Manuscript prepared for Atmos. Chem. Phys.
with version 2014/09/16 7.15 Copernicus papers of the LaTeX class copernicus.cls.
Date: 23 March 2015

Evaluation of the GEM-AQ model in the context of the AQMEII Phase 1 project

J. Struzewska[1], M. Zdunek[1], J. W. Kaminski[2,3], L. Łobocki[1], M. Porebska[1], M. Jefimow[1], and L. Gawuc[1]

[1]Institute of Environmental Engineering, Warsaw University of Technology, Warsaw, Poland
[2]EcoForecast Foundation, Warsaw, Poland
[3]Centre for Research in Earth and Space Science, Lasonde School of Engineering, York University, Toronto, Ontario, Canada

Correspondence to: J. Struzewska (joanna.struzewska@is.pw.edu.pl)

Abstract. In the scope of the AQMEII Phase 1 project the GEM-AQ model was run over Europe for the year 2006. The modelling domain was defined using a global variable resolution grid with a rotated equator and uniform resolution of $0.2^\circ \times 0.2^\circ$ over the European continent. Spatial distribution and temporal variability of the GEM-AQ model results were analysed for surface ozone and PM_{10} concentrations. Model results were compared with measurements available in the ENSEMBLE database. Statistical measures were used to evaluate performance of the GEM-AQ model. The mean bias error, the mean absolute gross error and the Pearson correlation coefficient were calculated for the maximum 8 h running average ozone concentrations and daily mean PM_{10} concentrations. The GEM-AQ model performance was characterised for station types, European climatic regions, and seasons. The best performance for ozone was obtained at suburban stations and the worst performance was obtained for rural stations where the model tends to underestimate. The best results for PM_{10} were calculated for urban stations, while over most of Europe concentrations at rural sites were too high. Discrepancies between modelled and observed concentrations were discussed in the context of emission data uncertainty as well as the impact of large scale dynamics and circulation of air masses. Presented analyses suggest that interpretation of modelling results is enhanced when regional climate characteristics are taken into consideration.

1 Introduction

The Air Quality Model Evaluation International Initiative (AQMEII; Galmarini et al, 2012) is a collaborative project aimed at improving our understanding of uncertainties and limitations of regional-scale air quality models. During Phase 1 of AQMEII, air quality simulations encompassing two

domains, Europe and North America, were carried out for 2006. Several results of this initiative have already been published. Solazzo et al (2012a) reported operational evaluation (Dennis et al, 2010) of particulate matter predictions by 10 models; this paper may serve as a general reference for the intercomparison methodology, data used and the participating models; evaluation of ozone predictions was presented by Solazzo et al (2012b). Appel et al (2012) presented and discussed results obtained with the Community Multiscale Air Quality (CMAQ) model using different boundary conditions. Brandt et al (2012) reported modelling results with the Danish Eulerian Hemispheric Model (DEHM). Pirovano et al (2012) discussed differences in simulations with the Comprehensive Air Quality Model with Extensions (CAMx) and the CHIMERE model. CAMx predictions were discussed by Nopmongcol et al (2012), while CMAQ simulations were done and analysed by Tagaris et al (2013).

As a part of Phase 1 of the AQMEII project, a simulation of air quality in Europe was conducted using the Global Environmental Multiscale – Air Quality (GEM-AQ, Kaminski et al, 2008). This model differs from other Phase 1 participating models in two aspects. First, it is a multiscale model that can cover the entire globe using a uniformly spaced latitude/longitude grid, a global variable resolution grid or a limited-area extent. Second, the atmospheric chemistry model is implemented on-line within the meteorological model, sharing the advection and subgrid transport schemes. As the GEM-AQ model was used in its global variable-resolution mode, this simulation required neither externally supplied meteorological fields nor lateral boundary conditions. The annual simulation presented here consisted of a series of daily runs, each initialized with a global meteorological objective analysis (Gauthier et al, 1999) done using the 3D-Var assimilation method, and utilizing air quality results from the previous day as initial conditions for the air quality module. Results from the GEM-AQ model simulations were already used by Solazzo et al (2013) to address diversity in multi-model ensembles.

In this paper we present a comprehensive operational evaluation of the GEM-AQ model. Concentrations of ozone and PM_{10} are compared with surface measurements.

The database of the AQMEII project contained hourly measurements of pollutant concentrations taken at rural, suburban and urban sites. Ozone concentrations were available from 472 rural, 391 suburban and 527 urban stations. For PM_{10} there were 119, 110 and 263 stations, correspondingly. To address the possible representativeness issues pertinent to the spatial resolution adopted in the project, we shall present the results stratified according to the station types.

2 Model description

As the GEM-AQ model used in this study has been extensively documented elsewhere (Côté et al, 1998a; Côté et al, 1998b; Mailhot et al, 2006; Kaminski et al, 2008; Gong et al, 2012), we shall focus on the model configuration choices used in this study. The meteorological component of the

system, the GEM model, is a medium-range operational weather forecast model of the Canadian Meteorological Centre.

2.1 Model formulation

The GEM model used in this study solves the hydrostatic primitive equations cast in spherical co-
60 ordinates (with a rotated equator plane in order to minimize distortions over the uniform resolution area), with a terrain following, hybrid pressure-type vertical coordinate comprising 28 levels, of which 8 may fall into a well-developed boundary layer. The top of the model domain is located at $10\,\mathrm{hPa}$ whereas the height of the lowest atmospheric level is approximately $40\,\mathrm{m}$. The model uses a semi-Lagrangian time discretization with a semi-implicit approximation of terms that give
65 rise to fast gravitational modes (Robert, 1985) – a feature crucial for its multiscale applications as well as for integration using variable-resolution meshes. This scheme permits using time steps several times longer than in, for instance, a split-explicit method. With a few exceptions, choices of physical parameterizations made for this study follow the 15-km version of the Canadian Regional Forecast System as documented by Mailhot et al (2006). The turbulence parameterization is based
70 on a turbulent kinetic energy budget with inclusion of statistical subgrid-scale cloudiness (Bélair et al, 2005) and the Bougeault–Lacarrere specification of the length scale (Bougeault and Lacarrere, 1989); surface energy budget is modelled with the force-restore equation (Deardorff, 1978). Gravity wave drag effects are taken into account using a modified McFarlane parameterization (McFarlane, 1987; McLandress and McFarlane, 1993). Condensation processes are handled with the Kain-Fritsch
75 deep convection scheme (Kain and Fritsch, 1990, 1993), the so-called Kuo Transient shallow convection parameterization (see Bélair et al, 2005), and the Sundqvist (Sundqvist, 1978) scheme for non-convective clouds. Solar and infrared radiation is modelled using Fouquart and Bonnel (1980) and Garand (1983) schemes, fully interactive with clouds.

The gas-phase chemistry package of GEM-AQ (Kaminski et al, 2008) describes 116 chemical and
80 19 photolysis reactions among 50 species or groups of species, and is based on the second version of the Acid Deposition and Oxidation Model (Venkatram et al, 1988; Lurmann, 1986) with extensions for free tropospheric chemistry. For a complete list of species and reactions, see Kaminski et al (2008). A simplified aqueous-phase reaction module allows for oxidation of SO_2 to sulphate. Aerosols are modelled with a sectional module CAM (Canadian Aerosol Module, Gong et al, 2003)
85 with 5 aerosols types: sulphate, black carbon, organic carbon, sea-salt and soil dust, size-segregated into 12 logarithmically spaced bins. The aerosol module includes parameterizations of nucleation, condensation, coagulation, sedimentation and dry deposition, in-cloud oxidation of SO_2 and scavenging, and below-cloud scavenging of aerosol species by rain and snow.

2.2 Model configuration

For this study the GEM-AQ model was run on the global variable grid with rotated equator and with the resolution of $0.2^\circ \times 0.2^\circ$ over the European continent. Number of grid points was set on the globe to 288×264 and 197×190 in the core part (Fig. 1). In the vertical, 28 sigma-hybrid layers extending to $10\,\mathrm{hPa}$ were used. The simulation was performed from 1 January to 31 December 2006, as a set of 30 h forecasts with a 6 h overlap. The integration time step of $600\,\mathrm{s}$ was used.

Emission data were prepared for the experiment by TNO (TNO, Utrecht, the Netherlands) using MACC (Monitoring Atmospheric Composition and Climate) methodology (Pouliot et al., 2012). Anthropogenic emissions included primary gaseous pollutants such as O_2, NO_x, CO, NMVOC, NH_3, CH_4 and particle pollution of the fine and coarse mode for individual SNAP (Standardized Nomenclature for Air Pollutants) sectors. Hourly biogenic emissions provided by the AQMEII project were used. Emissions outside the area provided by AQMEII were compiled using EDGAR 2.0 (Emission Database for Global Atmospheric Research, for 1990 base inventory year) and GEIA (Global Emissions Inventory Activity) global inventories (Olivier et al., 1999; Olivier and Berdowski, 2001). Anthropogenic emissions were distributed within the four lowest model layers (up to $\sim 630\,\mathrm{m}$) with different injection height profiles for each of the SNAP sectors. Temporal profiles modulating annual and diurnal variation of emission fluxes for each SNAP were used. Surface anthropogenic and biogenic emission fluxes were applied as a bottom boundary condition in the vertical diffusion equation.

3 Modelling results

Following the methodology used in previous publications describing the AQMEII Phase 1 results, the evaluation was undertaken with respect to station type (Hogrefe et al, 2013; Nopmongcol et al, 2012; Pirovano et al, 2012) and with respect to climatic differences between geographical regions in Europe (Solazzo et al, 2012a, b; Pirovano et al, 2012; Putaud et al, 2010). However, in contrast to previous publications, in our analysis four different climatic regions were chosen. The selected regions follow the Koeppen climate classifications for Europe (Fig. 2). Region I – Northern Europe, the Scandinavian Peninsula and Finland, reflects boreal continental climate. Region II – Western Europe, reflects maritime climate with the influence of an inflow from over the Atlantic Ocean. Region III – Central and Eastern Europe, is characterized mainly by transitional and warm summer continental climate (some countries were not included, as observations were not available in the database used). Region IV covers regions with the Mediterranean climate.

The following statistical measures were used to evaluate performance of the GEM-AQ model: mean bias error (MBE); mean absolute gross error (MAGE) and Pearson correlation coefficient.

3.1 Ozone

Analysis of ozone concentrations variability was based on daily maximum 8 h running average. Evaluation of the GEM-AQ model performance was done for 1386 stations available in the ENSEMBLE database (Galmarini et al, 2001, 2004).

3.1.1 Spatial distribution of ozone concentrations

Spatial and temporal variability of the modelled ozone concentrations as well as the mean bias error with respect to the type of station (rural/suburban/urban) were assessed on a seasonal basis.

Spatial distribution of model data and model performance statistics for maximum 8-h running average ozone concentrations during winter months (DJF) and mean bias error for three types of stations is shown in Fig. 3. The calculated ozone concentration over most of Europe is in the range 40–50 $\mu g\,m^{-3}$. Over the North Sea and the Baltic Sea concentrations are lower (30–40 $\mu g\,m^{-3}$). Lowest ozone levels (below 30 $\mu g\,m^{-3}$) were calculated over regions characterized with high NO_x emission (i.e. Benelux, the Po Valley, London, Paris). Concentrations higher than 50 $\mu g\,m^{-3}$ are modelled in Southern Europe, with the maximum (up to 70 $\mu g\,m^{-3}$) in mountain regions (the Alps, eastern part of the Carpathian Mountains, Pyrenees, and the Balkan Mountains).

During the winter (DJF) MBE for most of the rural stations varies in the range $-10 \div 10\,\mu g\,m^{-3}$. Over Scandinavia and the British Isles the underestimation is higher, up to $-20\,\mu g\,m^{-3}$. Measurements at rural stations located in the Alps are higher than modelled results. Suburban stations (available for the analysis) are located mainly in Western Europe. For most of these stations the bias is small and positive, in the range 0 to 10 $\mu g\,m^{-3}$. However, in the mountain regions the bias is up to 30 $\mu g\,m^{-3}$. Spatial coverage is most complete in Western Europe for urban stations. At most of the sites the bias is positive and small, in the range of 0–10 $\mu g\,m^{-3}$. The model underestimated ozone concentrations for a few stations located in Scandinavia. Overestimation of up to 30–40 $\mu g\,m^{-3}$ was over industrial regions in Europe (i.e. Southern Italy, South-Eastern and Northern France, the Ruhr region, Silesia and the Moravian Gate). In Northern Spain, in the same region the model underestimated ozone concentration levels for rural sites and overestimated for urban stations.

Spatial distribution of ozone concentrations and MBE for the spring (MAM) are shown in Fig. 4. Over most of Europe the maximum 8 h running average ozone concentrations were in the range of 80 to 100 $\mu g\,m^{-3}$. Over Scandinavia and the British Isles ozone concentrations were below 70 $\mu g\,m^{-3}$. Ozone concentrations higher than 100 $\mu g\,m^{-3}$ were calculated in mountain regions in South-Eastern Europe (the Apennine Mountains) and in the Alps. In Western and Central Europe elevated ozone concentrations were over industrial areas (i.e. Silesia and the Moravian Gate). MBE at rural stations is generally smaller than in winter months. For most sites the error is in the range $-10 \div 10\,\mu g\,m^{-3}$. In Central Europe and the British Isles the discrepancies were reduced as compared to the winter months, while in Scandinavia the underestimation is larger, especially near the Baltic Sea coast. In

the Alps the overestimation is smaller than in the winter months. For suburban stations the model performance is good. The underestimation of modelled results is mainly in coastal areas. In most of the urban sites the model overestimates ozone levels up to $20\,\mu g\,m^{-3}$ while the underestimation occurred at the costal stations.

