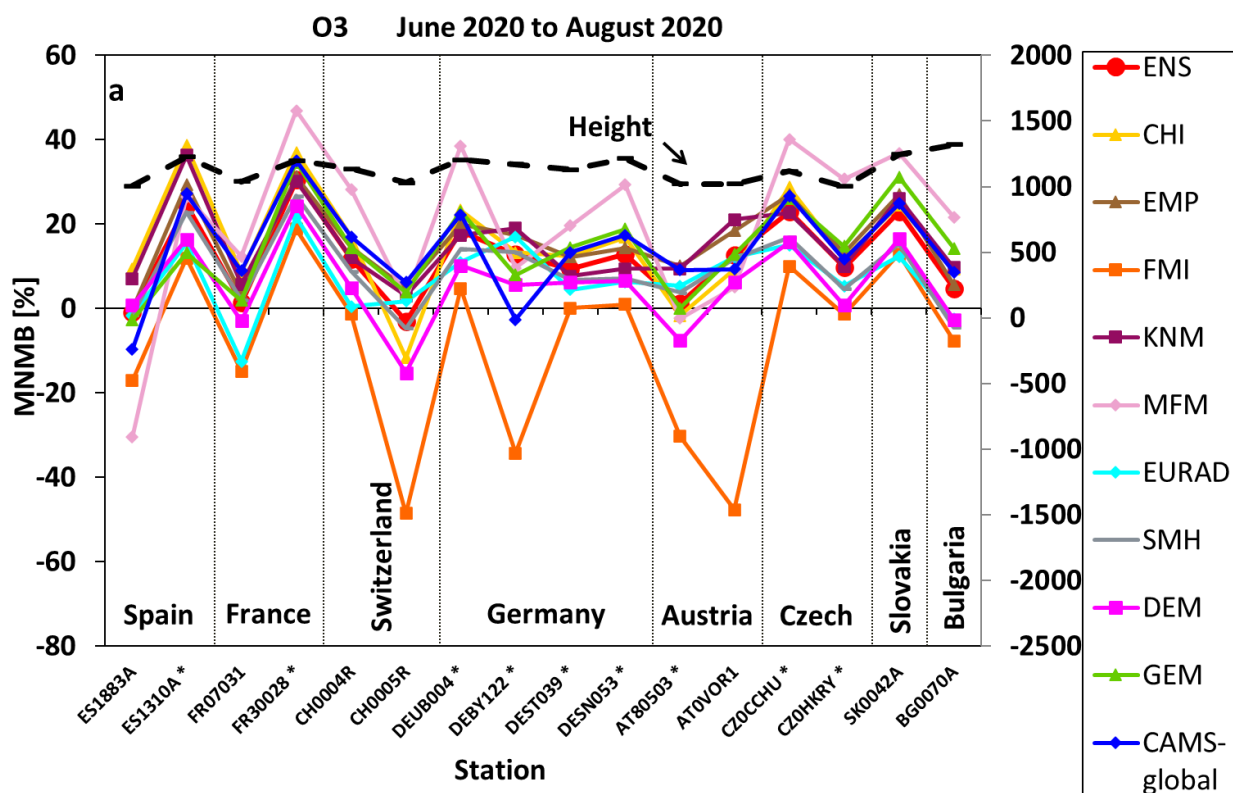
Fig. 8.4. Ozone MNMBs [%] for each one of the 9 models and CAMS-global (June to August 2020) for stations above 1000m altitude. With asterisks are denoted stations used in the assimilation process.

Figure 8.3 shows the Modified Normalized Mean biases (top) and correlation coefficients (bottom) at each of the remaining stations, moving from Spain to Bulgaria (i.e. from West to East) pertaining to the median of the regional ENSEMBLE forecast (D+0) and analysis (D+0) as well as CAMS-global (D+0). The ensemble median overestimates ozone levels during the period June 2020 to August 2020. Depending on the station the range of MNMB for the ENSEMBLE median D+0 forecast was found to be between 5% and 25%. From Figure 8.3 (bottom panel) it is obvious that the regional ENSEMBLE reproduces well the ozone variability. As it appears from Figs 8.3 (bottom panel) the correlation coefficients are highly significant (0.65<r<0.85). It should be noted that for the period June 2020 to August 2020 the ENSEMBLE analysis performs better than the ENSEMBLE D+0 forecast in terms of correlations (significantly higher correlations) and also that CAMS ENSEMBLE D+0 forecast shows higher correlations than CAMS Global D+0 forecast.
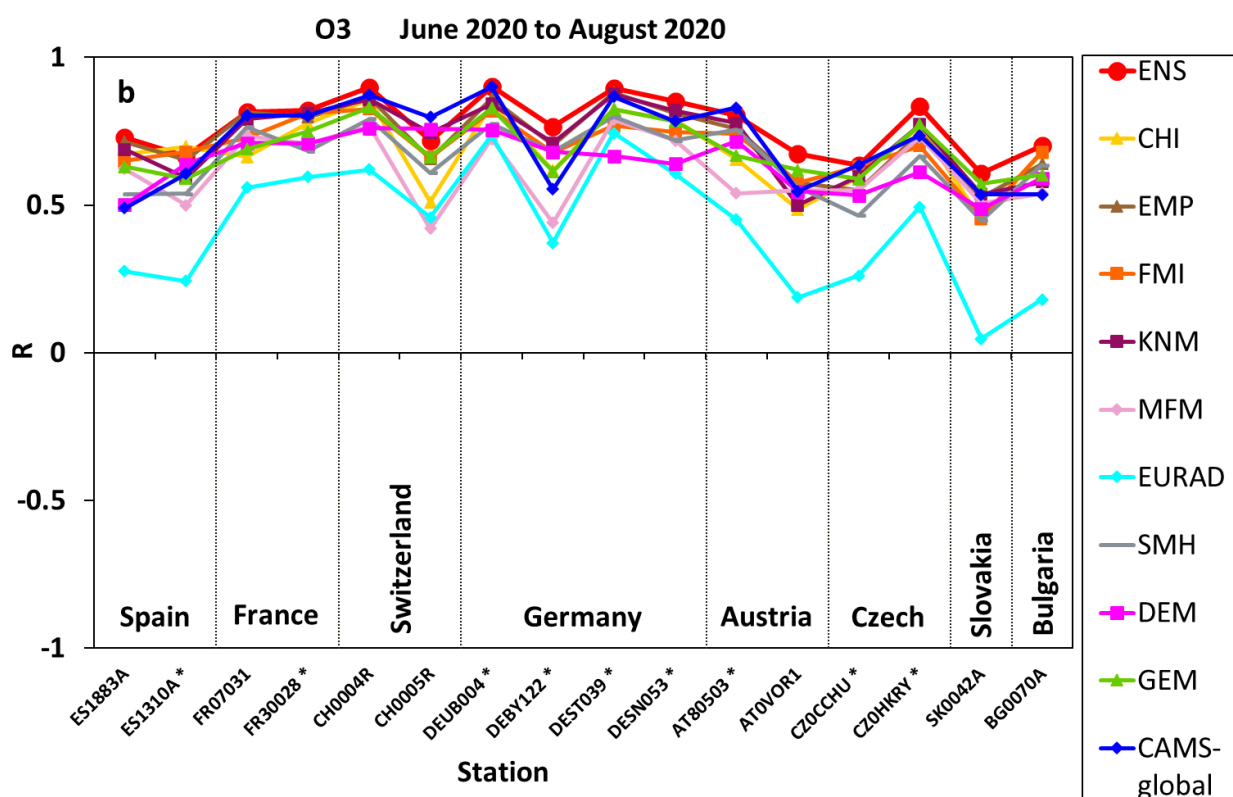
Fig. 8.5: Correlation coefficients (from hourly values) between observed and modelled O$_3$ from all models of the ENSEMBLE and CAMS-global (o-suite) for June to August 2020. With asterisks are denoted stations used in the assimilation process.

8.4   Results for the nine regional models

Figure 8.4 shows Modified Normalized Mean biases at each station with elevation greater than 1000 m above mean sea level, moving from Spain to Bulgaria (from West to East) pertaining to each one of the 9 model calculations, the regional ENSEMBLE as well as CAMS-global. All results are based on the forecast D+0 elevated ozone values. On top of the graph shown is the elevation of the station. Depending on the station the observed ozone levels are reproduced to within 5% and 25% from the regional ENSEMBLE. Figure 8.4 shows that MOCAGE deviate significantly from the ensemble median (strong positive offset up to 45% and 40% respectively) while the remaining models show scores closer to the ENSEMBLE. However, SILAM deviates significantly from all ensemble members at 4 individual stations (namely CH0005R, AT0VOR1, AT80503, DEBY122) in Switzerland, Austria, and Germany respectively, with strong negative biases down to -45%.

Finally, Figure 8.5 shows the correlations between observations and each model. It is clear that the regional ENSEMBLE reproduces well the ozone variability and has a better score than any of the individual models. The EURAD and the MOCAGE models show lower correlation while the remaining models show scores closer to the ENSEMBLE.

9 Comparison with ozone sonde observations

9.1 Summary

Free tropospheric ozone (<850 hPa) could be reproduced by the ENSEMBLE forecast and analysis with MNMBs between 3% and 15% during JJA 2020. The other models show MNMBs between -7% and 15% (forecasts and analysis).

9.2 Comparison approach

For the validation, the sonde profiles are compared to the model data closest in time. The model data is provided at the geographical coordinates of the sonde stations, the horizontal drift during the ascend of the sonde is considered negligible.

The model concentrations at the different height levels (0, 50, 250, 500, 1000, 2000, and 5000m above the ground) are matched to the respective sonde observations and are converted to mass mixing ratios. Pressure and temperature values needed for the conversion are taken from the sonde observations. For each station and all individual launches, the differences between observation and model are calculated. In order to be able to compare the profiles of different stations, this is done for fixed altitude levels between 0 and 6000m (interval for the surface 50m, above 100m, interval 100m). The sonde and model values are then aggregated to monthly means for each station and altitude level. For each month mean modified normalized biases (MNMB) are then calculated over all European stations for the free troposphere (<850 hPa).

9.3 Results for the ENSEMBLE

For the period August 2019 to August 2020, the ENSEMBLE forecast shows MNMBs between 3% and 15%. The ENSEMBLE analysis shows a similar behaviour with MNMBs between 3% and 14%, see Fig. 9.1.

Table 9.1: Sonde stations used in the validation for JJA2020

Station/location	Lat	lon	alt [m]
De Bilt	52.1	5.2	4
Hohenpeissenberg	47.8	11.2	976
Jokioinen	60.8	23.5	103
Legionow	52.4	20.97	96
Lerwik	60.14	-1.19	84
Madrid Spain	40.5	-3.8	631
Prag	50	14.4	302
Sodankyla	67	27	180
Uccle Belgium	51	4	100
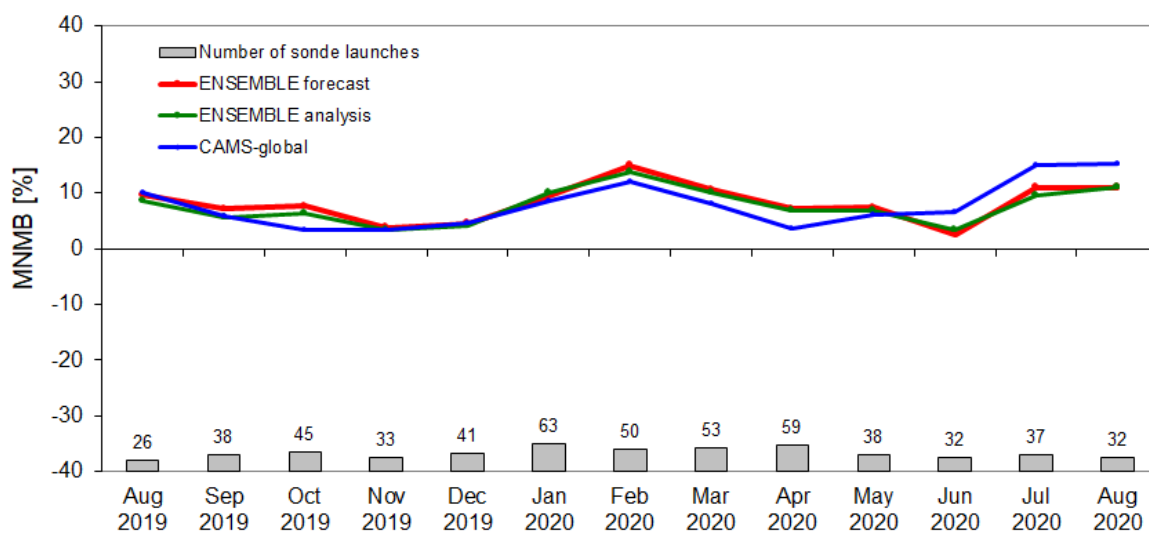
Figure 9.1 - MNMBs for the regional ENSEMBLE and CAMS-global between August 2019 and August 2020 for the free troposphere region (<850 hPa).
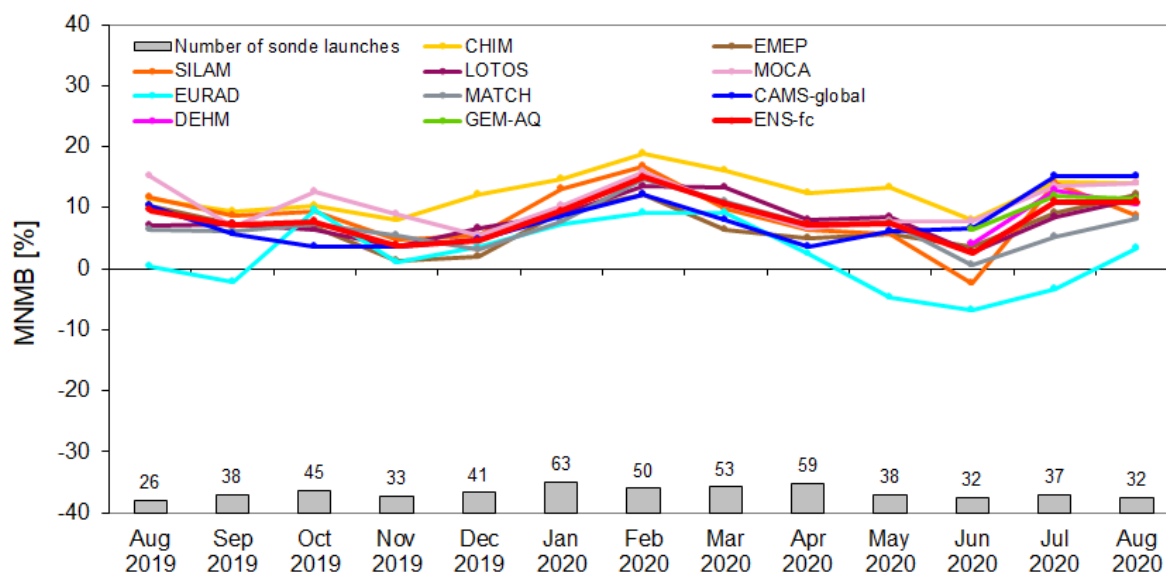
Figure 9.2. MNMBs for the individual regional model forecasts between August 2019 and August 2020 for the free troposphere region (pressure < 850 hPa).
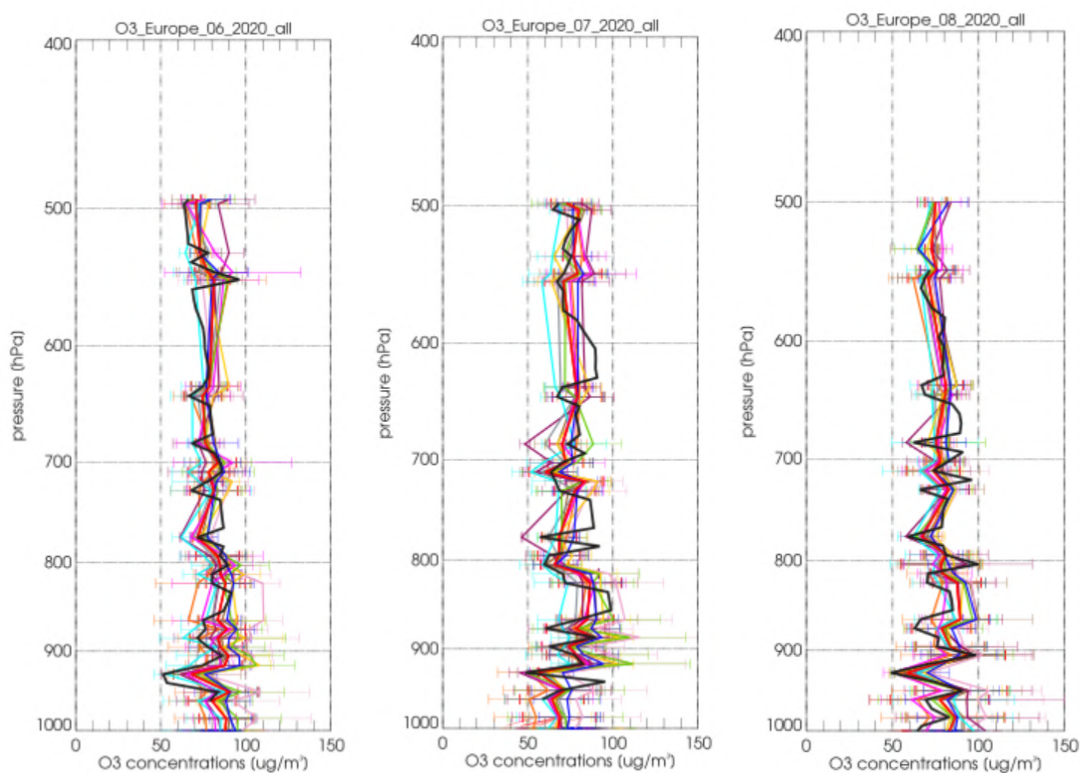
Figure 9.3 Averaged monthly profiles from the model forecast and sonde comparisons over European sonde stations for June, July, and August 2020 (red: ENSEMBLE, blue: CAMS-global, yellow: CHIMERE, brown: EMEP, orange: SILAM, purple: LOTOS-EUROS, cyan: EURAD-IM, pink: MOCAGE, grey: MATCH, fuchsia: DEHM, light green: GEM-AQ).
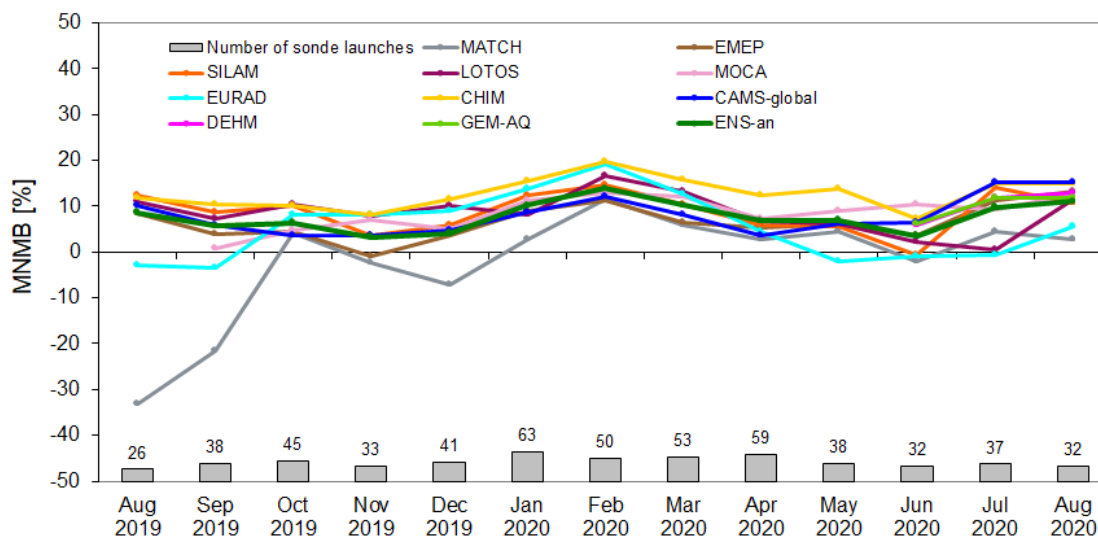
Figure 9.4. MNMBs for the regional model analyses between August 2019 and August 2020 for the free troposphere region (pressure < 850 hPa).