Summer (JJA) modelled ozone concentrations (Fig. 5) show uniform distribution in the range of 90 to $100\,\mu g\,m^{-3}$ in Western and Eastern Europe. Lower concentrations over Scandinavia are due to lower solar irradiance, lower temperature and lower emissions of ozone precursors. Over the south-eastern part of the British Isles concentrations over $80\,\mu g\,m^{-3}$ were calculated. Highest concentrations (over $120\,\mu g\,m^{-3}$) were calculated over the Po Valley and the Iberian Peninsula. Ozone concentration over the Atlantic Ocean is lower than in spring months, with values below $60\,\mu g\,m^{-3}$.

During summer months (JJA) the distribution of MBE for rural stations shows underestimation over the British Isles and the Iberian Peninsula larger than in spring months (-20 to $-10\,\mu g\,m^{-3}$). In Scandinavia and Central Europe modelled concentrations better agree with observed values. In South-Eastern Europe the model tends to overestimate observed concentrations. The model tends to overestimate ozone concentrations as compared to observation at suburban stations. In Western and Central Europe MBE is in the range of -10 to $10\,\mu g\,m^{-3}$, with highest values over the Benelux and industrial regions on the border of France and Germany. Over the Iberian Peninsula the model underestimates ozone concentrations. MBE calculated for urban stations is in most cases positive. However, MBE is significantly reduced as compared to winter and spring months.

The distribution of ozone concentrations in the fall is shown in Fig. 6. Over most of Europe the ozone concentrations exceed $60\,\mu g\,m^{-3}$. Lower concentrations were calculated near the British Isles, the North Sea, Scandinavia and Eastern Europe. Highest concentrations over $80\,\mu g\,m^{-3}$ are modelled in mountain regions in the southern part of the continent. Ozone concentrations below $40\,\mu g\,m^{-3}$ were calculated in South-Eastern Europe.

In autumn (SON) for most rural stations MBE is positive, in the range of 0 to $10\,\mu g\,m^{-3}$. Underestimation up to $-10\,\mu g\,m^{-3}$ was over Scandinavia and the British Isles and over the Iberian Peninsula. Significant overestimation was in the Alps (up to $30\,\mu g\,m^{3}$). Positive bias in the range of 10 to $20\,\mu g\,m^{-3}$ was over the industrial regions along the border of Germany and France and over the Netherlands. For urban and suburban stations the spatial distribution of MBE is similar. Bias is positive in the range of 0 to $20\,\mu g\,m^{-3}$. Highest error values were calculated in the Alps.

3.1.2 Temporal variability of ozone concentrations

In order to study the model performance on a daily basis, maximum 8 h running average concentrations were calculated and averaged over stations in four regions of the European continent characterized with different climatological conditions: Western Europe, Northern Europe, Central and Eastern Europe and Southern Europe (Fig. 7–14).

In Northern Europe there were only 45 stations available for the comparison. Model shows systematic underestimation during the autumn, winter and spring; MBE is $-16.8\,\mu g\,m^{-3}$ and MAGE is $20.7\,\mu g\,m^{-3}$. The differences between modelled and observed concentration values are much smaller from June to September (Fig. 7). In contrast to other regions where highest concentrations were observed from May to July, in this region the period with highest concentrations exceeding $100\,\mu g\,m^{-3}$ was observed at the end of April and the beginning of May, while during the summer months there were two high concentration episodes. In spite of the systematic bias, the correlation coefficient is 0.83, which shows good agreement in terms of changes related to exchange of air masses. Scatterplot representing concentrations averaged over days (Fig. 8) shows that levels in the range 60–80 were reproduced the best.

In Western Europe the agreement between model and observations was analysed for 791 stations. Temporal variability of the maximum 8 h running average ozone concentrations was captured very well. MBE is low $-0.4\,\mu g\,m^{-3}$ and MAGE $16.5\,\mu g\,m^{-3}$. Small overestimation was in January, March, September and October. High concentrations were observed in June and July (Fig. 9). Three high concentration episodes can be distinguished. During episodes the model underestimated peak values by $\sim 20\,\mu g\,m^{-3}$.

Short term variability is reproduced correctly and the correlation coefficient is 0.91. Scatterplots show that major discrepancies are for the highest concentrations during summer episodes ($> 100\,\mu g\,m^{-3}$) (Fig. 10).

Analysis of the ozone concentrations variability in Central Europe was undertaken for 251 stations. The model underestimated concentrations in March and April and during summer episodes (Fig. 11). MBE was $-1.5\,\mu g\,m^{-3}$ and MAGE $17.6\,\mu g\,m^{-3}$. First period with exceptionally high concentrations was in the beginning of May and was not reproduced by the model. Rapid increase of ozone concentrations in Central and Eastern Europe was connected with the inflow of aerosols from biomass fires in Eastern Europe, which plausibly contributed to changes of photodissociation rates. Other periods with high concentrations were caused by meteorological conditions favouring ozone production, and the increased ozone levels were captured by the model although modelled peak values were lower than observed. The value of the correlation coefficient is 0.89. The scatter plots show slightly larger variation than over Western Europe. Model tends to under-predict concentrations lower than $60\,\mu g\,m^{-3}$ and higher than $100\,\mu g\,m^{-3}$, while concentrations in the range 80–$100\,\mu g\,m^{-3}$ are in general over-predicted (Fig. 12).

In Southern Europe measurements from 303 stations were available for the comparison. The short term variability is reproduced well and the correlation coefficient is 0.96. During most of the year, modelled ozone levels show low negative bias $-1.4\,\mu g\,m^{-3}$ that is due to under-estimation of ozone concentrations in June, July and August (Fig. 13). Although concentration increase during episodes was consistent with observed variability, maximum values were 20–$40\,\mu g\,m^{-3}$ lower than observed. Even though MAGE is $19.9\,\mu g\,m^{-3}$, scatterplot for stations located in Southern Europe shows best

linear fit (Fig. 14). Scatterplot presenting annual average concentrations indicate that for observation at sites characterized with highest concentrations (above $100\,\mu g\,m^{-3}$), the model underestimated ozone levels.

3.2 PM_{10}

The analysis of model performance was undertaken for 492 stations available in the ENSEMBLE database for PM_{10} concentration. The 24 h averages were calculated based on hourly measurements and model results.

3.2.1 Spatial distribution of PM_{10} concentrations

The pattern of modelled daily averaged PM_{10} concentrations during winter months (DJF) and spatial distribution of the mean bias error for different types of stations is shown in Fig. 15. The calculated PM_{10} concentrations over Central and South-Eastern Europe are in the range 40–60 $\mu g\,m^{-3}$. Over the rest of the continent concentrations are lower – below $30\,\mu g\,m^{-3}$. Highest PM_{10} concentrations above $60\,\mu g\,m^{-3}$ are modelled over Romania and the eastern part of Germany, with maximum up to $70\,\mu g\,m^{-3}$.

MBE for most rural stations is positive and high – up to $30\,\mu g\,m^{-3}$. However, it should be noted that the rural stations available for the comparison are located mainly in Germany, the Czech Republic and Benelux. The overestimation was in Germany and the Czech Republic, while in Benelux, British Isles and Spain the model performs quite well. The highest overestimation is modelled over the eastern part of Germany for suburban stations. In Central and Western Europe the model performs well, while in Spain PM_{10} levels were underestimated up to $20\,\mu g\,m^{-3}$. Spatial coverage of urban monitoring sites is more complete. For stations located in Germany and the Czech Republic the model overestimated PM_{10} levels. However, in Central and Southern Europe the model results are lower than the observed concentrations.

Spatial distribution of daily average PM_{10} concentrations and MBE for the spring (MAM) is shown in Fig. 16. Over most of Europe the maximum 24 h PM_{10} concentrations were low – in the range of 15–30 $\mu g\,m^{-3}$. Highest concentrations were calculated over Central Europe, with maximum values over Poland and the eastern part of Germany. PM_{10} concentrations lower than $10\,\mu g\,m^{-3}$ were modelled over Scandinavia and over the south-western part of the continent.

The distribution of MBE at rural stations shows good agreement over most of Europe, in spite of overestimations over the Czech Republic and Germany. For most of the suburban monitoring sites MBE was lower than $10\,\mu g\,m^{-3}$, except for Eastern Germany where the model tends to overestimate, and for some stations in Spain, where modelled PM_{10} concentrations were lower than observed. For urban stations the pattern of MBE spatial distribution was similar to MBE calculated at suburban stations. MBE varies in a range -10 to $10\,\mu g\,m^{-3}$.

Summer (JJA) PM_{10} concentrations (Fig. 17) show uniform distribution in the range of 20–30 $\mu g\,m^{-3}$ in Western and Eastern Europe. Lower concentrations (below 20 $\mu g\,m^3$) were calculated over Scotland, Scandinavia and North-Eastern Europe. Highest concentrations above 30 $\mu g\,m^{-3}$ were calculated over Central Europe, with the maximum over the eastern part of Germany. MBE distribution for all types of stations shows a similar pattern. Over Germany the model overestimates PM_{10} levels, while over Spain the model underestimated mainly for suburban sites. Over the rest of the continent the modelled and observed PM_{10} concentrations agree well, with small positive bias lower than 10 $\mu g\,m^{-3}$.

The distribution of PM_{10} concentrations in autumn (SON) is shown in Fig. 18. The maximum 24 h averaged PM_{10} concentrations were in the range of 20 to 40 $\mu g\,m^{-3}$ over most of Europe. As in other periods, highest concentrations were calculated over Central Europe, with maximum values exceeding 60 $\mu g\,m^{-3}$ over the eastern part of Germany. The lowest modelled PM_{10} concentrations (lower than 20 $\mu g\,m^{-3}$) occur over Scandinavia and over the south-western part of the continent.

MBE distribution is similar to that in winter. MBE for rural stations is positive and shows highest discrepancies over the eastern part of Germany. For urban and suburban stations the overestimation of PM_{10} levels occurs over Germany, Czech and industrial regions in Poland. In Western Europe the agreement is better, with MBE below 10 $\mu g\,m^{-3}$. MBE over the Iberian Peninsula varies in the range -20 to 10 $\mu g\,m^{-3}$.

3.2.2 Temporal variability of PM_{10}

A detailed analysis of the temporal variability was undertaken for four climatological regions in Europe. In Northern Europe (for 19 stations) the model overestimated concentrations in January, while in February and March the agreement between model and observations was good. In April and the beginning of May there was an episode of elevated PM_{10} concentrations which was not captured by the model. Till mid-August observed and modelled PM_{10} levels agree well. In autumn the model tends to overestimate on average 5–10 $\mu g\,m^{-3}$. Although the modelled concentrations reflect observed values with MBE 0.6 $\mu g\,m^{-3}$ and MAGE 10.3 $\mu g\,m^{-3}$, the short and mid term variability was reproduced only during some months and the correlation coefficient is 0.32.

In Western Europe the agreement between model and observations was analysed for 251 stations. Temporal variability of 24 h average PM_{10} concentrations was captured very well, with the correlation coefficient of 0.73. However, the model tends to overestimate PM_{10} concentrations for all seasons – MBE is 8.9 $\mu g\,m^{-3}$ and MAGE 16.7 $\mu g\,m^{-3}$. Highest overestimation was modelled during autumn and winter months: January, February, October and December. Best agreement between modelled and observed PM_{10} concentrations was in spring months, while in summer there was a small systematic overestimation $\sim 10\,\mu g\,m^{-3}$.

In Central Europe 129 stations were available for the comparison. The model underestimated concentrations during severe episodes in January and the beginning of February. In spring observed

PM_{10} concentrations are reproduced correctly by the model. From June there is a small systematic overestimation averaging 10 to 15 $\mu g\,m^{-3}$. During the last quarter of the year the model overestimated PM_{10} concentrations, MBE was 5.9 $\mu g\,m^{-3}$, MAGE 24.2 $\mu g\,m^{-3}$ and the correlation coefficient was 0.72. Overall variability of PM_{10} levels was captured. However, the model overestimated up to 30 $\mu g\,m^{-3}$ during three periods in October, November and December that were not supported by observations.

In Southern Europe analysis was undertaken for 93 stations. The temporal variability is not reproduced by the model. However, some incidental agreement leads to the correlation coefficient of 0.56. Negative bias $-9.4\,\mu g\,m^{-3}$ is due to the underestimation of PM_{10} levels (Fig. 22), with MAGE 16.2 $\mu g\,m^{-3}$. Overall modelled PM_{10} levels are lower than observed. Model overestimated only during a singe episode at the end of May.

4 Discussion

Spatial and temporal averaging is a common methodology used in model performance analysis for long-term simulations. However, it is clear that the averaging leads to error compensation and does not reflect the model's ability to reproduce specific features of concentrations distribution. The choice of the averaging period (e.g. month, seasonal, annual) impacts the value of the bias error. Also, similar "average performance" might be obtained for cases with small systematic errors and large positive and negative errors.

4.1 Ozone

For rural background stations (472 sites) the GEM-AQ model underestimated ozone concentrations during the cold season (January to mid March and November–December). During the summer underestimation of peak values ranges from 30 to 40 $\mu g\,m^{-3}$. Although the correlation coefficient is high (0.92), MBE is $-7\,\mu g\,m^{-3}$ and MAGE is 18 $\mu g\,m^{-3}$ (Table 1). For suburban (391 sites) and urban stations (527 sites), the characteristics of variability range and agreement with measurement is comparable. In January, March, September and October the model slightly overestimated ozone concentrations. In June and July during pollution episodes modelled concentrations are lower than observed by 10 to 20 $\mu g\,m^{-3}$. Results for other months show very good agreement in terms of average concentration levels and short term variability, which is confirmed with very high correlation coefficient – 0.93 for both types of stations. MBE is positive 1.9 $\mu g\,m^{-3}$ for suburban stations and 2.5 $\mu g\,m^{-3}$ for urban stations, with MAGE 17.2 and 17.4 $\mu g\,m^{-3}$, respectively. In contrast to Hogrefe et al (2013) that reported higher correlation coefficient for ozone concentrations for rural sites than for "locally influenced sites" the GEM-AQ model performance in terms of the Pearson correlation coefficient was almost at the same level for all types of stations (0.92 – rural sites, 0.93 – suburban and urban sites).