9.4 Results for individual regional models

Between August 2019 and August 2020 regional model forecasts show MNMBs in the range of -7 % and 19%, see Fig. 9.2.

Results for the regional model analyses

Similar to the results of the individual models' forecasts, the analyses show MNMBs between -4% and 20% (Fig. 9.4). Only the MATCH model showed larger negative MNMBs in August and September 2019 up to -33%, which is in contrast to the forecast. This issue is attributed to not properly processed background error statistics. Correction was implemented in the first week of January 2020.
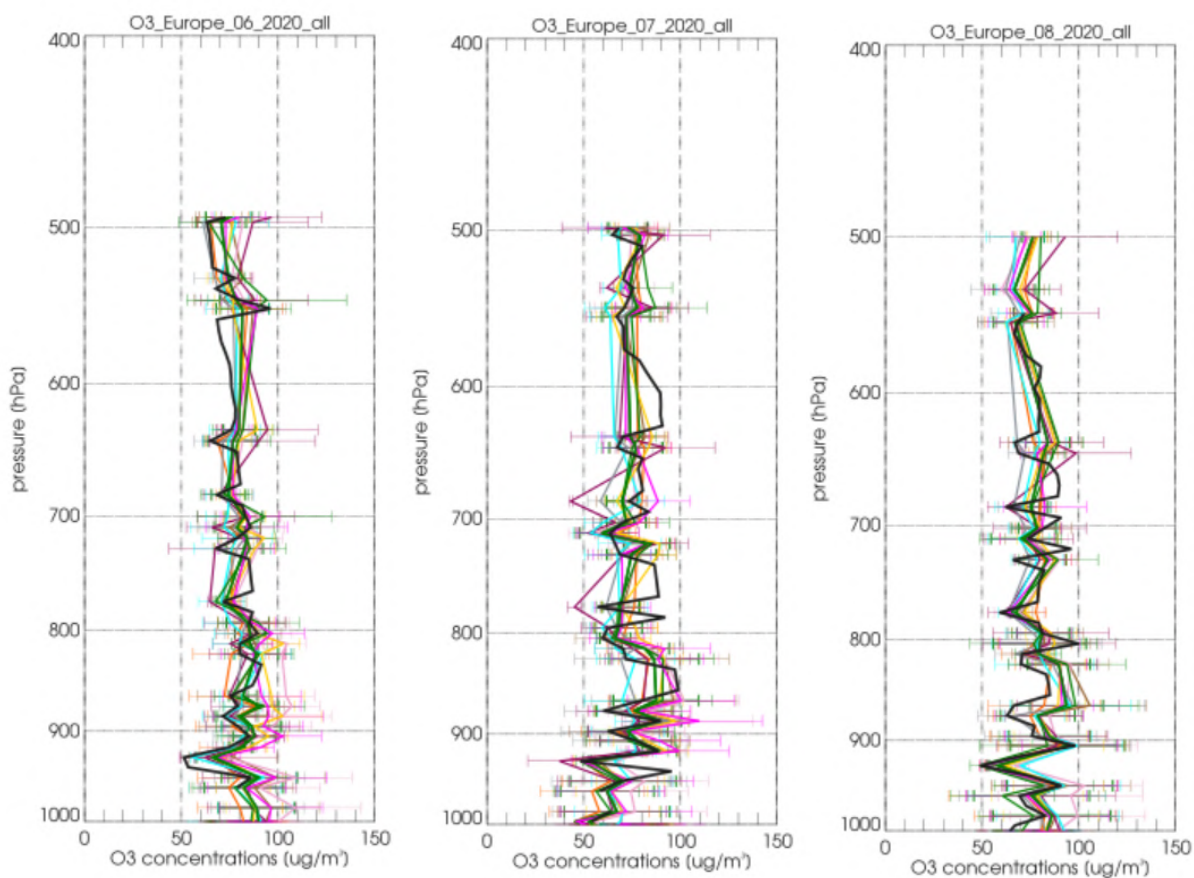
Figure 9.5. Averaged monthly profiles from the model analysis and sonde comparisons over European sonde stations for June, July, and August 2020 (green: ENSEMBLE, blue: CAMS-global, yellow: CHIMERE, brown: EMEP, orange: SILAM, purple: LOTOS-EUROS, cyan: EURAD-IM, pink: MOCAGE, grey: MATCH, fuchsia: DEHM, light green: GEM-AQ).

10 Comparison with GAW stations

10.1 Summary

The model concentrations of O_3 and CO at higher model levels were compared with five GAW high-altitude stations in mountainous terrain. As for the EEA air quality e-reporting stations, differences between the regional model orography and the true altitude of the station were used for this model level selection. Good results were obtained for the ENSEMBLE for ozone with small biases and good correlations. The CHIMERE model shows slightly larger MNMBs in both the analysis and the forecast. For CO, especially the ENSEMBLE corresponds well to the observations, however, small underestimates are found. The MATCH and SILAM models behave differently in the analysis than in the forecast, showing larger overestimations for CO than the rest of the models. The time series and correlation coefficients for CO and O_3 show that the ENSEMBLE reproduces for a large part the variability observed.

10.2 Comparison method

Hourly O_3 and CO concentration values in µg/m^3 are extracted from the nine models and are compared to the GAW measurements, which were converted from volume mixing ratios (ppb) into concentrations by using pressure and temperature values at the respective pressure levels from the IFS model.

The altitude of the stations Hohenpeissenberg (HPB), Jungfraujoch (JFJ), Monte Cimone (CMN), Sonnblick (SNB) and Zugspitze (ZUG) in the model has been extracted from the orography as used in the LOTOS-EUROS model, see Table 10.1. For the level choice, the GAW stations' altitudes together with the best correlation of the corresponding levels were taken into account. Uncertainties due to the choice of level (calculated as mean differences between the chosen level and one up/down for the period JJA 2020 for the ENSEMBLE) are up to ±11 µg/m^3 for CO and up to ±9 µg/m^3 for O_3.

Table 10.1 - Validation set-up for June – August 2020.

| station	altitude station [m]	altitude model [m]	level choice (range 0-7)	altitude at level [m]
HPB	985	813	2	1063
JFJ	3580	1837	5	3837
CMN	2165	602	4	1602
SNB	3105	1687	3	2187
ZUG	2670	1348	4	2348
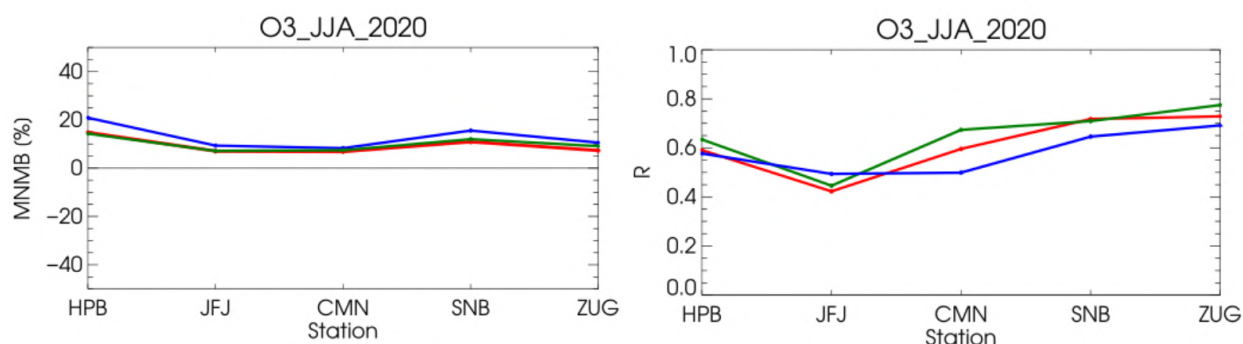
Figure 10.1. MNMBs [%] (left) and correlation coefficients (right) for ozone (red: ENSEMBLE forecast, green: ENSEMBLE analysis, blue: CAMS-global) for the period June to August 2020.
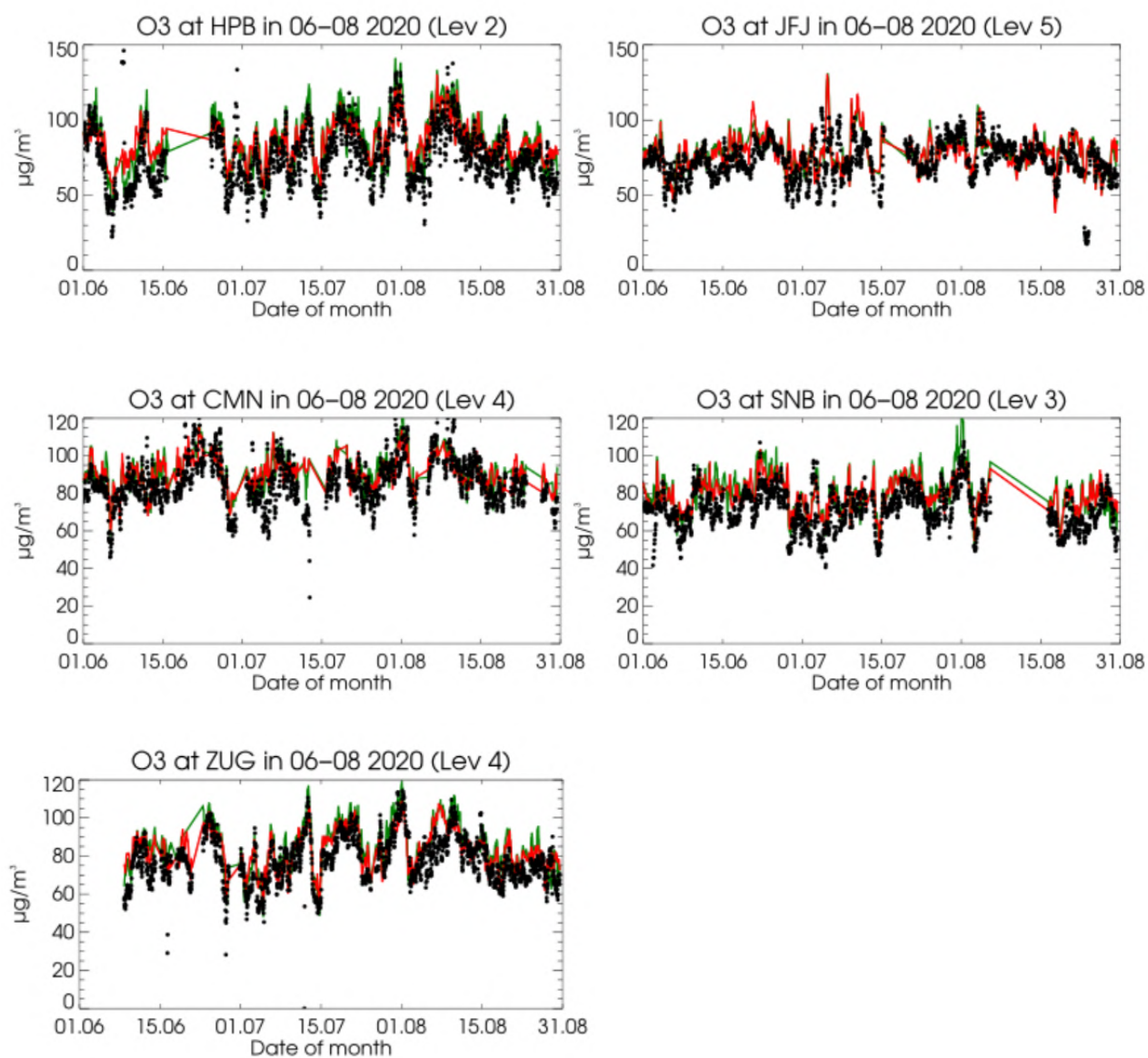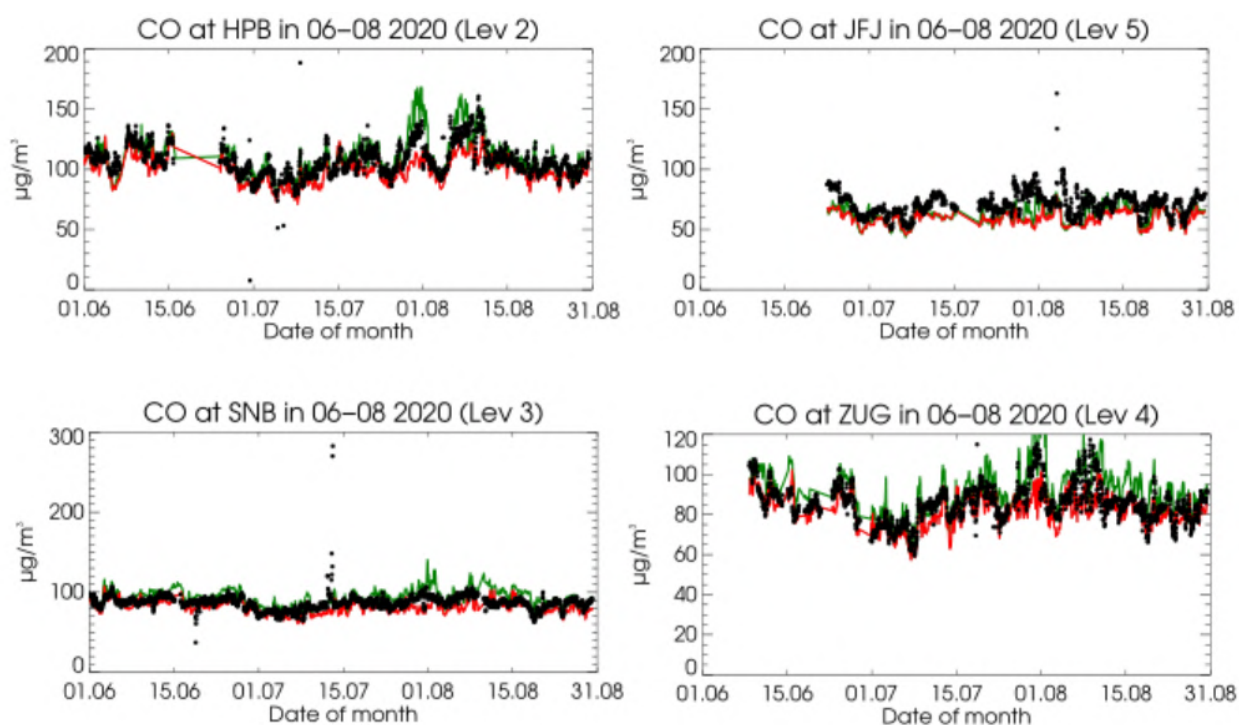

Figure 10.2. Time series plots for the ENSEMBLE forecast (red) and ENSEMBLE analysis (green) for surface O_3 in comparison with high altitude stations for the period June to August 2020.
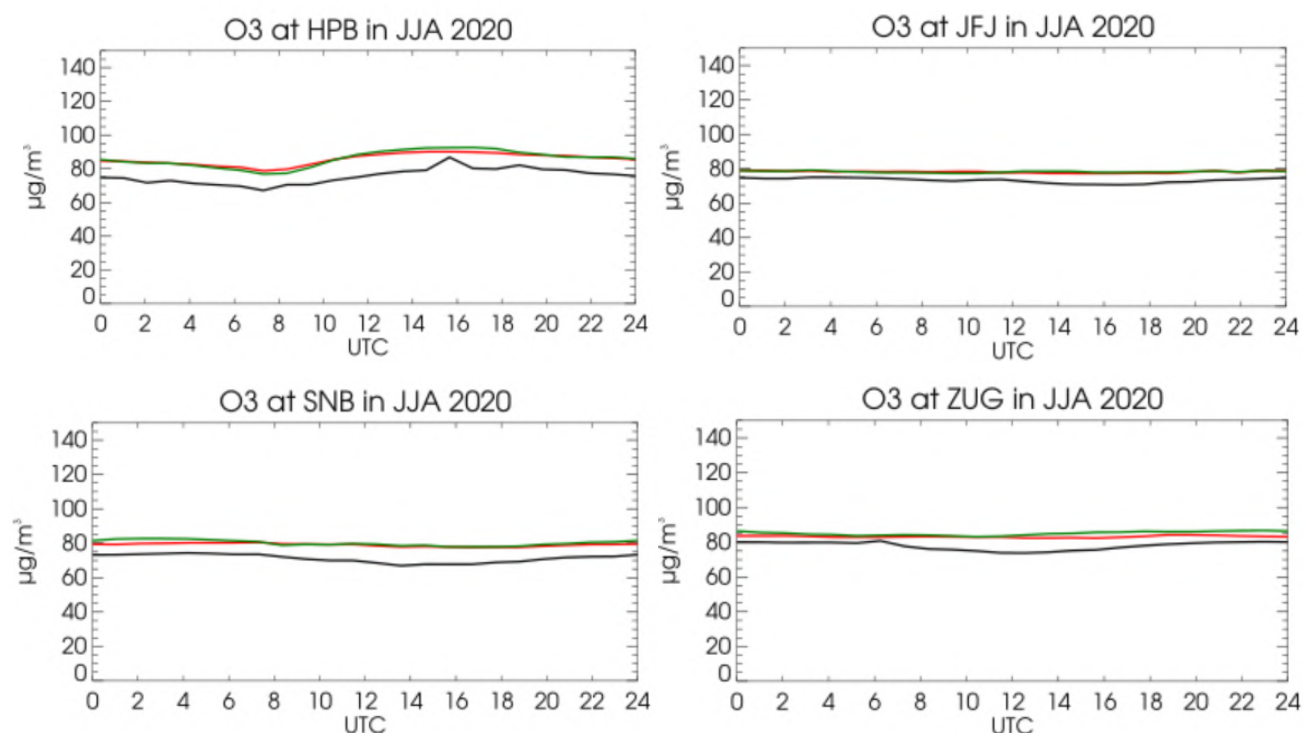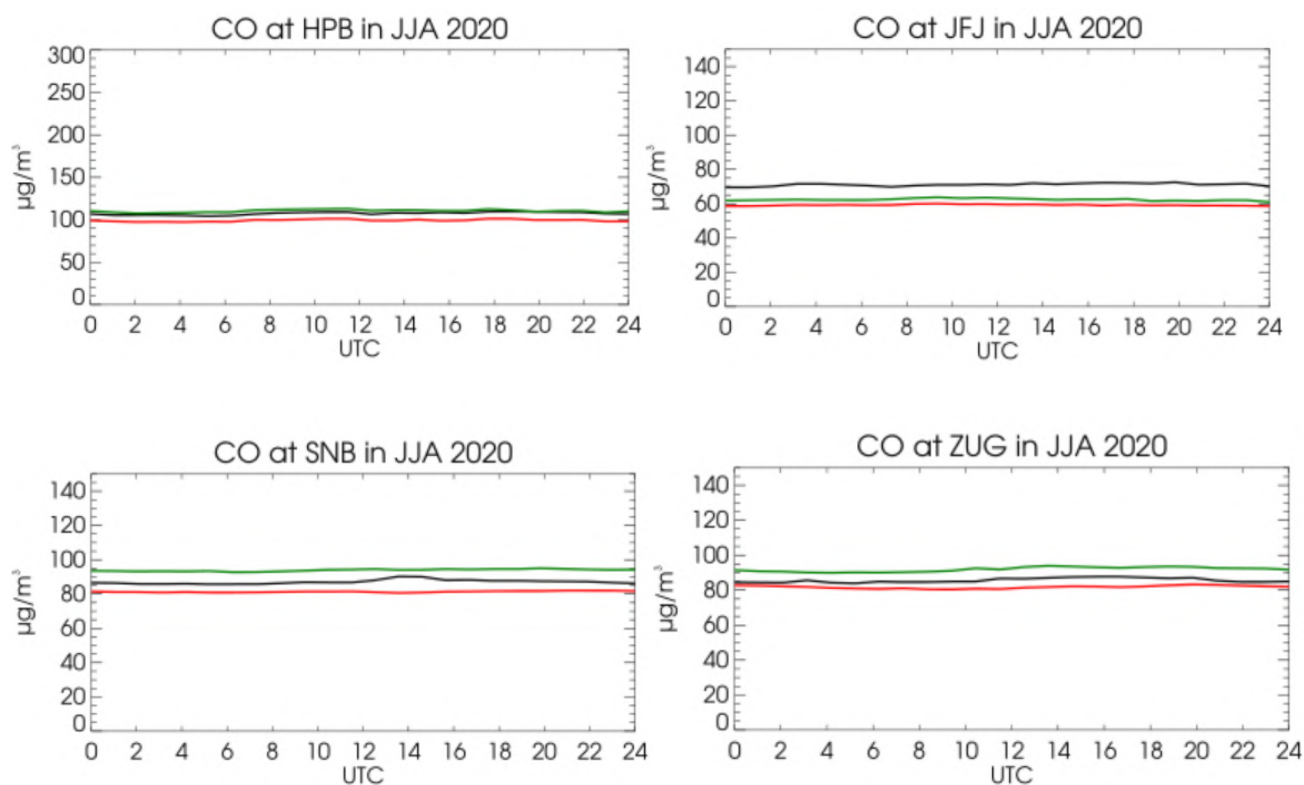
Figure 10.3. Mean diurnal cycle of O_3 for the ENSEMBLE forecast (red) and ENSEMBLE analysis (green) compared to the observations (black) for the period JJA2020.
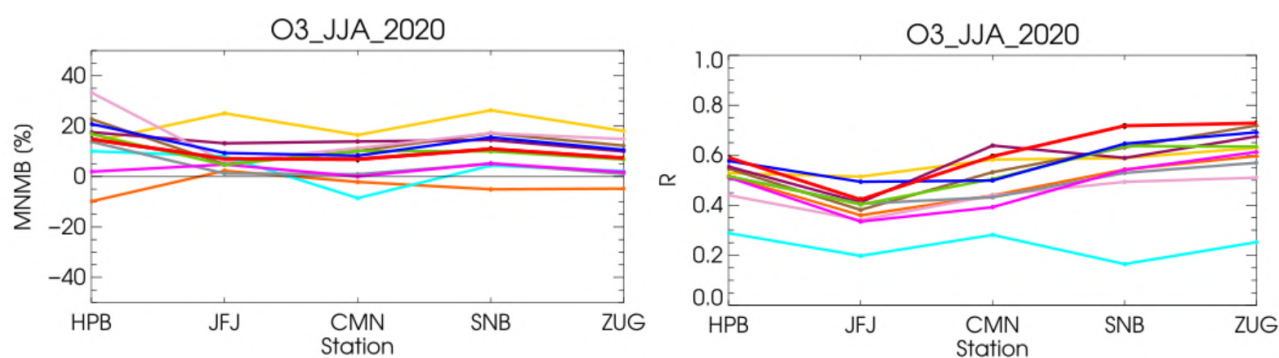
Figure 10.4. MNMBs [%] (left) and correlation coefficients (right) for all model forecasts for ozone (red: ENSEMBLE, blue: CAMS-global, yellow: CHIMERE, brown: EMEP, orange: SILAM, purple: LOTOS-EUROS, cyan: EURAD-IM, pink: MOCAGE, grey: MATCH). Altitudes are listed in Table 10.1.