The evaluation results reported by Pirovano et al (2012) for CAMx and CHIMERE models showed that for both models, ozone concentrations at rural stations are reproduced not as well as at suburban and urban stations. This has been confirmed by the GEM-AQ model results. While MAGE is relatively constant (18.2 – rural, 17.2 and 17.4 – suburban and urban respectively), there are significant differences in MBE values for different types of stations.

Analysis of seasonal variability of model performance was undertaken by Nopmongcol et al (2012) using results from the CAMx model for January and July. Modelled ozone concentrations were systematically underpredicted. CAMx performance was the worst at urban stations in January while in July the bias was lower. Also, Pirovano et al (2012) reported strong underestimation of ozone concentrations during the first part of the year for CAMx and CHIMERE models. CMAQ results, described in Appel et al (2012), indicate that model performance for the daytime ozone varied seasonally. In the winter CMAQ overestimated ozone concentration by 8 % and in the spring and summer ozone levels were underestimated on average ~ 4 and 2 %, respectively. In the fall model performance was worst, with significant overestimation by 30 %. Results from the GEM-AQ model also show seasonality in MBE and the Pearson correlation coefficient. However, MAGE is relatively constant for all seasons – on average $17.5 \pm 0.8\,\mu\mathrm{g\,m^{-3}}$. For winter and autumn the model slightly overestimated, with MBE 0.93 and $0.09\,\mu\mathrm{g\,m^{-3}}$, respectively. Although, MAGE and MBE are smallest in autumn, the correlation coefficient is lower than in winter (0.78 in autumn, 0.87 in winter). In the spring and summer the model tends to underestimate. However, with nearly the same MAGE, model results for summer are noticeably better with MBE of $1.4\,\mu\mathrm{g\,m^{-3}}$ and correlation coefficient 0.92 while in spring the bias is larger $-3.2\,\mu\mathrm{g\,m^{-3}}$ and correlation coefficient is lower 0.75.

Analysis of model performance for different regions showed differences between models. Best performance for CAMx and CHIMERE models (Pirovano et al, 2012) was for Southern Europe, while North-Western and Eastern Europe were characterized with a negative bias in the range 10–30 %, especially in winter. In terms of correlation coefficient, the CHIMERE model performed well in Southern, North-Western and Eastern Europe, while CAMx showed better results in Eastern Europe. The CMAQ model (Appel et al, 2012) overestimated daytime ozone concentrations in the south-western part of the domain and underestimated the north-eastern part including British Isles in winter. Largest overestimation occurred over Northern Italy (Po Valley), while largest underestimation was in the Czech Republic and Poland.

The GEM-AQ model underestimated daytime ozone concentrations with the highest values in Northern Europe (MBE $-16.8\,\mu\mathrm{g\,m^{-3}}$) especially during the cold part of the year. One possible reason relates to uncertainties of NOx emissions in Scandinavia, but also over British Isles and the north-western part of Europe, which may contribute to overestimated titration processes in the model. Another source of possible underestimation is too weak transport of ozone from upper troposphere in high latitudes. However, such analysis is beyond the scope of the presented study, which

is focused on surface air quality and does not include the analysis of the vertical structure of the atmosphere.

For the rest of the domain the agreement between modelled and observed ozone concentrations is good. However, summer episodes are underestimated. During winter and spring months the MBE distribution shows significant overestimation of ozone concentrations over the Alps, which indicates too intensive downward mixing in the mountain regions, where in the cold part of the year concentrations are highest. In spring and autumn there is a systematic overestimation for suburban and urban stations in Benelux, Germany and France that results in positive MBE 0.4 for Region II, while for other regions MBE is negative. In general, best model performance was achieved in the summer.

In winter and autumn topography plays a very important role in the distribution of higher ozone concentration levels, although modelled concentrations seem to be overestimated. Higher concentrations are over the southern part of Europe (with the exception of the Po Valley) and the Mediterranean Sea. In spring concentrations are significantly higher in Southern, Central and Eastern Europe, while in Western Europe and the British Isles the increase is not that significant due to the inflow of relatively clean Atlantic air masses. In summer there is further increase of ozone concentrations over land, with the maximum over Region IV. Elevated ozone concentrations are also over the Mediterranean Sea and the Black Sea. In contrast to ozone distribution patterns in spring ozone concentrations are lower over North-Eastern and Eastern Europe. The reason for this effect might confirm the hypothesis of advective nature of ozone episodes in Eastern Europe and the role of high pressure system blocks during summer months (Struzewska and Kaminski, 2008). However, due to the lack of measurements, model results cannot be evaluated in this region. In autumn concentrations decrease with lowest values in the north-eastern part of Europe.

4.2 $\mathbf{PM_{10}}$

Surface PM_{10} concentrations at rural background stations (119 sites) were systematically overestimated in all season. In January and February the model captured quite well periods with high PM_{10} concentrations, but peak values were overestimated. Best model performance was from mid-February to May. From June to September the systematic bias was relatively constant ~ 10–$15\,\mu\mathrm{g\,m}^{-3}$. In autumn MBE is largest and the model overestimated significantly for all periods with observed higher concentrations. The correlation coefficient is high (0.72), MBE is $10.8\,\mu\mathrm{g\,m}^{-3}$ and MAGE is $18.2\,\mu\mathrm{g\,m}^{-3}$ (Table 2).

At suburban stations (110 sites) model performance is very good with MBE $4.7\,\mu\mathrm{g\,m}^{-3}$ and MAGE $18.4\,\mu\mathrm{g\,m}^{-3}$. From January to April the agreement of modelled and observed time series, averaged over all suburban stations in the domain, is good except for a short winter episode. At the beginning of May the model underestimated PM_{10} concentrations by about 10 to $20\,\mu\mathrm{g\,m}^{-3}$. From June to September modelled and observed concentrations at suburban stations averaged over the domain vary in the range 20–$40\,\mu\mathrm{g\,m}^{-3}$. There is a small systematic overestimation of modelled

PM_{10} – on average $5\,\mu g\,m^{-3}$. As in the case of rural stations, starting in October the discrepancies between model and measurements increase. Model generated three PM_{10} episodes from October to December, which were not observed. In spite of discrepancies, the correlation coefficient is relatively high (0.75), which confirms good agreement in terms of average concentration level and short term variability.

For urban stations (263 sites) the agreement between modelled and observed PM_{10} concentration levels is good. MAGE is comparable to the value obtained for suburban stations ($18.3\,\mu g\,m^{-3}$), MBE is small ($1.3\,\mu g\,m^{-3}$) and the correlation coefficient is 0.72. The model underestimates observed PM_{10} concentration during two severe pollution episodes at urban stations in January and at the beginning of February. From the second week of February concentration decrease and the agreement between modelled and observed daily average PM_{10} levels was good. In May the model did not capture an increase of PM_{10} concentrations. From June to September the model tends to overestimate PM_{10} levels by 5 to $10\,\mu g\,m^{-3}$. In October and December the model showed significant overestimation, while in November the agreement was good.

Daily mean PM_{10} concentrations averaged over suburban and urban stations show good agreement of the modelled and observed concentrations averaged over the domain for values below $40\,\mu g\,m^{-3}$. At suburban stations, modelled PM_{10} concentrations in the range of 60 to $100\,\mu g\,m^{-3}$ were overestimated, while for urban stations highest concentrations, above $60\,\mu g\,m^{-3}$, were underestimated.

In contrast to the GEM-AQ model results, Solazzo et al (2012a) reported that most models used in their analyses had difficulties in reproducing elevated PM concentrations during winter. Most models underestimated PM_{10} levels over Europe, although during summer the performance was better. In the Mediterranean region highest concentrations was in summer months. Results from the CAMx model (Nopmongcol et al, 2012) were substantially underestimated and performance was poor for both January and July, with similar magnitude of error statistics. Also, the DEHM model (Brandt et al., 2012) underestimated PM_{10} concentrations.

The GEM-AQ model tends to overestimate PM_{10} daily mean concentrations. The model performance is clearly better for urban stations. With nearly the same MAGE $18.3 \pm 8\,\mu g\,m^{-3}$ highest overestimation was for rural background stations $10.8\,\mu g\,m^{-3}$. For suburban and urban stations the performance is better – 4.7 and $1.3\,\mu g\,m^{-3}$ and the correlation coefficient is 0.72 and 0.75, respectively.

The work by Pirovano et al (2012) shows that best performance of CAMx and CHIMERE models in terms of PM_{10} concentrations was in North-Western Europe for rural-background stations. In Southern and Eastern Europe model results were underestimated. Appel et al (2012) reported that in winter the domain averaged MBE is $-21.5\,\mu g\,m^{-3}$ for the CMAQ model. For other seasons the underestimation is lower – in the range of $-11 \div -16\,\mu g\,m^{-3}$. The smallest bias was in Northern

France. In spring and summer the bias spatial pattern was similar to the winter case and the bias tends to improve in autumn.

In Southern Europe modelled concentrations were systematically underestimated for all seasons with the bias $-9.41\,\mu g\,m^{-3}$. However, for other regions GEM-AQ tends to overestimate PM_{10} daily mean concentrations. In winter in Northern and Western Europe the model overestimates, while in Central and Eastern Europe there was a strong underestimation. These lead to error compensation in terms of averaged MBE ($0.3\,\mu g\,m^{-3}$) for winter months and relatively high MAGE of $22.5\,\mu g\,m^{-3}$. In the spring model performance is the best in terms of MBE ($1.9\,\mu g\,m^{-3}$) and MAGE ($14.8\,\mu g\,m^{-3}$), but the correlation coefficient is lower than in other seasons (0.51). In autumn, the model overestimated over all regions but the correlation coefficient is relatively high (0.79). The worst performance in terms of correlation coefficient was over Northern Europe (0.32) and Southern Europe (0.56).

The distribution of MBE errors for PM_{10} clearly shows that anthropogenic emission data were overestimated over Germany and the Czech Republic. MBE in this region (especially over the eastern part of Germany) was positive for all types of stations and for all seasons. This impacted the average model bias in Region II. In Region III modelled PM_{10} concentrations were underestimated during severe winter pollution episodes in January and February 2006. During this period low temperatures were observed and differences between modelled and observed PM_{10} levels might be caused by too low emission estimates that did not account for household heating. In addition, as the lowest model layer height is at $\sim 27\,m$, the structure of the stable boundary layer over urban regions might not be fully reproduced. Underestimation of PM_{10} in Southern Europe is most probably due to underestimated mineral dust emissions and transport from North Africa. However, this does not relate to AQMEII emission estimates but to an on-line dust emission module (Marticorena et al, 1995) in the GEM-AQ model. Further work will be undertaken to revise geophysical fields describing soil properties that are used for dust uptake. Highest error was for rural stations, which clearly indicates that emissions in remote regions were overestimated. For urban stations in January, February and March the model systematically underestimated PM_{10} concentrations. The temporal variability of MBE for different types of stations is similar.

Seasonal variability of PM_{10} concentrations is dominated by the distribution and intensity of anthropogenic and natural sources. In winter PM_{10} concentrations are highest over Central-Eastern Europe, although the model tends to overestimate over Germany. Also, elevated concentrations are present over the Northern Atlantic due to sea salt generation during winter storms. In spring, the winter maximum over the Atlantic dissipates and concentration over Northern Africa and Southern Europe is increasing. Highest concentrations remain over Central Europe. Similar pattern was calculated for summer months, with lower maximum values but for higher background. In autumn, PM_{10} concentrations over the Atlantic Ocean and over the North Sea increase and over Southern Europe – decrease. In Eastern Europe PM_{10} background is also higher than in spring and summer months.

5 Summary and Conclusions

In the scope of the AQMEII Phase 1, the GEM-AQ model was run over Europe for the year 2006. Modelling domain was defined using a global variable resolution grid with a rotated equator. The uniform part of the domain with resolution of $0.2^\circ \times 0.2^\circ$ was positioned over the European continent.

Modelled concentrations for ozone maximum 8 h running average and daily mean PM_{10} were analysed in terms of spatial distribution and temporal variability. Model results were evaluated against measurements available in the ENSEMBLE database. For better understanding of the model performance in terms of station representativeness, emission estimates and climate characteristics, the concentration data were averaged for:

- all stations in the domain for the whole year and for each season (DJF/MAM/JJA/SON)
- different types of stations (rural/suburban/urban)
- four climatic regions of Europe

In summary:

1. For ozone values of the mean absolute gross error and the Pearson correlation coefficient are similar for all station types. However, differences in the mean bias error are significant. The best performance (MBE 1.9 µg m^{-3}) for ozone was obtained for suburban locations and slightly higher overestimation (MBE 2.5 µg m^{-3}) was calculated for urban sites. The worst performance was obtained for rural stations where the model tends to underestimate (MBE -7.0 µg m^{-3}).

2. For PM_{10} values of the mean absolute gross error and the Pearson correlation coefficient are similar for all station types. However, differences in the mean bias error are significant. The best results were modelled for urban stations (MBE 1.3 µg m^{-3}), while over most of Europe concentrations at rural sites were overestimated (MBE 10.8 µg m^{-3}) by the model.

3. Although the representativeness of urban stations for model results at the resolution of 20 km may seem questionable, the model performance was worst for rural stations for both analysed species. This may indicate that the emission estimates in rural areas are less accurate smaller than in cities. Perhaps emission accuracy influences model performance more significantly than the sub-grid local features of the emission field.

4. A possible explanation for the systematic underestimation of ozone concentrations over Scandinavia and the Baltic Sea could be an insufficient transport of ozone from the upper troposphere or errors in emission estimates over the Baltic Sea and the North Sea. These hypotheses require further investigation.

5. Largest systematic differences between the GEM-AQ model performances for different types of stations were from January to March and in December, when ozone levels depend to a large extent on dynamical factors. During the summer months, when photochemical production dominates, the model performance for different types of stations is comparable.

6. In regions where the monitoring network is sufficiently dense, errors in emission inventories can be linked to errors in modelling results with greater confidence. Systematic overestimation of PM_{10} concentrations indicates that over Germany and the Czech Republic anthropogenic emissions are overestimated. Limited availability of PM_{10} measurements in Scandinavia and lack of data for France, Italy and the Eastern European countries does not allow evaluation of emission uncertainties in these regions. The overestimation of ozone concentrations in industrial regions of Western Europe in spring and autumn indicates high uncertainty of $\mathrm{NO_x}$ emission estimation in this region, at least in terms of annual temporal variability.