10.3 Ozone

The ENSEMBLE forecast shows positive MNMBs between 6% and 14% and correlation coefficients ranging between 0.42 and 0.72 for the period June 2020 to August 2020. The ENSEMBLE analysis shows almost the same MNMBs (between 7% and 14%) and slightly better correlation coefficients ranging from 0.44 and 0.77 (see Fig. 10.1).
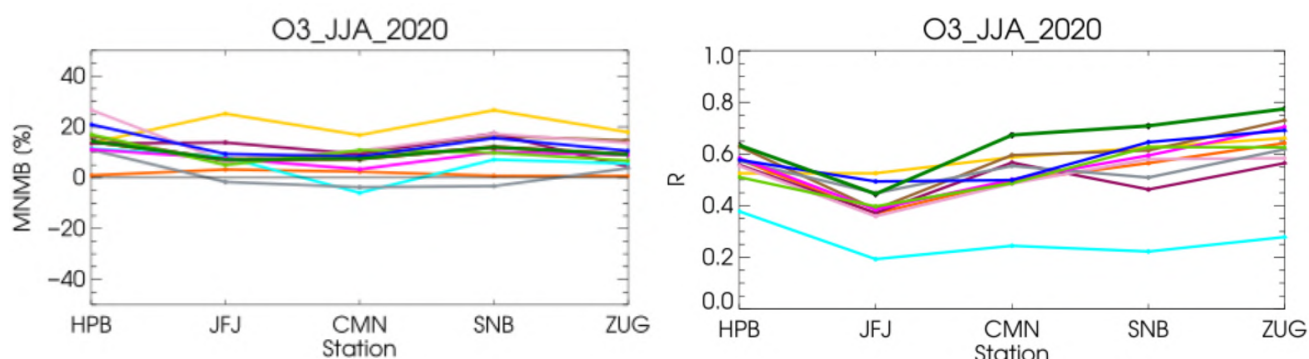
Figure 10.5. MNMBs [%] (left) and correlation coefficients (right) for all models analyses for ozone (green: ENSEMBLE, blue: CAMS-global, yellow: CHIMERE, brown: EMEP, orange: SILAM, purple: LOTOS-EUROS, cyan: EURAD-IM, pink: MOCAGE, grey: MATCH, fuchsia: DEHM, light green: GEM-AQ). Altitudes are listed in Table 10.1.
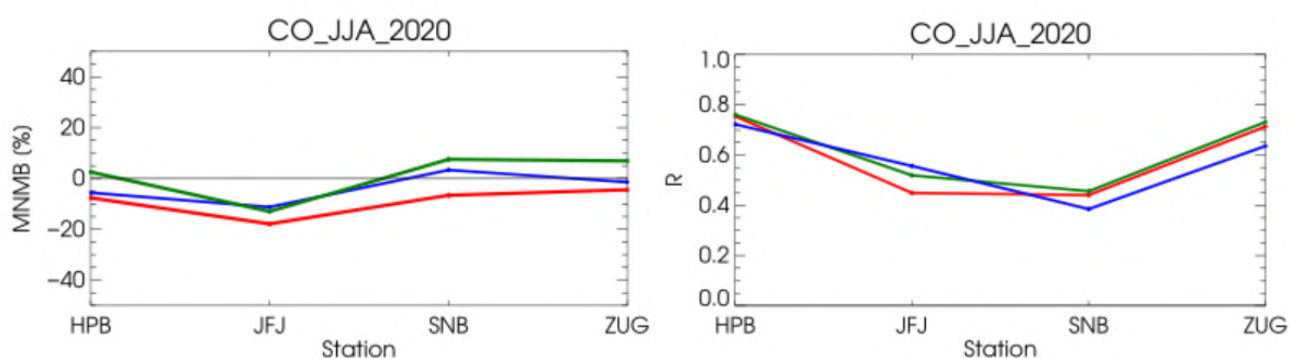
Figure 10.6. MNMBs [%] (left) and correlation coefficients (right) for the ENSEMBLE for CO (red: ENSEMBLE forecast, green: ENSEMBLE analysis, blue: CAMS-global).

The time series plots show a good correspondence between model and observations (Fig. 10.2). Diurnal cycles (Fig. 10.3) are not very pronounced at most GAW station locations. For HPB, an increase of ozone after sunrise is captured by the analysis and forecast.

Results for individual model forecasts:

The models show MNMBs in the range of 33% and -10% for the period June to August 2020 (Fig. 10.4). The CHIMERE model shows larger positive MNMBs than the other models. During JJA 2020, correlation coefficients vary between 0.4 and 0.72. An exception is the EURAD model with low correlation coefficients (below 0.3).

Results for the individual model analyses:

For the individual model analyses (Fig. 10.5), MNMBs range between 30% and -5% and are very similar to the forecast. Same as for the forecast, correlation coefficients range between 0.4 and 0.8, except for the EURAD model, which shows lower correlation.

Figure 10.7. Time series plots for the ENSEMBLE forecast (red) and ENSEMBLE analysis (green) for surface CO in comparison with high altitude stations.

Figure 10.8. Mean diurnal cycle of CO for the ENSEMBLE forecast (red) and ENSEMBLE analysis (green) compared to the observations (black) for the period JJA2020.
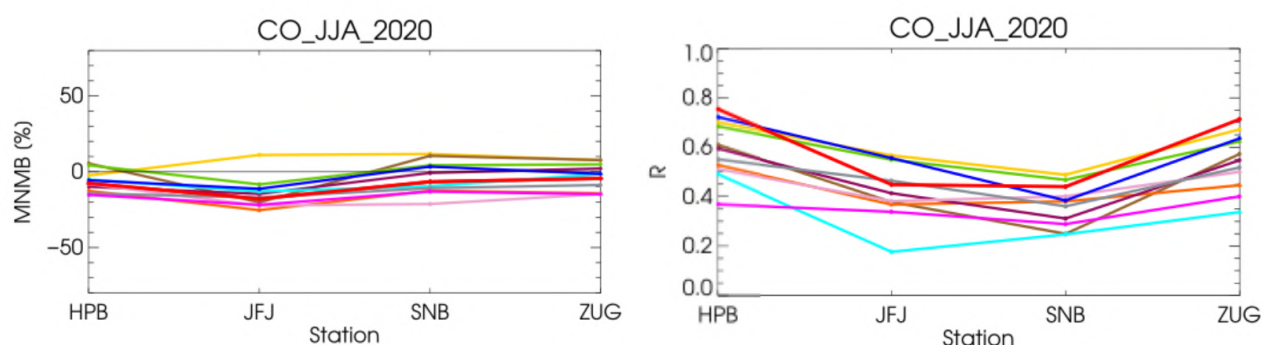
Figure 10.9. CO MNMBs [%] (left) and correlation coefficients (right) for all regional model forecasts, the ENSEMBLE and CAMS-global (red: ENSEMBLE forecast, blue: CAMS-global, yellow: CHIMERE, brown: EMEP, orange: SILAM, purple: LOTOS-EUROS, cyan: EURAD-IM, pink: MOCAGE, grey: MATCH, fuchsia: DEHM, light green: GEM-AQ). Altitude ranges are listed in Table 10.1.
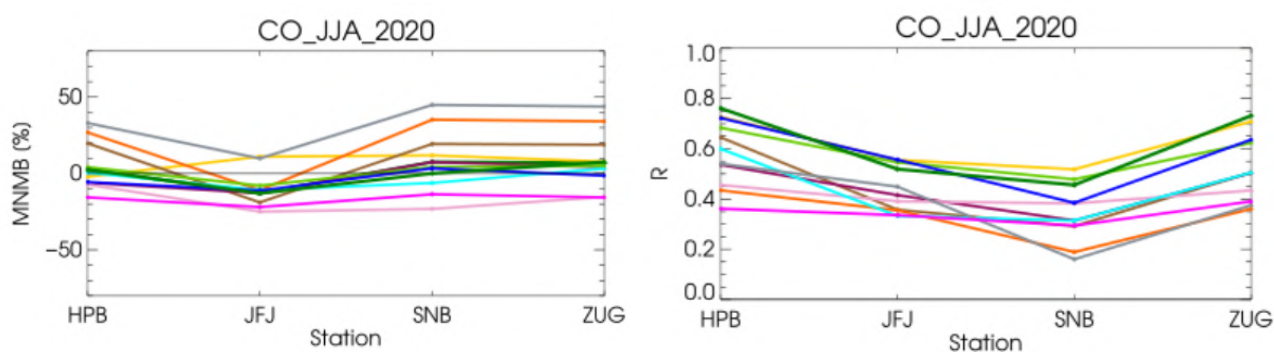
Figure 10.10. CO MNMBs [%] (left) and correlation coefficients (right) for all regional model analyses, the ENSEMBLE and CAMS-global (green: ENSEMBLE analysis, blue: CAMS-global, yellow: CHIMERE, brown: EMEP, orange: SILAM, purple: LOTOS-EUROS, cyan: EURAD-IM, pink: MOCAGE, grey: MATCH, fuchsia: DEHM, light green: GEM-AQ). Altitude ranges are listed in Table 10.1.