7. Seasonal differences between distribution patterns of the concentration fields relate to large scale dynamics. Elevated ozone concentrations during autumn and winter are calculated for mountain regions as a result of the transport from the upper troposphere in the model. In spring and summer photochemical production dominates and highest concentrations are calculated over regions characterized with highest emissions. PM_{10} concentration patterns correspond to the distribution of anthropogenic emissions. Also, changes in the wind field that drive natural emissions (sea salt, dust) play an important role.

8. Highest discrepancies between modelled and observed concentrations are for periods characterized with highest concentration levels (ozone – summer, PM_{10} – winter).

9. Modelled ozone distribution patterns over North-Eastern and Eastern Europe show lower concentrations in the summer than in the spring. The reason for this effect might confirm the hypothesis of advective nature of ozone episodes in Eastern Europe and the role of high pressure system blocks during summer months

In conclusion, spatial distribution and seasonal variability of air pollution species depend on regional climate and are strongly modulated by anthropogenic emission fluxes. Although "climatology of air pollution species" may not coincide with climate classifications for Europe, the presented analyses confirm that interpretation of modelling results is enhanced when regional climate characteristics are taken into consideration. Thus, data stratification should be recommended for model analysis and evaluation methodology.

Acknowledgements. This study was sponsored by the Polish Ministry of Science and Higher Education grant 821/N-AQMEII/2010/0 (Identification of atmospheric pollutants transformation mechanisms in regional-scale modelling). JWK was partially supported by the Canadian Foundation for Climate and Atmospheric Sciences.

The AQMEII community (http://aqmeii.jrc.ec.europa.eu/) is kindly acknowledged for providing observational data.

References

Appel, K. W., Chemel, C., Roselle, S. J., Francis, X. V., Hu, R.-M., Sokhi, R. S., Rao, S. T., and Galmarini, S.: Examination of the Community Multiscale Air Quality (CMAQ) model performance over the North American and European domains, Atmos. Environ., 53, 142–155, doi:10.1016/j.atmosenv.2011.11.016, 2012.

Bélair, S., Mailhot, J., Girard, C., and Vaillancourt, P.: Boundary Layer and Shallow Cumulus Clouds in a Medium-Range Forecast of a Large-Scale Weather System, Mon. Weather Rev., 133, 1938–1960, 2005.

Bougeault, P. and Lacarrere, P.: Parameterization of orography-induced turbulence in a mesobeta–scale model, Mon. Weather Rev., 117, 1872–1890, 1989.

Brandt, J., Silver, J. D., Frohn, L. M., Geels, C., Gross, A., Hansen, A. B., Hansen, K. M., Hedegaard, G. B., Skjøth, C. A., Villadsen, H., Zare, A., and Christensen, J. H.: An integrated model study for Europe and North America using the Danish Eulerian Hemispheric Model with focus on intercontinental transport of air pollution, Atmos. Environ., 53, 156–176, doi:10.1016/j.atmosenv.2012.01.011, 2012.

Côté, J., Gravel, S., Méthot, A., Patoine, A., Roch, M., and Staniforth, A.: The operational CMC–MRB Global Environmental Multiscale (GEM) Model. Part I: Design considerations and formulation, Mon. Weather Rev., 126, 1373–1395, 1998a.

Côté, J., Desmarais, J.-G., Gravel, S., Méthot, A., Patoine, A., Roch, M., and Staniforth, A.: The Operational CMC–MRB Global Environmental Multiscale (GEM) Model. Part II: Results, Mon. Weather Rev., 126, 1397–1418, 1998b.

Deardorff, J. W.: Efficient prediction of ground surface temperature and moisture with inclusion of a layer of vegetation, J. Geophys. Res., 83, 1889–1903, 1978.

Dennis, R., Fox, T., Fuentes, M., Gilliland, A., Hanna, S., Hogrefe, C., Irwin, J., Rao, S. T., Scheffe, R., Schere, K., Steyn, D., and Venkatram, A.: A framework for evaluating regional-scale numerical photochemical modeling systems, Environ. Fluid Mech., 10, 471–489, 2010.

Fouquart, Y. and Bonnel, B.: Computations of solar heating of the earth's atmosphere: A new parameterization, Contrib. Atmos. Phys., 53, 35–62, 1980.

Galmarini, S., Bianconi, R., Bellasio, R., and Graziani, G.: Forecasting the consequences of accidental releases of radionuclides in the atmosphere from ensemble dispersion modelling. J. Environ. Radioactiv., 57, 203–219, 2001.

Galmarini, S., Bianconi, R., Klug, W., Addis, R., Andronopoulos, S., Baklanov, A., Bartniki, J., Bartzis, J. C., Bellasio, R., Bompay, F., Buckley, R., Bouzom, M., Champion, H., D'Amours, R., Davakis, E., Eleveld, H., Geertsema, G. T., Glaab, H., Kollax, M., Ilvonen, M., Manning, A., Pechinger, U., Persson, C., Polreich, E., Potemski, S., Prodanova, M., Saltbones, J., Slaper, H., Sofiev, M. A., Syrakov, D., Sørensen, J. H., Auwera, Van der, L., Valkama, I., and Zelazny, R.: Ensemble dispersion forecasting, Part 1: Concept, approach and indicators, Atmos. Environ., 38, 4607–4617, doi:10.1016/j.atmosenv.2004.05.030, 2004.

Galmarini, S., Rao, S. T., and Steyn, D. G.: AQMEII: An international initiative for the evaluation of regional-scale air quality models – Phase 1 Preface, Atmos. Environ., 53, 1–3, 2012.

Garand, L.: Some improvements and complements to the infrared emissivity algorithm including a parameterization of the absorption in the continuum region, J. Atmos. Sci., 40, 230–244, 1983.

Gauthier, P., Charette, C., Fillion, L., Koclas, P., and Laroche, S.: Implementation of a 3-D assimilation system at the Canadian Meteorological Centre. Part I: The global analysis, Atmos. Ocean, 37, 103–156, 1999.

Gong, S. L., Barrie, L. A., Blanchet, J.-P., Salzen, K. v., Lohmann, U., Lesins, G., Spacek, L., Zhang, L. M., Girard, E., Lin, H., Leaitch, R., Leighton, H., Chylek, P., and Huang, P.: Canadian aerosol module: A size-segregated simulation of atmospheric aerosol processes for climate and air quality models 1. Module development, J. Geophys. Res., 108, 4007, doi:10.1029/2001JD002002, 2003.

Gong, S. L., Lavoué, D., Zhao, T. L., Huang, P., and Kaminski, J. W.: GEM-AQ/EC, an on-line global multi-scale chemical weather modelling system: model development and evaluation of global aerosol climatology, Atmos. Chem. Phys., 12, 8237–8256, doi:10.5194/acp-12-8237-2012, 2012.

Hogrefe, C., Roselle, S., Mathur, R., Rao, S. T., Galmarini, S.: Space-time analysis of the Air Quality Model Evaluation International Initiative (AQMEII) Phase 1 air quality simulations, J. Air Waste Manage., 64, 388–405, doi:10.1080/10962247.2013.811127, 2013.

Kain, J. S. and Fritsch, J. M.: A one-dimensional entraining/detraining plume model and its application in convective parameterization, J. Atmos. Sci., 47, 2784–2802, 1990.

Kain, J. S. and Fritsch, J. M.: Convective parameterization for mesoscale models: The Kain–Fritsch scheme. The representation of cumulus convection in numerical models. Meteor. Monogr., No. 24, Amer. Meteor. Soc., 165–170, 1993.

Kaminski, J. W., Neary, L., Struzewska, J., McConnell, J. C., Lupu, A., Jarosz, J., Toyota, K., Gong, S. L., Côté, J., Liu, X., Chance, K., and Richter, A.: GEM-AQ, an on-line global multiscale chemical weather modelling system: model description and evaluation of gas phase chemistry processes, Atmos. Chem. Phys., 8, 3255–3281, doi:10.5194/acp-8-3255-2008, 2008.

Lurmann, F. W., Lloyd, A. C., and Atkinson, R.: A chemical mechanism for use in long-range transport/acid deposition computer modeling, J. Geophys. Res., 91, 10905–10936, 1986.

Mailhot, J., Bélair, S., Lefaivre, L., Bilodeau, B., Desgagné, M., Girard, C., Glazer, A., Leduc, A., Méthot, A., Patoine, A., Plante, A., Rahill, A., Robinson, T., Talbot, D., Tremblay, A., Vaillancourt, P., Zadra, A., and Qaddouri, A.: The 15 km version of the Canadian regional forecast system, Atmos. Ocean, 44, 133–149, 2006.

Marticorena, B. and Bergametti, G.: Modeling the atmospheric dust cycle: 1. Design of a soil-derived dust emission scheme, J. Geophys. Res., 100(D8), doi:10.1029/95JD00690, 16415–16430, 1995.

McFarlane, N. A.: The effect of orographically excited gravity wave drag on the general circulation of the lower stratosphere and troposphere, J. Atmos. Sci., 44, 1775–1800, 1987.

McLandress, C. and McFarlane, N. A.: Interactions between orographic gravity wave drag and forced stationary planetary waves in the winter Northern Hemisphere middle atmosphere, J. Atmos. Sci., 50, 1966–1990, 1993.

Nopmongcol, U., Koo, B., Tai, E., Jung, J., Piyachaturawat, P., Emery, C., Yarwood, G., Pirovano, G., Mitsakou, C., and Kallos, G.: Modeling europe with CAMx for the Air Quality Model Evaluation International Initiative (AQMEII), Atmos. Environ., 53, 177–185, doi:10.1016/j.atmosenv.2011.11.023, 2012.

Olivier, J. G. J. and Berdowski, J. J. M.: Global emissions sources and sinks, in: The Climate System, edited by: Guicherit, R., and Heij, B., 33–78, A.A. Balkema Publishers/Swets and Zeitlinger Publishers, Lisse, the Netherlands, 2001.

Olivier, J. G. J., Bouwman, A. F., Berdowski, J. J. M., Veldt, C., Bloos, J. P. J., Visschedijk, A. J. H., van der Maas, C. W. M., and Zandveld, P. Y. J.: Sectoral emission inventories of green-house gases for 1990 on a per country basis as well as on $1^\circ \times 1^\circ$, Environ. Sci. Policy, 2, 241–263, 1999.

Pirovano, G., Balzarini, A., Bessagnet, B., Emery, C., Kallos, G., Meleux, F., Mitsakou, C., Nopmongcol, U., Riva, G. M., and Yarwood, G.: Investigating impacts of chemistry and transport model formulation on model performance at European scale, Atmos. Environ., 53, 93–109, doi:10.1016/j.atmosenv.2011.12.052, 2012.

Pouliot, G., Pierce, T., Denier van der Gon, H., Schaap, M., Moran, M., and Nopmongcol, U.: Comparing emission inventories and model-ready emission datasets between Europe and North America for the AQMEII project, Atmos. Environ., 53, 4–14, 2012.

Putaud, J. P., Van Dingenen, R., Alastuey, A., Bauer, H., Birmili, W., Cyrys, J., Flentje, H., Fuzzi, S., Gehrig, R., Hansson, H. C., Harrison, R. M., Herrmann, H., Hitzenberger, R., Hüglin, C., Jones, A. M., Kasper-Giebl, A., Kiss, G., Kousa, A., Kuhlbusch, T. A. J., Löschau, G., Maenhaut, W., Molnar, A., Moreno, T., Pekkanen, J., Perrino, C., Pitz, M., Puxbaum, H., Querol, X., Rodriguez, S., Salma, I., Schwarz, J., Smolik, J., Schneider, J., Spindler, G., ten Brink, H., Tursic, J., Viana, M., Wiedensohler, A., and Raes, F.: A European aerosol phenomenology – 3: physical and chemical characteristics of particulate matter from 60 rural, urban, and kerbside sites across Europe. Atmos. Environ., 44, 1352e2310, doi:10.1016/j.atmosenv.2009.12.011, 2010.

Robert, A., Yee, T. L. and Ritchie, H.: A Semi-Lagrangian and Semi-Implicit Numerical Integration Scheme for Multilevel Atmospheric Models. Mon. Wea. Rev., 113, 388–394, doi:, 1985.

Solazzo, E., Riccio, A., Kioutsioukis, I., and Galmarini, S.: Pauci ex tanto numero: reduce redundancy in multi-model ensembles, Atmos. Chem. Phys., 13, 8315–8333, doi:10.5194/acp-13-8315-2013, 2013.

Solazzo, E., Bianconi, R., Pirovano, G., Matthias, V., Vautard, R., Moran, M. D., Appel, K. W., Bessagnet, B., Brandt, J., Christensen, J. H., Chemel, C., Coll, I., Ferreira, J., Forkel, R., Francis, X. V., Grell, G., Grossi, P., Hansen, A. B., Miranda, A. I., Nopmongcol, U., Prank, M., Sartelet, K. N., Schaap, M., Silver, J. D., Sokhi, R. S., Vira, J., Werhahn, J., Wolke, R., Yarwood, G., Zhang, J., Rao, S. T., and Galmarini, S.: Operational model evaluation for particulate matter in Europe and North America in the context of AQMEII, Atmos. Environ., 53, 75–92, doi:10.1016/j.atmosenv.2012.02.045, 2012a.

Solazzo, E., Bianconi, R., Vautard, R., Appel, K. W., Moran, M. D., Hogrefe, C., Bessagnet, B., Brandt, J., Christensen, J. H., Chemel, C., Coll, I., Van der Gon, H. D., Ferreira, J., Forkel, R., Francis, X. V., Grell, G., Grossi, P., Hansen, A. B., Jericevic, A., Kraljevic, L., Miranda, A. I., Nopmongcol, U., Pirovano, G., Prank, M., Riccio, A., Sartelet, K. N., Schaap, M., Silver, J. D., Sokhi, R. S., Vira, J., Werhahn, J., Wolke, R., Yarwood, G., Zhang, J., Rao, S. T., and Galmarini, S.: Model evaluation and ensemble modelling of surface-level ozone in Europe and North America in the context of AQMEII, Atmos. Environ., 53, 60–74, 2012b.

Struzewska, J. and Kaminski, J. W.: Formation and transport of photooxidants over Europe during the July 2006 heat wave – observations and GEM-AQ model simulations. Atmos. Chem. Phys., 8, 721–736, doi: , 2008.

Sundqvist, H.: A parameterization scheme for non-convective condensation including prediction of cloud water content, Q. J. Roy. Meteor. Soc., 104, 677–690, 1978.

Tagaris, E., Sotiropoulou, R. E. P., Gounaris, N., Andronopoulos, S., and Vlachogiannis, D.: Air quality over Europe: modelling gaseous and particulate pollutants, Atmos. Chem. Phys., 13, 9661–9673, doi:10.5194/acp-13-9661-2013, 2013.

Venkatram, A., Karamchandani, P. K., and Misra, P. K.: Testing a comprehensive acid deposition model, Atmos. Environ., 22, 737–747, 1988.

Table 1: Error statistics for GEM-AQ model performance – maximum 8 h-running average of ozone concentrations in $\mu g\,m^{-3}$.

	MBE	MAGE	Correlation
All stations	−0.9	17.6	0.93
Seasonal:			
Winter	−0.8	17.6	0.71
Spring	0.3	18.0	0.66
Summer	−6.7	17.7	0.96
Autumn	3.6	17.1	0.93
Station type			
Rural	−7.0	18.2	0.92
Suburban	1.9	17.2	0.93
Urban	2.5	17.4	0.93
Regions			
Region I	−16.8	20.7	0.83
Region II	0.4	16.5	0.91
Region III	−1.5	17.6	0.89
Region IV	−1.4	19.9	0.96

Table 2: Error statistics for GEM-AQ model performance – 24 h-running average of PM_{10} concentrations in $\mu g\,m^{-3}$.

	MBE	MAGE	Correlation
All stations	4.3	18.3	0.74
Seasonal:			
Winter	3.4	24.4	0.74
Spring	0.2	15.0	0.71
Summer	6.3	15.7	0.81
Autumn	7.4	18.3	0.8
Station type			
Rural	10.8	18.2	0.72
Suburban	4.7	18.4	0.75
Urban	1.3	18.3	0.72
Regions			
Region I	0.6	10.3	0.32
Region II	8.9	16.7	0.73
Region III	5.9	24.2	0.72
Region IV	−9.4	16.2	0.56

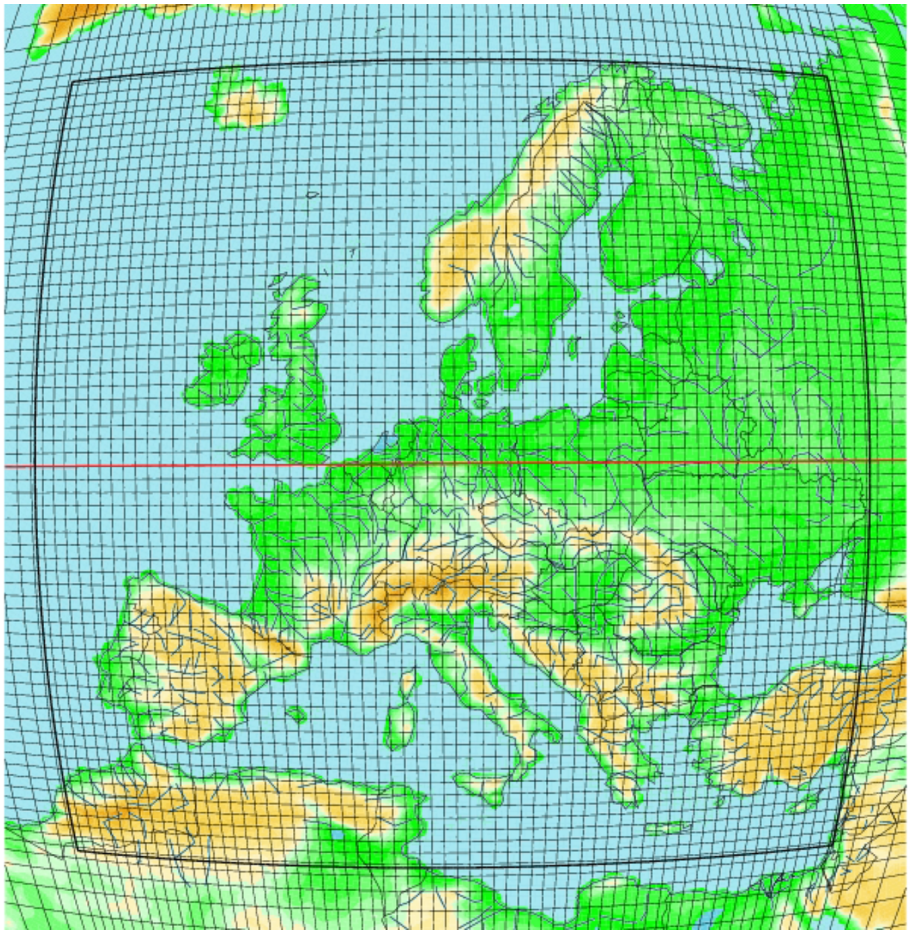


Figure 1: GEM-AQ computational domain configuration. Global variable grid with rotated equator (red line). Thicker black line borders the central region with the resolution of $0.2^\circ \times 0.2^\circ$.

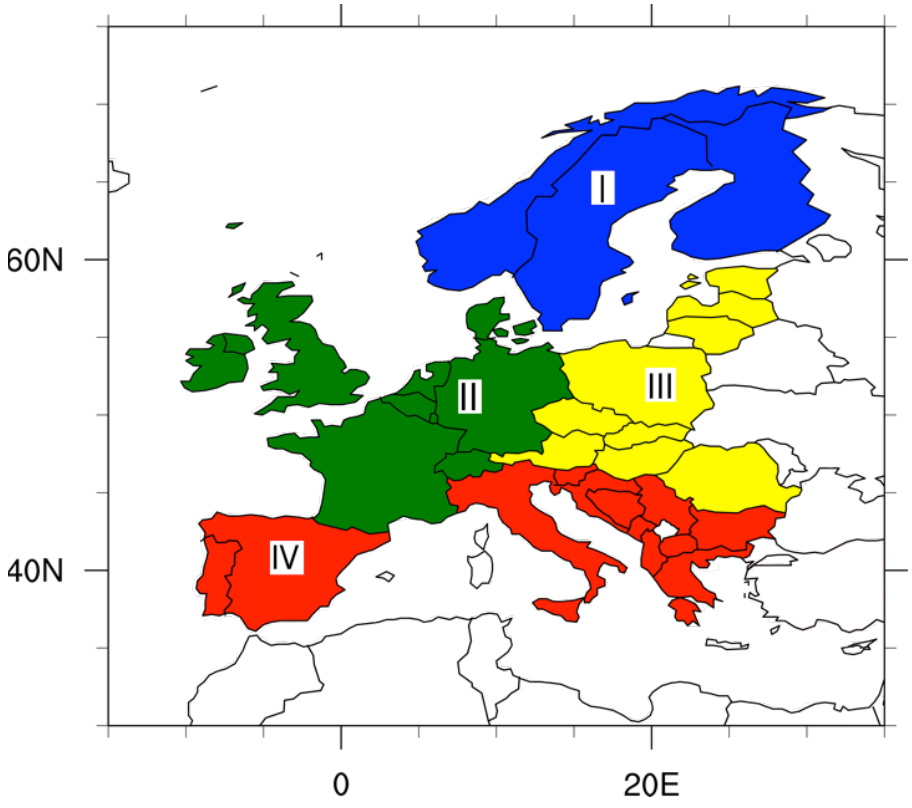


Figure 2: Regions selected for analysis of the GEM-AQ model results.

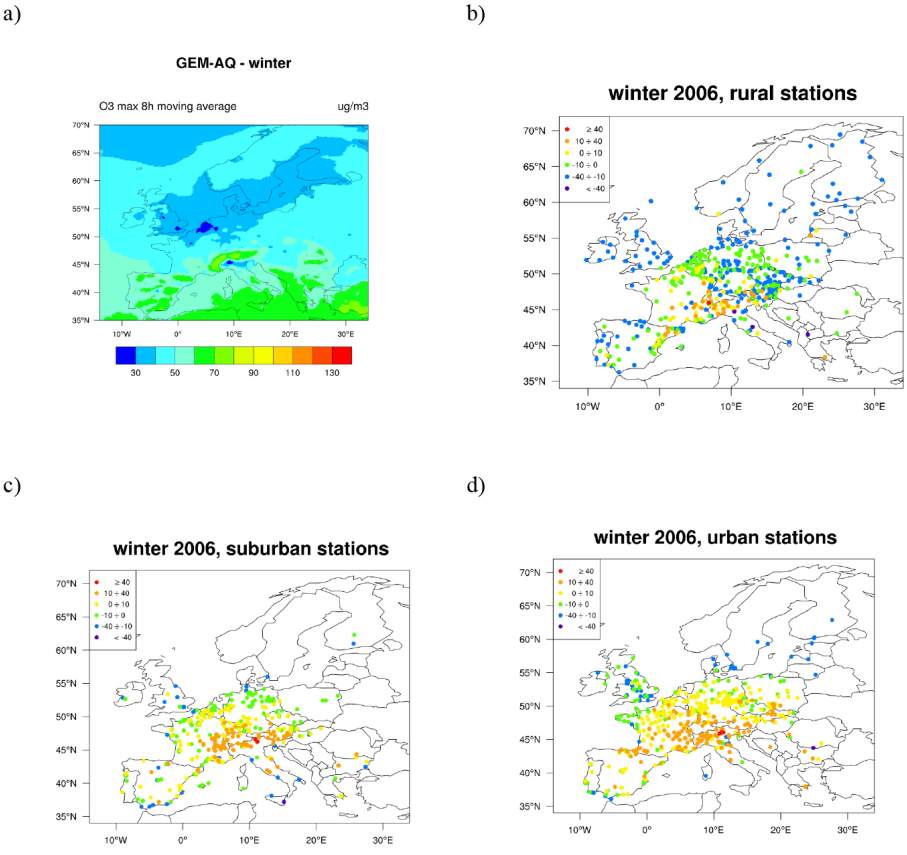
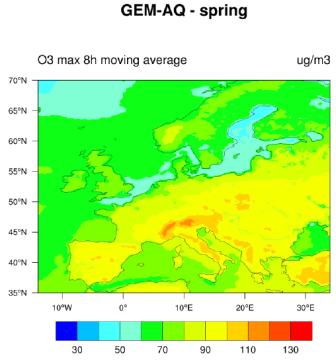
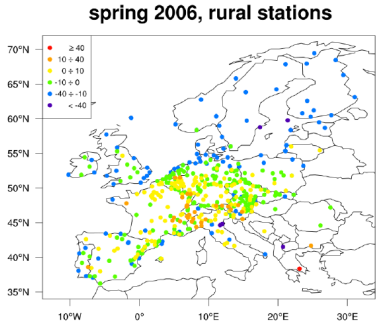


Figure 3: Winter (DJF) maximum 8 h running average ozone concentration in $\mu g\,m^{-3}$ **(a)**, MBE at rural stations **(b)**, MBE at suburban stations **(c)** MBE at urban stations **(d)**.

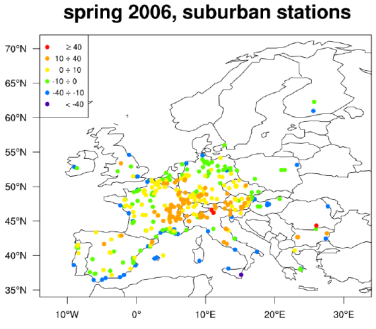
a)



b)



c)



d)

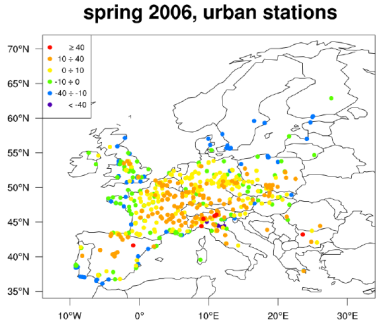
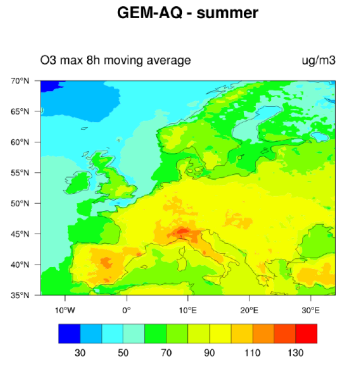
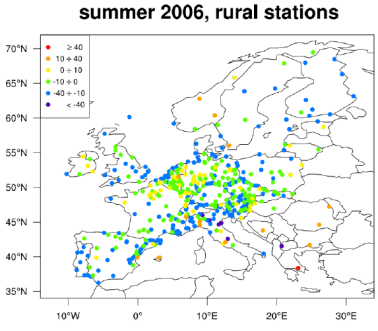


Figure 4: Spring (MAM) maximum 8 h running average ozone concentration in $\mu g\,m^{-3}$ **(a)**, MBE at rural stations **(b)**, MBE at suburban stations **(c)** MBE at urban stations **(d)**.