10.4 Carbon monoxide

For CO, the ENSEMBLE forecast and analysis MNMBs range between -12% and 2% for the analysis and between -17% and -4% for the forecast during June 2020 - August 2020. Correlation coefficients are between 0.45 and 0.76 for the analysis and between 0.44 and 0.75 for the forecast (Fig. 10.6). The time series plots (Fig. 10.7) show mostly a very good agreement between model and observations, except for a small vertical offset for some stations.

The comparison between mean diurnal observations and mean model concentrations show that the forecast partly has a negative offset for the high-altitude stations (Fig. 10.8). The analysis partly overestimates CO (SNB, ZUG).

CO results for individual model forecasts:

CO show MNMBs between -25% and 11% (Fig. 10.9). The CHIMERE model shows the largest positive MNMBs of all models, see Fig. 10.9. Correlation coefficients are between 0.16 and 0.75 for the forecasts. The EURAD model mostly shows the weakest performance.

Results for the individual model analyses:

CO mixing ratios range between 35% and -24%, except for the MATCH model which shows large positive MNMBs (up to 44%) (Fig. 10.10). The MATCH model and the SILAM model behave differently in analysis and forecast, with larger positive biases in the analysis, see Figs. 10.9 and 10.10.

 Correlation coefficients between the models vary significantly and range between 0.16 and 0.75.

11 Comparisons with MOPITT CO

11.1 Summary

During JJA 2020 MOPITT shows low CO values over entire region. The ENSEMBLE forecast date are in good agreement with the observations. The underestimation over the land part of the domain and overestimation over the ocean part are within 10% with some regional exceptions. The overestimated spot by 20% can be seen over Northern Italy in August 2020. The ENSEMBLE analysis data are very similar to the forecast data, except some overestimation (10%) in the middle part of domain in July-August and larger area with a positive bias over Northern Italy in August.

11.2 Method

CO total column forecasts over Europe from nine regional models and the model ENSEMBLE are compared with CO total column retrievals from MOPITT Version 8 (thermal infrared radiances) (Emmons et. al., 2009). Modelled CO data were converted from µg/m³ to VMR by using temperature obtained from CAMS-global (o-suite) model. Pressure at the middle of the layers was also interpolated from the global model. Regional model data are available from the surface up to altitude of 5 km. For the comparison with satellite retrievals, the averaging kernels were applied to the modelled data.

Regional model data up to 5 km were merged with CAMS-global data above 5 km in order to minimize uncertainty error. We performed several confidence tests to establish the method. To check the error due to coarse sampling of the profiles up to 5 km as provided by the regional models, CAMS-global data were sampled at the height levels of the regional models up to 5 km and merged with the CAMS-global original levels above 5 km. Comparison of this results with the original CAMS-global data showed that the errors due to coarse sampling of the profiles up to 5 km were very small. Both results showed slight underestimation of the MOPITT data. CAMS-global values up to 5 km sampled at the height levels as the regional models and ENSEMBLE data without merging with the levels above 5km show overestimation of the satellite data over almost entire region. From this we concluded that error due to missing values above 5 km is significant and merging the regional data with CAMS-global values above 5 km is necessary for the proper comparison.

MOPITT shows relatively low CO values during June-July-August. Slightly higher values can be seen over UK, north of France and Germany and eastern part of domain and over the Mediterranean Sea in June. The ENSEMBLE forecast data show underestimation over the land part of the domain and overestimation over the ocean part by about 10% with some regional exceptions. The negative bias is slightly higher in June. Small overestimated spot by 20% can be seen over Northern Italy in August. All the models have very similar patterns to ENSEMBLE with some exceptions. SILAM and DEHM show bigger negative bias, especially during June for SILAM and during entire season for DEHM (up to 30%). CHIMERE and GEMAQ are in slightly better agreement with the observations. The analysis data are very similar to the forecast data with some exceptions. MATCH and SILAM show overestimation over the central part of domain (up to 30%, especially strong in MATCH). MOCAGE and DEHM show bigger negative bias (up to 30%), especially in June. CHIMERE shows overestimation up to 40% over UK, north of France and north-east spot in domain in June.
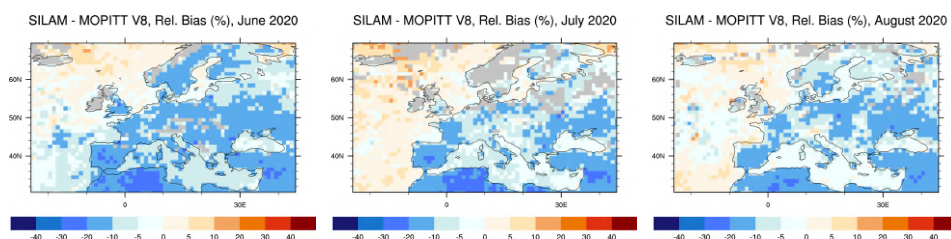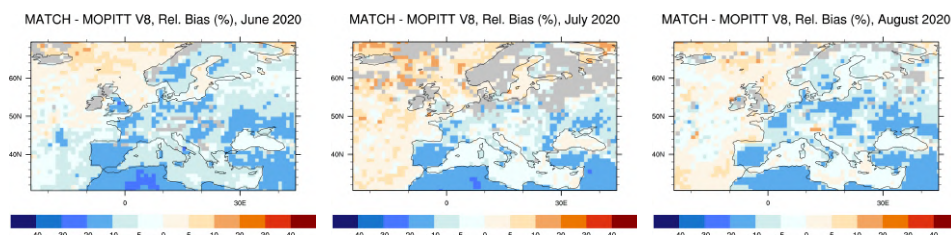
MOPITT
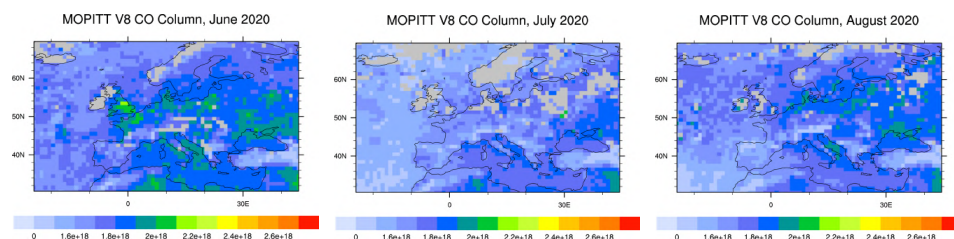
ENS-MOPITT
Forecast (0H24H)
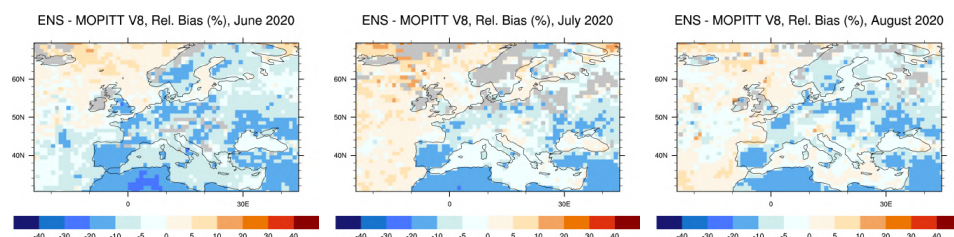
EMEP-MOPITT
Forecast (0H24H)
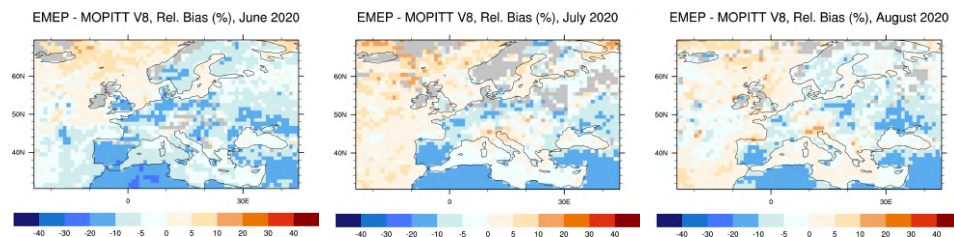
EURAD-MOPITT
Forecast (0H24H)
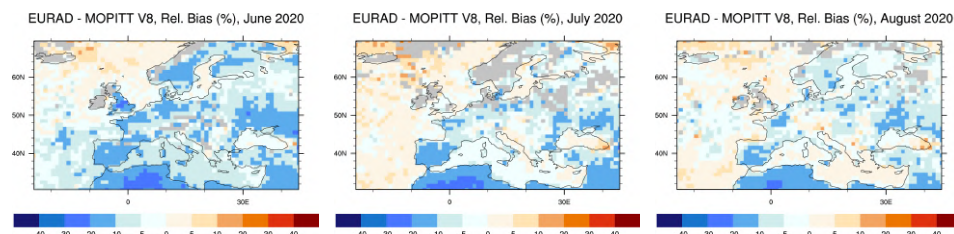
CHIMERE-MOPITT
Forecast (0H24H)
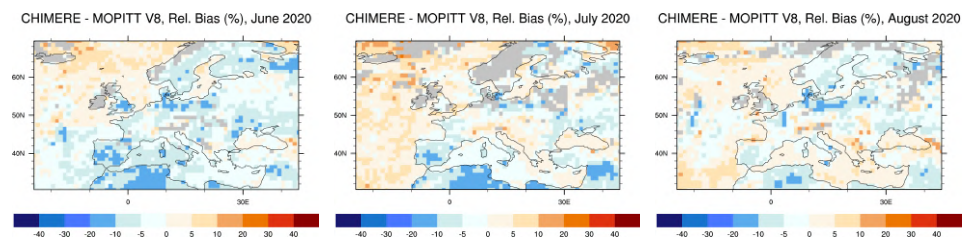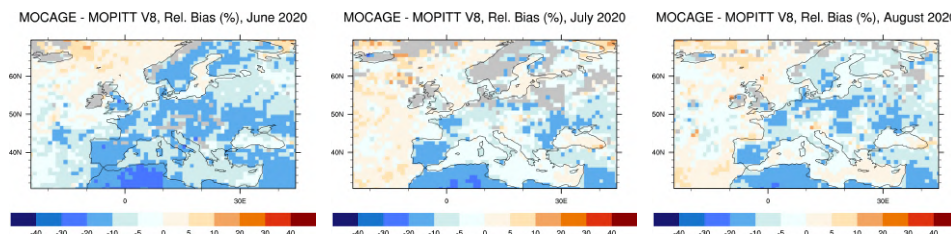
LOTOS-EUROS-MOPITT
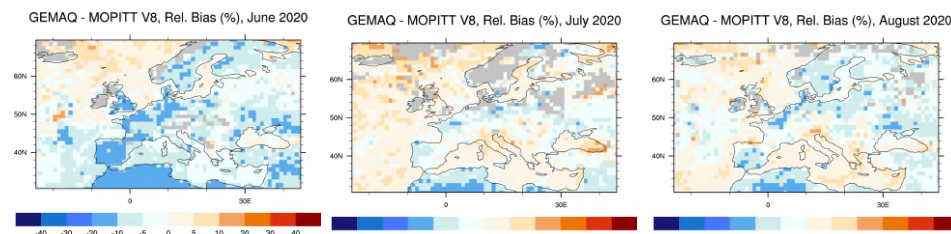Forecast (0H24H)
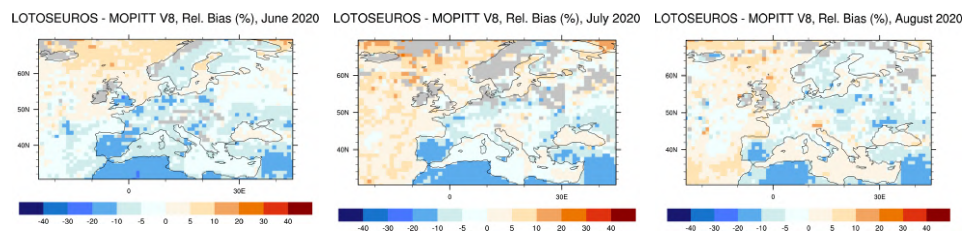
SILAM-MOPITT
Forecast (0H24H)