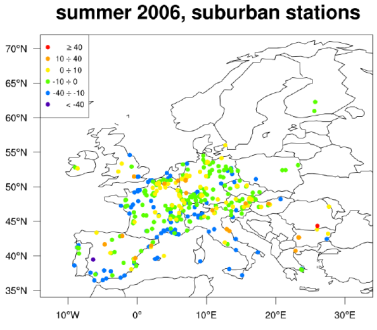
a)



b)



c)



d)

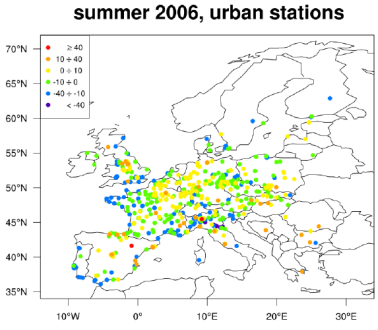
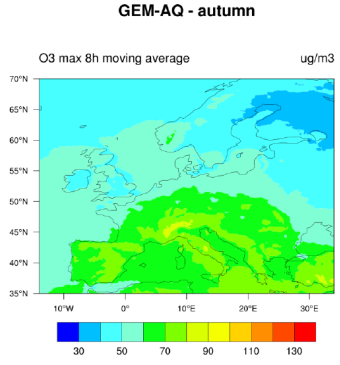
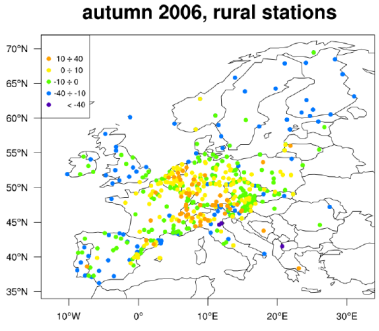


Figure 5: Summer (JJA) maximum 8 h running average ozone concentration in $\mu g\,m^{-3}$ **(a)**, MBE at rural stations **(b)**, MBE at suburban stations **(c)** MBE at urban stations **(d)**.

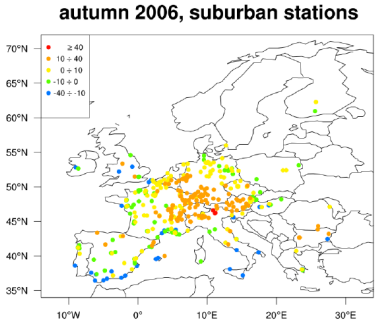
a)



b)



c)



d)

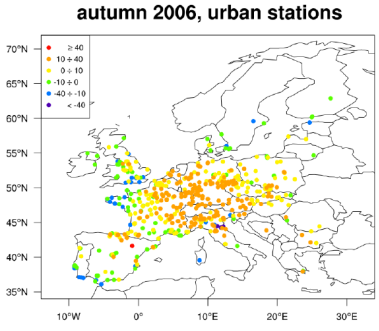


Figure 6: Autumn (SON) maximum 8 h running average ozone concentration in $\mu g\,m^{-3}$ **(a)**, MBE at rural stations **(b)**, MBE at suburban stations **(c)** MBE at urban stations **(d)**.

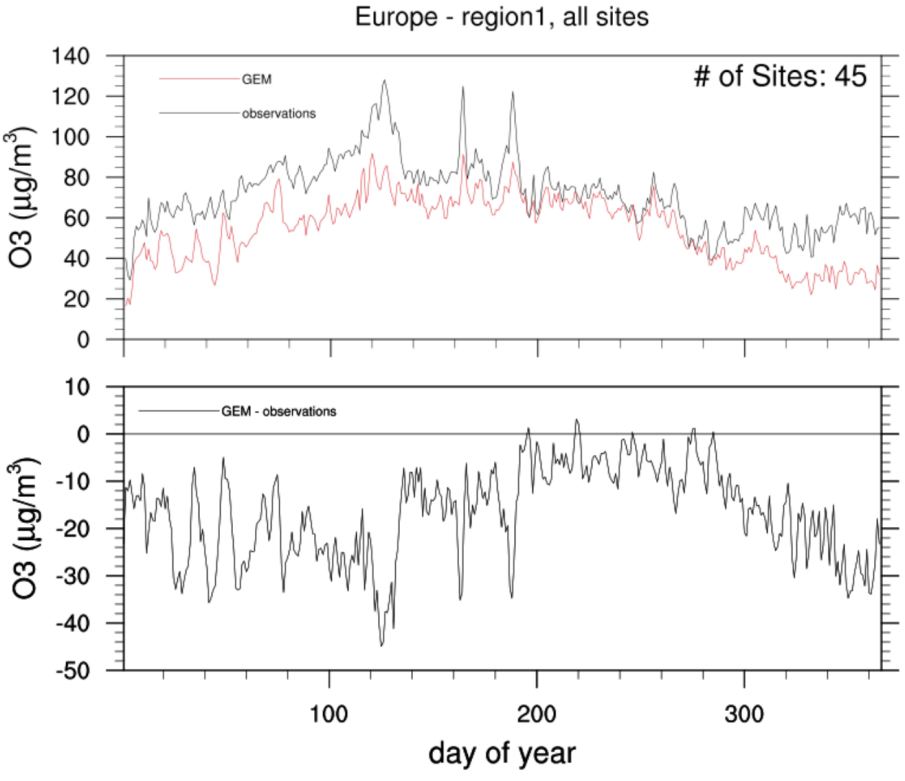


Figure 7: Time series of observed and modelled maximum 8 h running averaged ozone concentration **(a)** averaged for all stations in Northern Europe, **(b)** MBE.

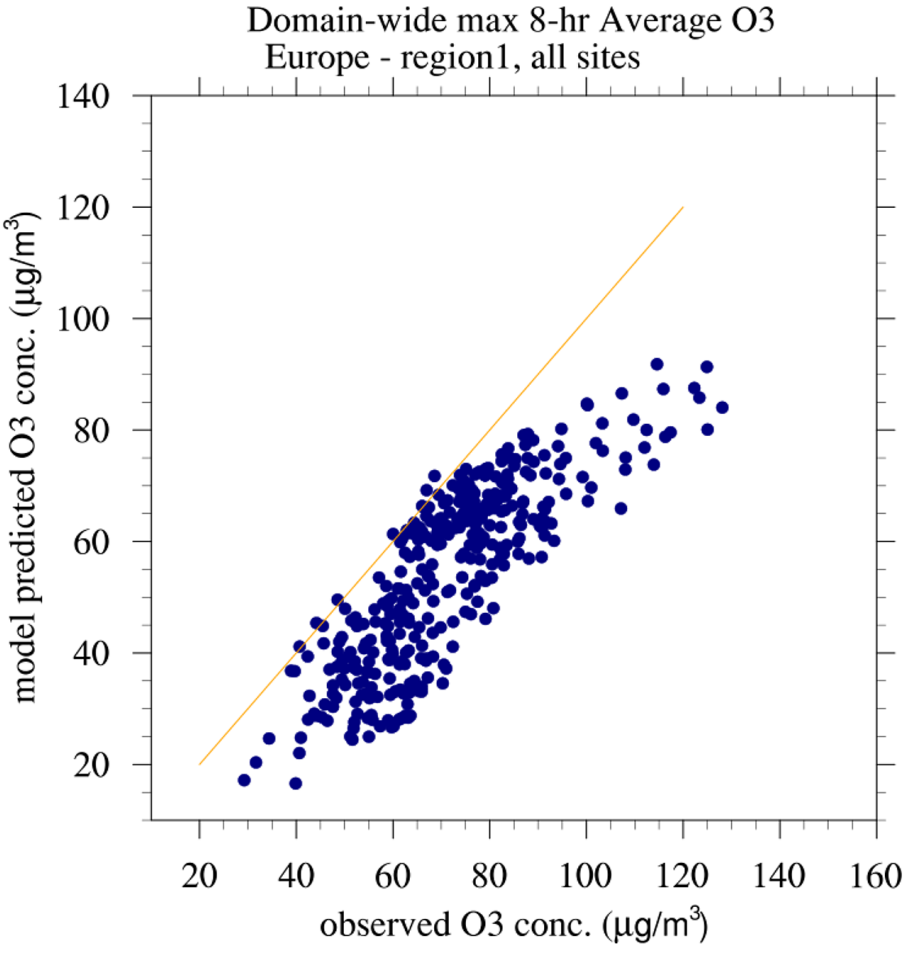


Figure 8: Scatterplot of maximum 8 h running average ozone concentration in Northern Europe.

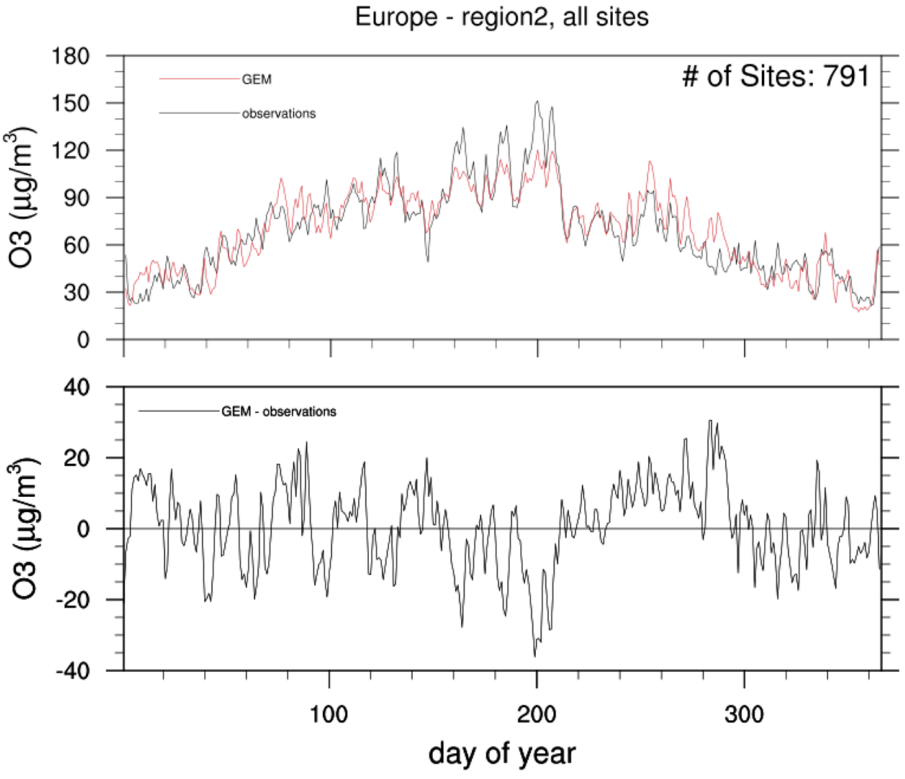


Figure 9: Time series of observed and modelled maximum 8 h running averaged ozone concentration **(a)** averaged for all stations in Western Europe, **(b)** MBE.

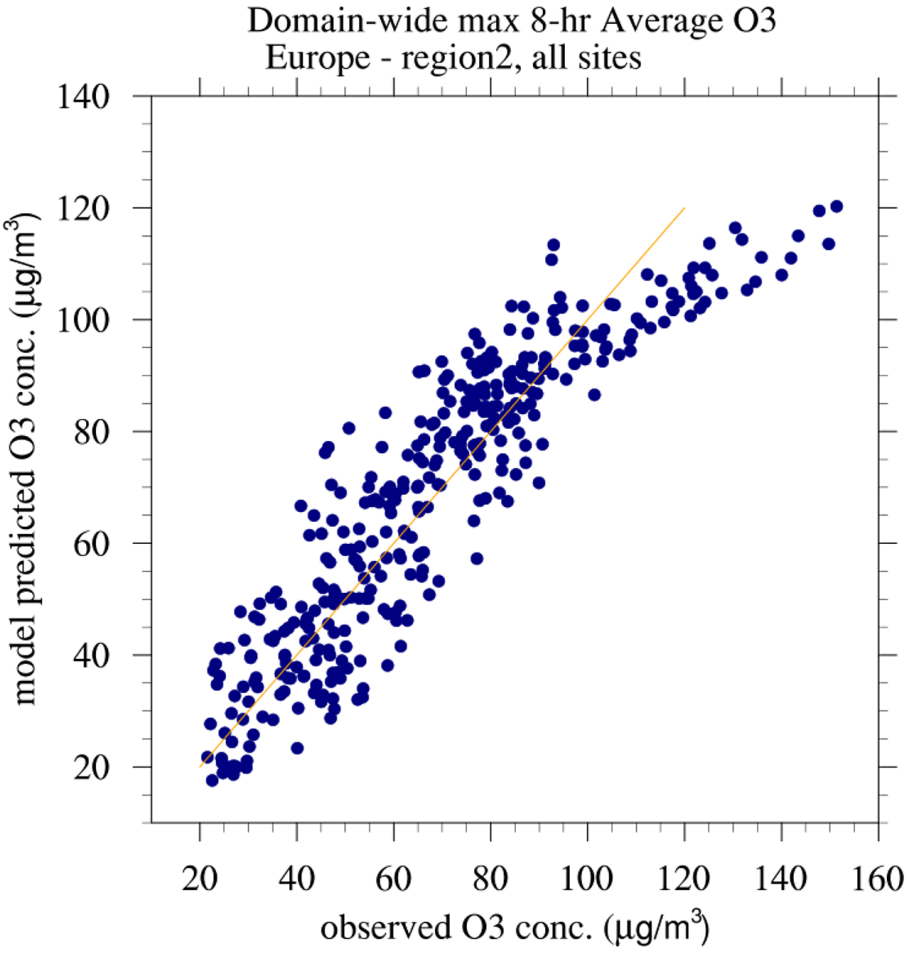


Figure 10: Scatterplot of maximum 8 h running average ozone concentration in Western Europe.

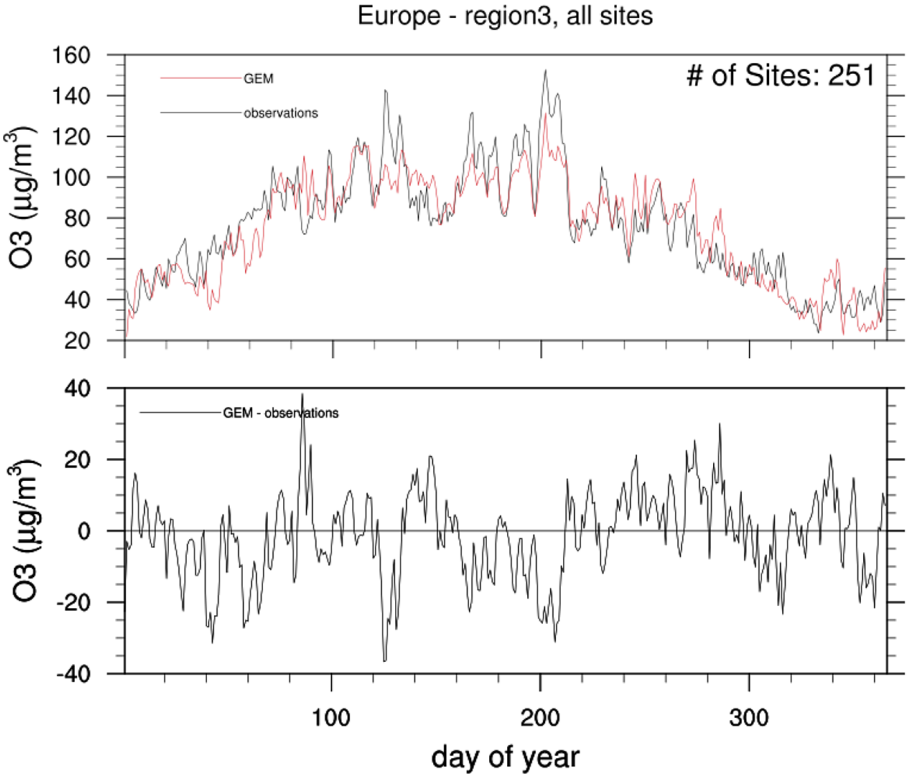


Figure 11: Time series of observed and modelled maximum 8 h running averaged ozone concentration **(a)** averaged for all stations in Central Europe, **(b)** MBE.

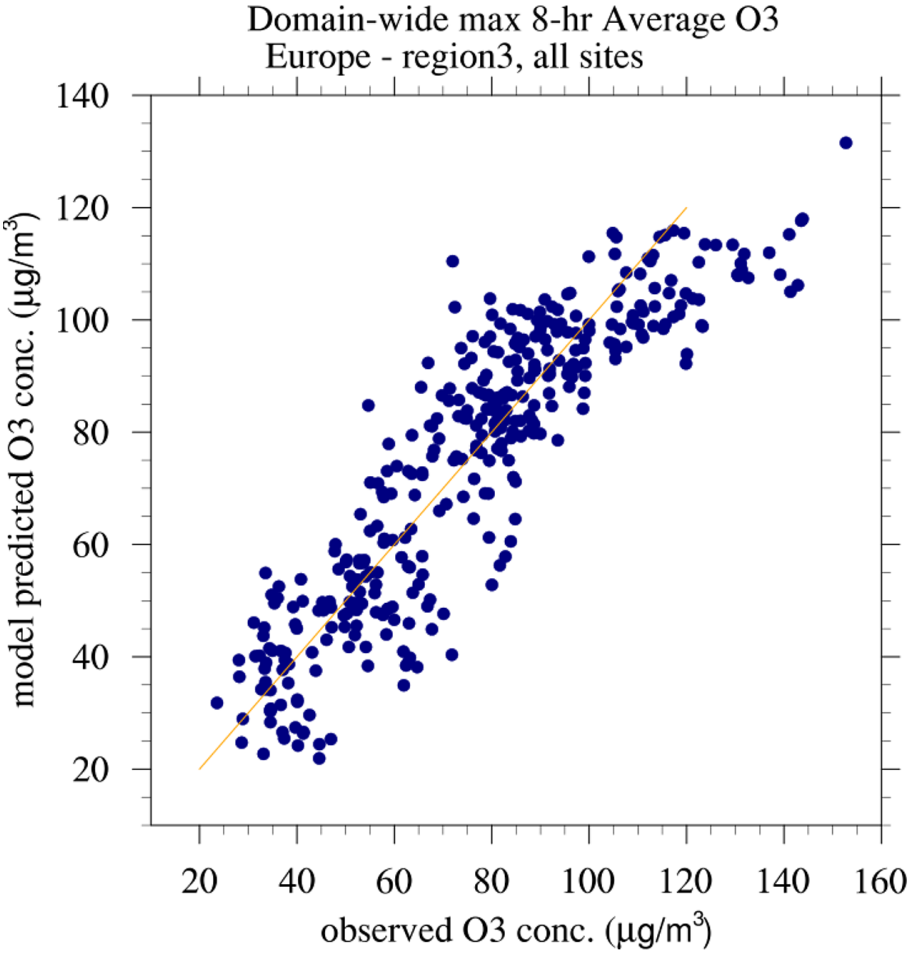


Figure 12: Scatterplot of maximum 8 h running average ozone concentration in Central Europe.

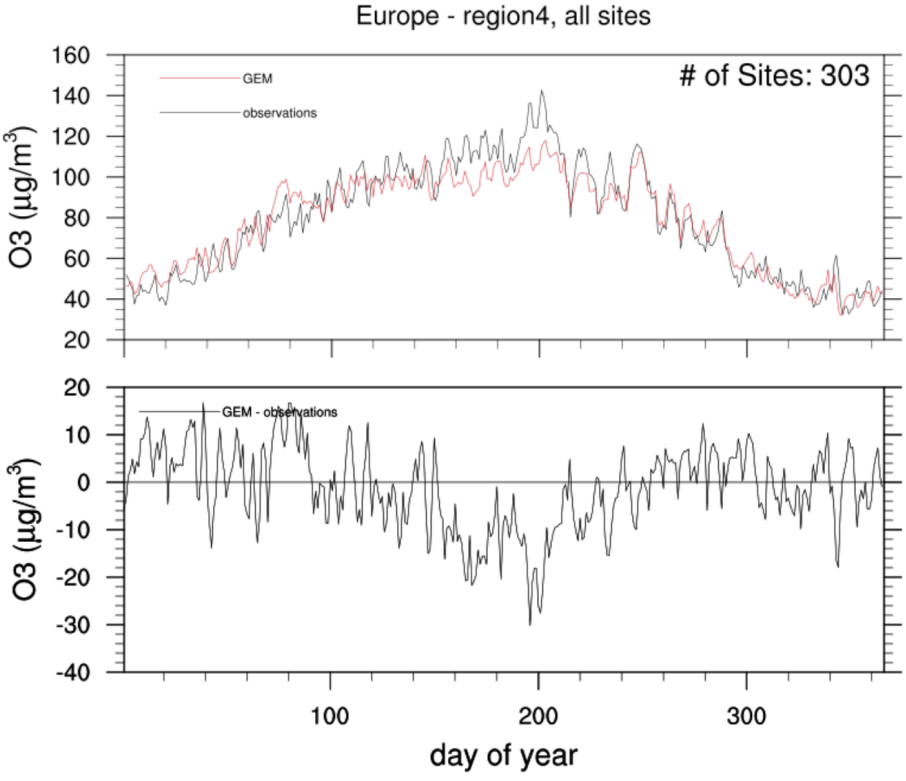


Figure 13: Time series of observed and modelled maximum 8 h running averaged ozone concentration **(a)** averaged for all stations in Southern Europe, **(b)** MBE.

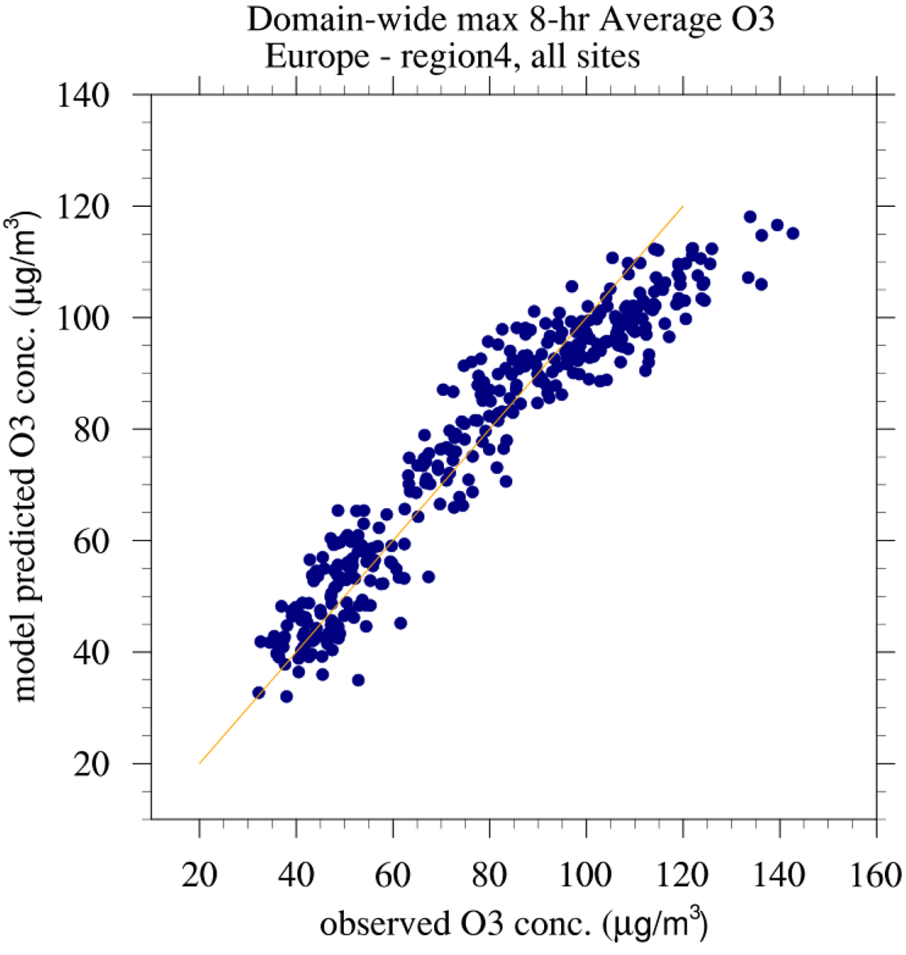
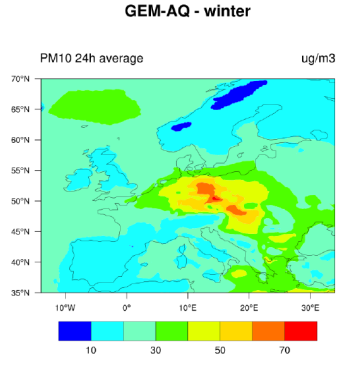
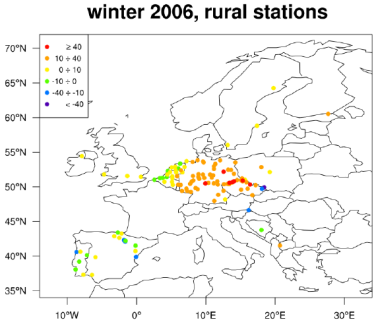


Figure 14: Scatterplot of maximum 8 h running average ozone concentration in Central Europe.

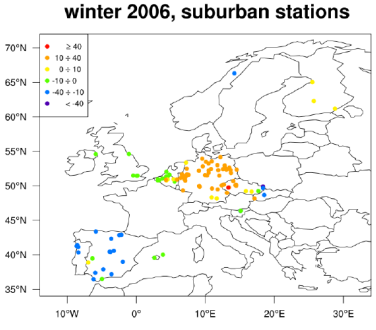
a)



b)



c)



d)

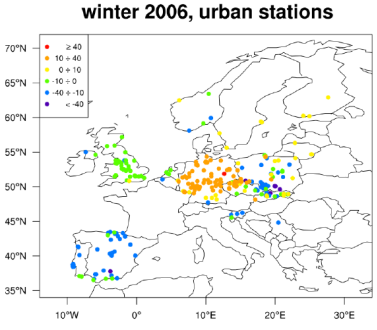
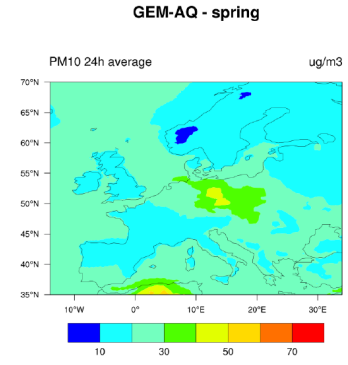
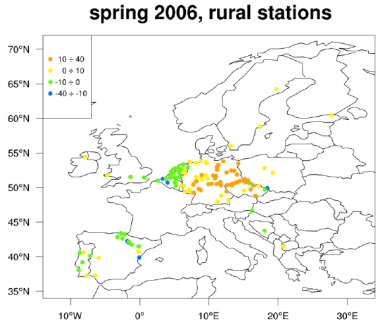


Figure 15: Winter (DJF) 24 h average PM_{10} concentrations in $\mu\mathrm{g\ m}^{-3}$ **(a)**, MBE at rural stations **(b)**, MBE at suburban stations **(c)** MBE at urban stations **(d)**.

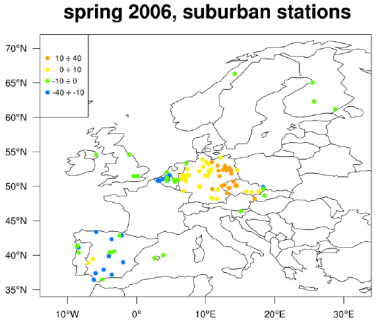
a)



b)



c)



d)

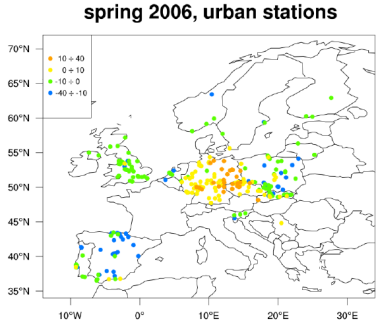
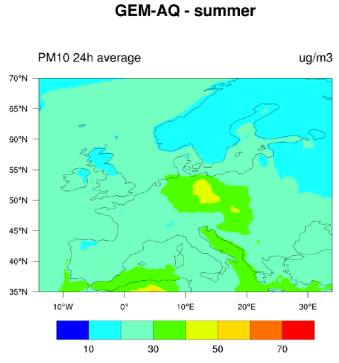
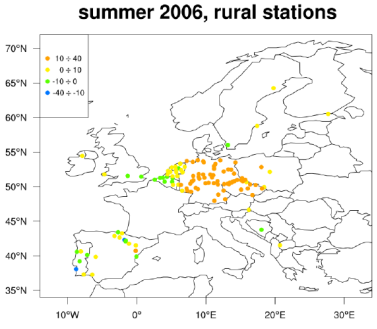


Figure 16: Spring (MAM) 24 h average PM_{10} concentrations in $\mu g\,m^{-3}$ **(a)**, MBE at rural stations **(b)**, MBE at suburban stations **(c)** MBE at urban stations **(d)**.