MATCH-MOPITT
Forecast (0H24H)

MOCAGE-MOPITT
Forecast (0H24H)

GEMAQ-MOPITT
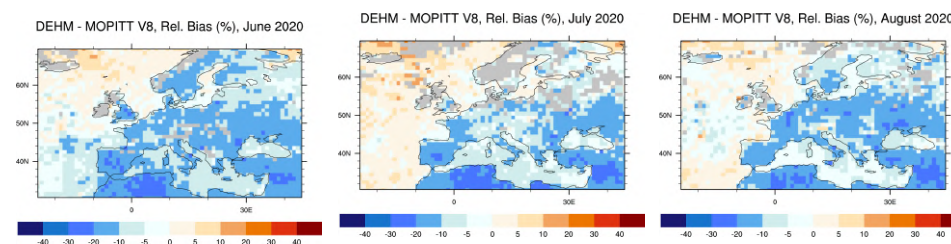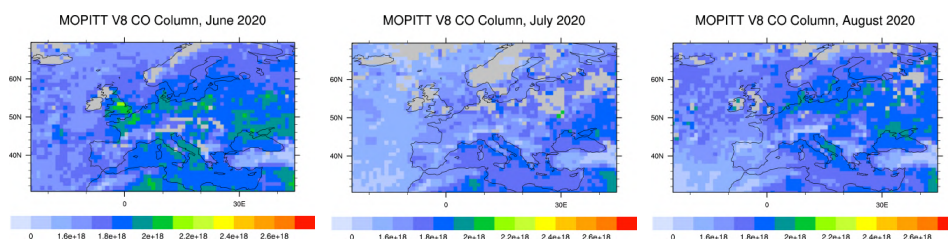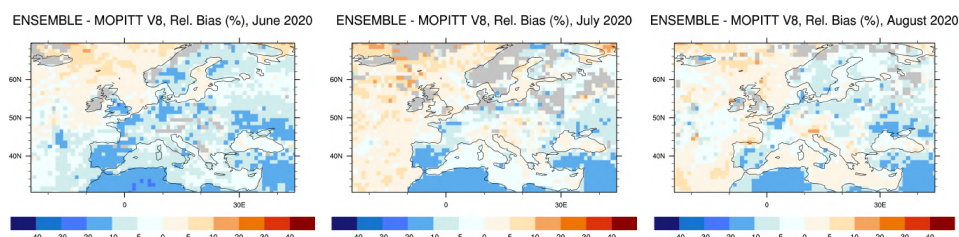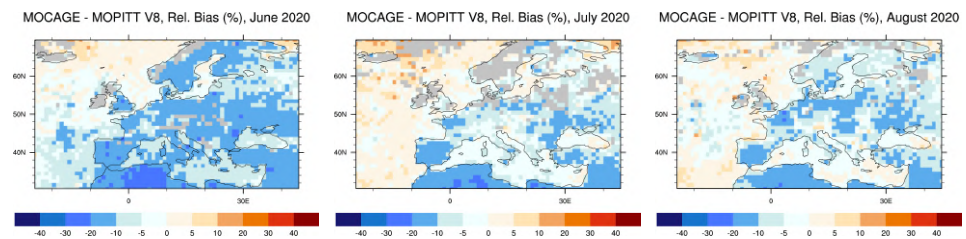Forecast (0H24H)

DEHM-MOPITT
Forecast (0H24H)

Figure 11.1. CO total column for MOPITT V8 satellite retrievals (top row, in molecules/cm^2), relative difference between the regional forecasts of the nine models and the ENSEMBLE and MOPITT (other rows) for June (left column), July (middle column) and August 2020 (right column). Grey colour indicates missing values.

MOPITT

ENS-MOPITT
Analysis (-24H-1H)

EMEP-MOPITT
Analysis (-24H-1H)
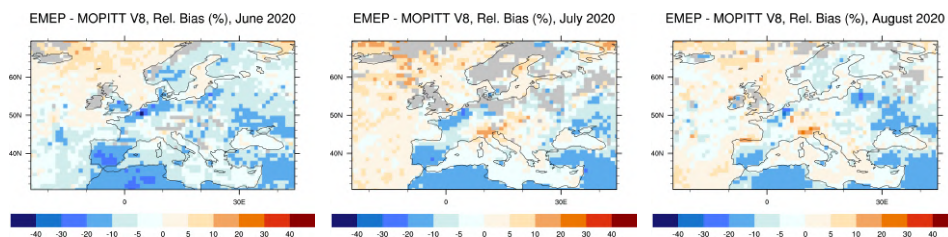
EURAD-MOPITT
Analysis (-24H-1H)
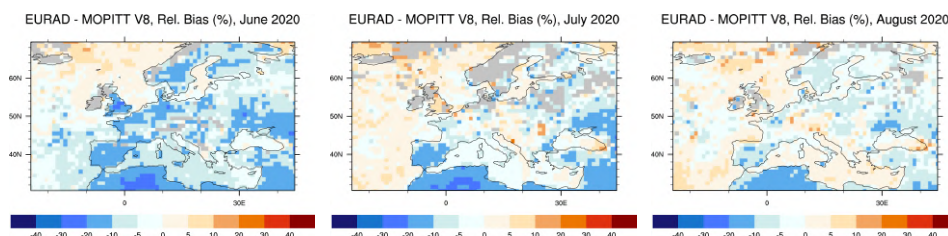
CHIMERE-MOPITT
Analysis (-24H-1H)
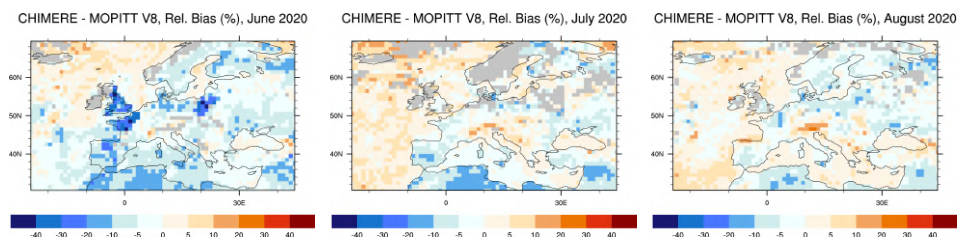
LOTOS-EUROS-MOPITT
Analysis (-24H-1H)
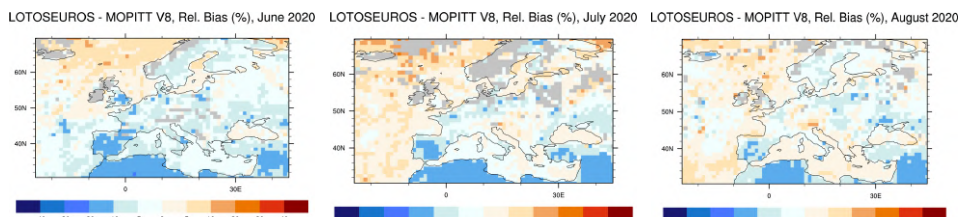
SILAM-MOPITT
Analysis (-24H-1H)
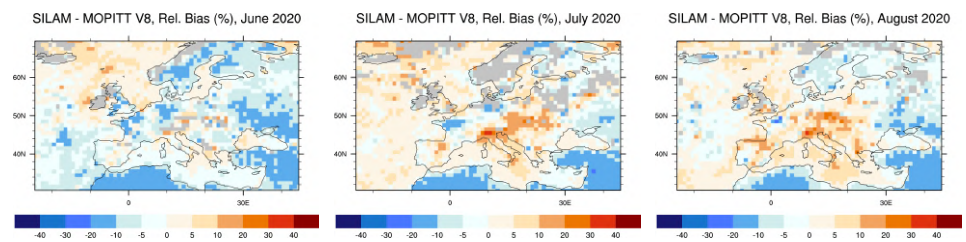
MATCH-MOPITT
Analysis (-24H-1H)
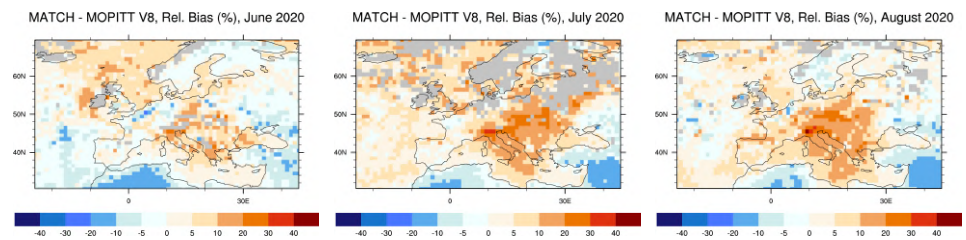
MOCAGE-MOPITT
Analysis (-24H-1H)