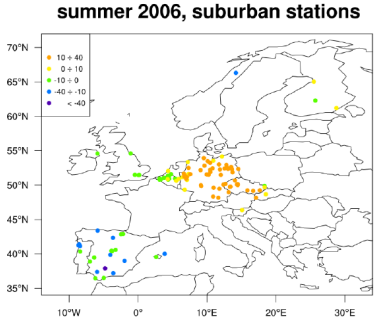
a)



b)



c)



d)

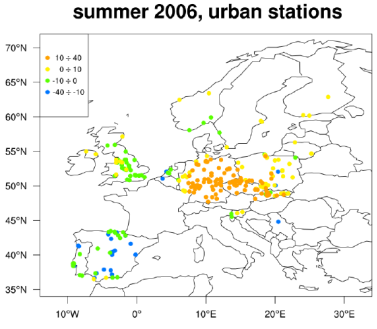
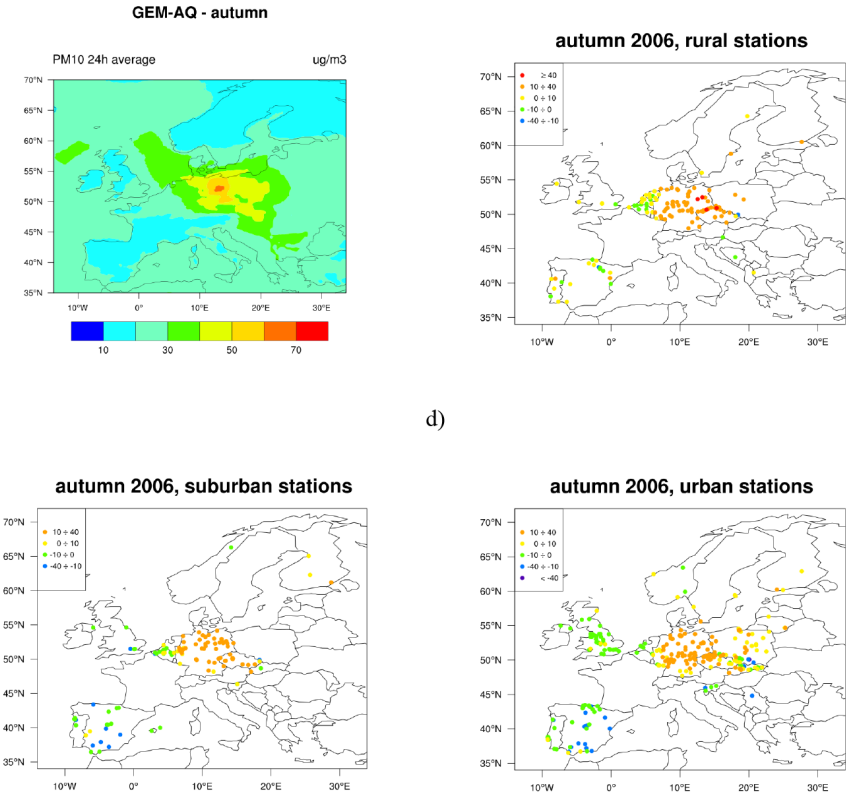


Figure 17: Summer (JJA) 24 h average PM_{10} concentrations in $\mu\mathrm{g\,m}^{-3}$ **(a)**, MBE at rural stations **(b)**, MBE at suburban stations **(c)** MBE at urban stations **(d)**.

a) b)



c) d)

Figure 18: Autumn (SON) 24 h average PM_{10} concentrations in $\mu g\,m^{-3}$ **(a)**, MBE at rural stations **(b)**, MBE at suburban stations **(c)** MBE at urban stations **(d)**.

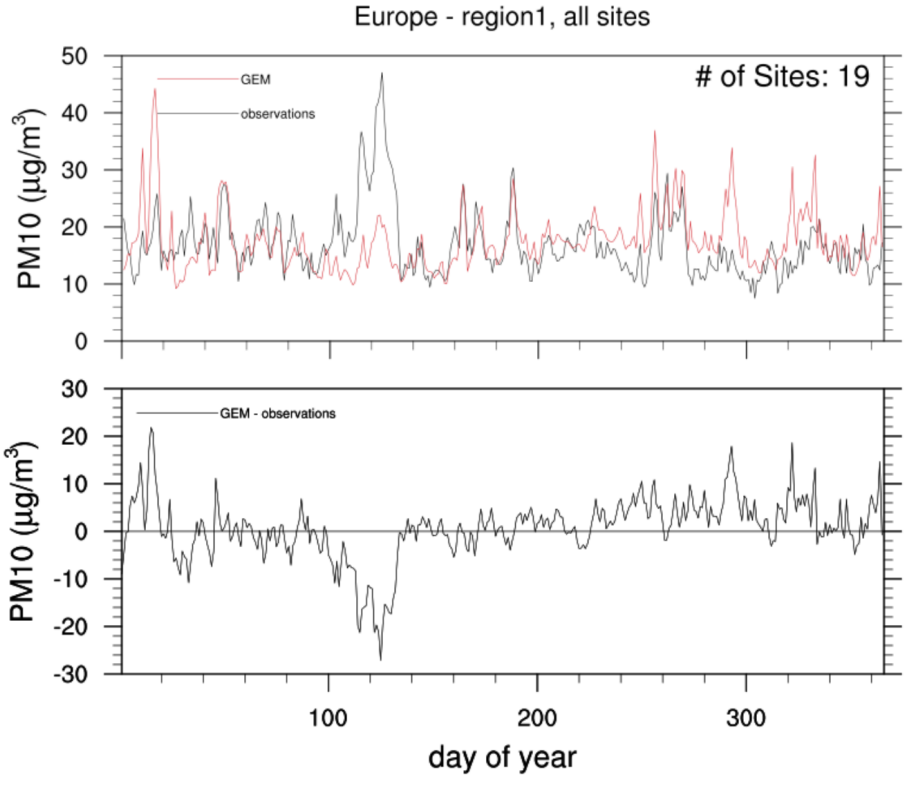


Figure 19: Time series of observed and modelled PM_{10} 24 h concentration averaged for all stations in Northern Europe and MBE.

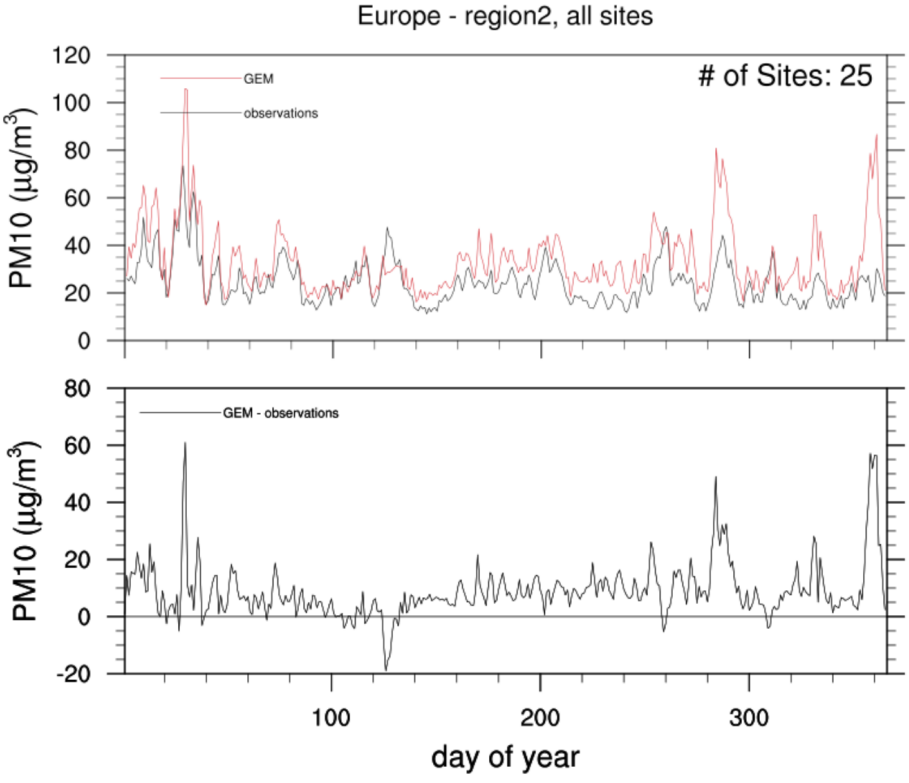


Figure 20: Time series of observed and modelled PM_{10} 24 h concentration averaged for all stations in Western Europe and MBE.

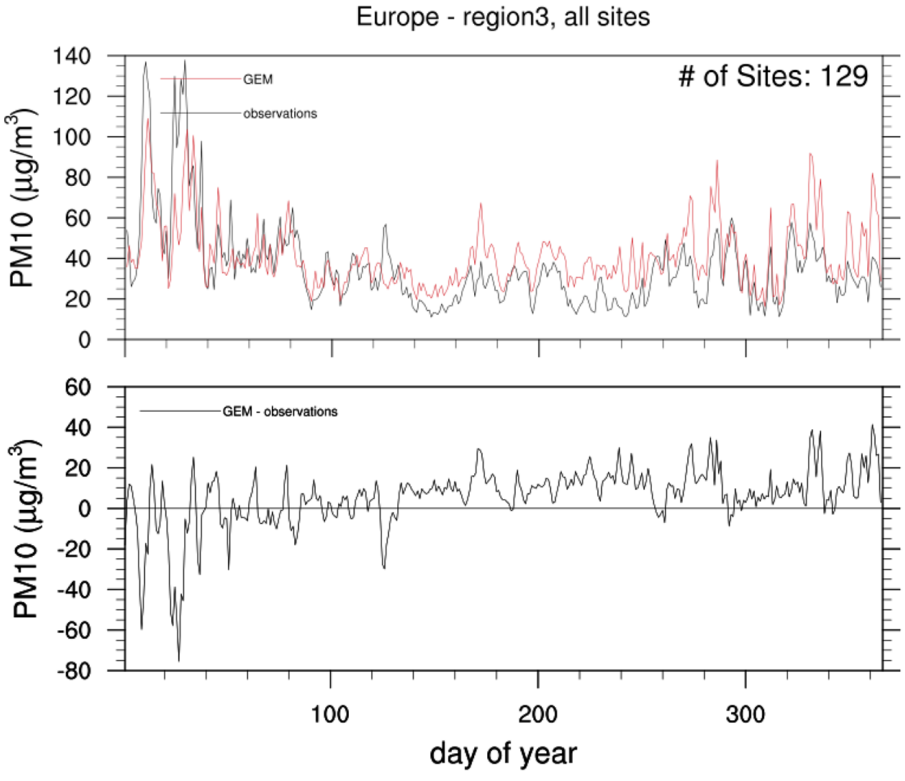


Figure 21: Time series of observed and modelled PM_{10} 24 h concentration **(a)** averaged for all stations in Central Europe, **(b)** MBE.

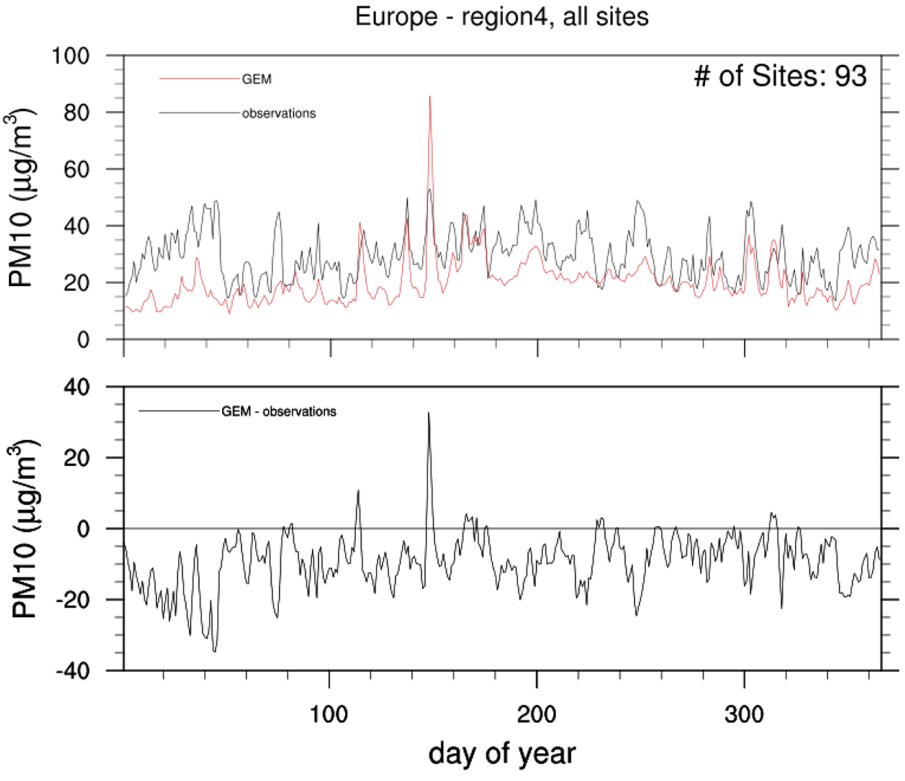


Figure 22: Time series of observed and modelled PM_{10} 24 h concentration **(a)** averaged for all stations in Southern Europe, **(b)** MBE.