Figure 11.2. CO total column for MOPITT V8 satellite retrievals (top row, in molecules/cm^2), relative difference between the regional analyses of the nine models and the ENSEMBLE and MOPITT (other rows) for June (left column) July (middle column) and August 2020 (right column). Grey colour indicates missing values.

12 Summary of issues for individual models

This section provides a short overview of the main, model specific findings of this report.

CHIMERE

Ozone and CO cross sections at the lateral boundaries show features that are not in line with what is seen in the other models (figs 3.11 and 3.12). This indicates a probable issue with the CAMS-global boundary condition implementation. Comparison with the GAW and stations also indicates relatively high MNMBs for ozone (fig 10.4) and CO (fig 10.9).

EURAD

The ozone cross section at the northern boundary shows features that are not in line with what is seen in the other models (fig 3.11). Some relatively large underestimations of ozone are spotted in the free troposphere compared to the other models and IAGOS measurements (fig 5.5). Ozone correlation coefficients at high altitude stations and GAW stations are much lower than for the other models in both forecasts and analyses (figs 8.5, 10.4 and 10.5). Similar results are found in the correlation coefficient for CO based on comparisons at the GAW stations for the forecasts (fig 10.9).

LOTOS-EUROS

Differences between forecasts and analyses for ozone at 5000m seem to be large compared to other models and exhibit some unexpected spatial features.

MATCH

Although the situation is considerably improved compared to previous quarters, the model seems to be still over-correcting for CO concentrations as evidenced by the MNMBs at the GAW stations (fig 10.10) and MOPITT (fig 11.2)

MOCAGE

The model exhibits considerably higher ozone concentrations over central Europe at 250m compared to CAMS-global and most other models. In support of this, comparisons with high-altitude EEA Air Quality e-reporting surface stations for ozone show quite high MNMBs (fig 8.4). High PM10 concentrations are spotted over the Atlantic, up to 250m above the surface (fig 3.7). The western boundary for PM10 is quite high, while the southern boundary is very low (fig 3.13), indicating possible issues with the implementation of the CAMS-global boundary conditions.

SILAM

Comparison with high-altitude EEA Air Quality e-reporting surface stations for ozone, reveal large negative biases over central Europe (fig 8.4).

13 Acknowledgements

14 References

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

Ackermann, J. (1998). The extinction-to-backscatter ratio of tropospheric aerosol: A numerical study. Journal of atmospheric and oceanic technology, 15(4), 1043-1050.

ACTRIS Deliverable WP6/D6.21, http://www.actris.net/Portals/97/deliverables/PU/WP6_D6.21_M45v2.pdf

Cattrall, C., Reagan, J., Thome, K., & Dubovik, O. (2005). Variability of aerosol and spectral lidar and backscatter and extinction ratios of key aerosol types derived from selected Aerosol Robotic Network locations. Journal of Geophysical Research: Atmospheres, 110(D10).

Chin, M., Ginoux, P., Kinne, S., Torres, O., Holben, B. N., Duncan, B. N., ... & Nakajima, T. (2002). Tropospheric aerosol optical thickness from the GOCART model and comparisons with satellite and Sun photometer measurements. Journal of the atmospheric sciences, 59(3), 461-483.

Eskes, H.J., S. Basart, A. Benedictow, Y. Bennouna, A.-M. Blechschmidt, S. Chabrillat, Y. Christophe, E. Cuevas, H. Flentje, K. M. Hansen, J. Kapsomenakis, B. Langerock, M. Ramonet, A. Richter, M. Schulz, N. Sudarchikova, A. Wagner, T. Warneke, C. Zerefos, Observation characterisation and validation methods document, Copernicus Atmosphere Monitoring Service (CAMS) report, December 2019, doi: 10.24380/0nsd- wb26.

Eskes, H., Huijnen, V., Arola, A., Benedictow, A., Blechschmidt, A.-M., Botek, E., Boucher, O., Bouarar, I., Chabrillat, S., Cuevas, E., Engelen, R., Flentje, H., Gaudel, A., Griesfeller, J., Jones, L., Kapsomenakis, J., Katragkou, E., Kinne, S., Langerock, B., Razinger, M., Richter, A., Schultz, M., Schulz, M., Sudarchikova, N., Thouret, V., Vrekoussis, M., Wagner, A., and Zerefos, C.: Validation of reactive gases and aerosols in the MACC global analysis and forecast system, Geosci. Model Dev., 8, 3523-3543, doi:10.5194/gmd-8-3523-2015, 2015.

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.

Joly, Mathieu, and Vincent-Henri Peuch, Objective classification of air quality monitoring sites over Europe, Atmospheric Environment 47, 111-123, 2012.

Katragkou, E., Zanis, P., Tsikerdekis, A., Kapsomenakis, J., Melas, D., Eskes, H., Flemming, J., Huijnen, V., xsxcInness, A., Schultz, M. G., Stein, O., and Zerefos, C. S.: Evaluation of near-surface ozone over Europe from the MACC reanalysis, Geosci. Model Dev., 8, 2299-2314, doi:10.5194/gmd-8-2299-2015, 2015.

Marécal, V., Peuch, V.-H., Andersson, C., Andersson, S., Arteta, J., Beekmann, M., Benedictow, A., Bergström, R., Bessagnet, B., Cansado, A., Chéroux, F., Colette, A., Coman, A., Curier, R. L., Denier van der Gon, H. A. C., Drouin, A., Elbern, H., Emili, E., Engelen, R. J., Eskes, H. J., Foret, G., Friese, E., Gauss, M., Giannaros, C., Guth, J., Joly, M., Jaumouillé, E., Josse, B., Kadygrov, N., Kaiser, J. W., Krajsek, K., Kuenen, J., Kumar, U., Liora, N., Lopez, E., Malherbe, L., Martinez, I., Melas, D., Meleux, F., Menut, L., Moinat, P., Morales, T., Parmentier, J., Piacentini, A., Plu, M., Poupkou, A., Queguiner, S., Robertson, L., Rouïl, L., Schaap, M., Segers, A., Sofiev, M., Tarasson, L., Thomas, M., Timmermans, R., Valdebenito, Á., van Velthoven, P., van Versendaal, R., Vira, J.,

and Ung, A.: A regional air quality forecasting system over Europe: the MACC-II daily ensemble production, Geosci. Model Dev., 8, 2777-2813, doi:10.5194/gmd-8-2777-2015, 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.

Mortier, A., Goloub, P., Derimian, Y., Tanré, D., Podvin, T., Blarel, L., ... & Ndiaye, T. (2016). Climatology of aerosol properties and clear-sky shortwave radiative effects using Lidar and Sun photometer observations in the Dakar site. Journal of Geophysical Research: Atmospheres.

Müller, D., Ansmann, A., Mattis, I., Tesche, M., Wandinger, U., Althausen, D., & Pisani, G. (2007). Aerosol-type-dependent lidar ratios observed with Raman lidar. Journal of Geophysical Research: Atmospheres, 112(D16).

Omar, A. H., Winker, D. M., Vaughan, M. A., Hu, Y., Trepte, C. R., Ferrare, R. A., ... & Kuehn, R. E. (2009). The CALIPSO automated aerosol classification and lidar ratio selection algorithm. Journal of Atmospheric and Oceanic Technology, 26(10), 1994-2014.

Pappalardo, G., A. Amodeo, A. Apituley, A. Comeron, V. Freudenthaler, H. Linne, A. Ansmann, J. Bösenberg, G. D'Amico, I. Mattis, L. Mona, U. Wandinger, V. Amiridis, L. Alados-Arboledas, D. Nicolae, and Wiegner, M.: EARLINET: towards an advanced sustainable European aerosol lidar network, Atmos. Meas. Tech., 7, 2389–2409, www.atmos-meas-tech.net/7/2389/2014/, doi:10.5194/amt-7-2389-2014, 2014.

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

Vinken, G. C. M., Boersma, K. F., van Donkelaar, A., and Zhang, L.: Constraints on ship NOx emissions in Europe using GEOS-Chem and OMI satellite NO2 observations, Atmos. Chem. Phys., 14, 1353-1369, doi:10.5194/acp-14-1353-2014, 2014.

Wagner, A., M. Schulz, Y. Christophe, M. Ramonet, H. J. Eskes, S. Basart, A. Benedictow, Y. Bennouna, A.-M. Blechschmidt, S. Chabrillat, E. Cuevas, A. El- Yazidi, H. Flentje, K.M. Hansen, U. Im, J. Kapsomenakis, B. Langerock, A. Richter, N. Sudarchikova, V. Thouret, T. Warneke, C. Zerefos, Validation report of the CAMS near-real-time global atmospheric composition service: Period September - November 2019, Copernicus Atmosphere Monitoring Service (CAMS) report, CAMS84_2018SC2_D1.1.1_SON2019_v1.pdf, February 2020, doi:10.24380/xzkk-bz05.

Copernicus Atmosphere Monitoring Service