Performance of European chemistry transport models as function of horizontal resolution

M. Schaap[1,*], C. Cuvelier[2], C. Hendriks[1], B. Bessagnet[3], J. Baldasano[4], A. Colette[3], P. Thunis[2], D. Karam[2], H. Fagerli[5], A. Graff[6], R. Kranenburg[1], A. Nyiri[5], M.T. Pay[4], L. Rouïl[3], M. Schulz[5], D. Simpson[5,7], R. Stern[8], E. Terrenoire[3], P. Wind[5]

[1] TNO, Dept. of Climate, Air and Sustainability, P.O. Box 80015, NL-3508TA Utrecht, The Netherlands

[2] European Commission, DG Joint Research Centre, Institute for Environment and Sustainability, I-21020 Ispra (Va), Italy

[3] INERIS, Institut National de l'Environnement Industriel et des Risques, Parc Technologique, ALATA, F-60550 Verneuil-en-Halatte, France

[4] Earth Science Department, Barcelona Supercomputing Center – Centro Nacional de Supercomutación, c/ Jordi Girona 29, E-08034 Barcelona, Spain

[5] EMEP MSC-W, Norwegian Meteorological Institute, P.O. Box 43, Blindern, N-0313, Oslo, Norway

[6] Umweltbundesamt, Postfach 1406, D-06813 Dessau-Roßlau, Germany

[7] Dept. Earth & Space Sciences, Chalmers University of Technology, SE412 96, Gothenburg, Sweden

[8] Freie Universität Berlin, Institut für Meteorologie und Troposphärische Umweltforschung, Carl-Heinrich-Becker Weg 6-10, D-12165 Berlin, Germany

[*] Corresponding author:

M. Schaap

TNO

Dept. of Climate, Air and Sustainability

P.O.Box 80015, 3508 TA Utrecht

The Netherlands

Tel: +31 (0) 88 8662074

Mob: +31 (0) 6 11783060

Fax: +31 (0) 88 8662044

martijn.schaap@tno.nl

Abstract

Air pollution causes adverse effects on human health as well as ecosystems and crop yield and also has an impact on climate change trough short-lived climate forcers. To design mitigation strategies for air pollution, 3D Chemistry Transport Models (CTMs) have been developed to support the decision process. Increases in model resolution may provide more accurate and detailed information, but will cubically increase computational costs and pose additional challenges concerning high resolution input data. The motivation for the present study was therefore to explore the impact of using finer horizontal grid resolution for policy support applications of the European Monitoring and Evaluation Programme (EMEP) model within the Long Range Transboundary Air Pollution (LRTAP) convention. The goal was to determine the "optimum resolution" at which additional computational efforts do not provide increased model performance using presently available input data. Five regional CTMs performed four runs for 2009 over Europe at different horizontal resolutions.

The models' responses to an increase in resolution are broadly consistent for all models. The largest response was found for NO_2 followed by PM_{10} and O_3. Model resolution does not impact model performance for rural background conditions. However, increasing model resolution improves the model performance at stations in and near large conglomerations. The statistical evaluation showed that the increased resolution better reproduces the spatial gradients in pollution regimes, but does not help to improve significantly the model performance for reproducing observed temporal variability. This study clearly shows that increasing model resolution is advantageous, and that leaving a resolution of 50 km in favour of a resolution between 10 and 20 km is practical and worthwhile. As about 70% of the model response to grid resolution is determined by the difference in the spatial emission distribution, improved emission allocation procedures at high spatial and temporal resolution are a crucial factor for further model resolution improvements.

Keywords:

Air quality; Chemistry transport models; Resolution; Model comparison; Air Pollution; Model Evaluation; EMEP; CHIMERE; LOTOS-EUROS; CMAQ; Particulate Matter; Ozone; Nitrogen Oxides;

1. Introduction

Air pollution is associated with adverse effects on human health through population exposure to particulate matter and ozone (Dockery et al., 1993; Bell et al., 2004), loss of biodiversity through acidification and eutrophication (Bobbink et al., 1998), decreased crop yields (Adams et al., 1982; Mills et al., 2011) as well as climate change through interactions of short-lived climate forcers with the earth's radiation balance and carbon and nitrogen cycles (Ainsworth et al., 2012, Kiehl and Briegleb, 1993, Simpson et al., 2014b). Air pollutants like ozone (O_3), particulate matter (PM) and nitrogen oxides (NO_x) play a key role in several of these issues. To guide the design of mitigation strategies 3D chemistry transport models (CTMs) have been developed (e.g., Bessagnet et al., 2004; Schaap et al., 2008; Byun and Schere, 2006; Simpson et al., 2012). For this application CTMs should accurately predict the concentration distributions and temporal variability of pollutants, as well as the response of these concentration distributions to emission changes due to mitigation options. As cheap mitigation options have already been implemented in Europe, further mitigation strategies are anticipated to become increasingly expensive (Wagner et al., 2000). Hence, the quality of the underpinning data and models to develop cost-effective mitigation strategies needs to be as high as possible.

The European Monitoring and Evaluation Programme Meteorological Synthesizing Centre - West (EMEP MSC-W) models (Eliassen and Saltbones, 1983; Berge and Jakobsen, 1998; Simpson et al., 2012) have been instrumental to the development of air quality policies in Europe since the late 1970s. In the 1990s the EMEP models became the reference tools for atmospheric dispersion calculations as input to Integrated Assessment Models RAINS (Regional Air Pollution Information and Simulation) and GAINS (Greenhouse Gas and Air Pollution Interactions and Synergies) (Schöpp et al., 1999; Reis et al., 2012), which support the development of air quality polices in the European Union. From 1999 until 2012, the EMEP model has been run on a resolution of 50x50 km^2 for policy support purposes. Partly as a result of the work to be presented here, starting 2013 a grid size of 28 km is used for source-receptor calculations. However, rapid computational technology advancements in the past two decades have enabled CTM applications at even higher resolutions. In Europe, Eulerian CTMs currently use resolutions between 12-25 km for operational European wide applications (e.g., Pay et al., 2010; Zhang et al., 2012; Mues et al., 2013), 4-10 km for application to a single country (e.g., Vieno et al., 2010; Baldasano et al., 2011; Hendriks et al. 2013) and reaching 1 km for some European regions (Pay et al., 2014). A major motivation for the present study was therefore to assess the impact of using finer grid resolution for policy support applications of the EMEP model within the Long Range Transboundary Air Pollution (LRTAP) convention. As an increase in horizontal model resolution will increase the computational costs cubically and poses additional challenges

concerning high resolution input data and model formulation, it is important to determine the "optimum resolution" at which additional efforts do not pay off in terms of increased model performance for the application at hand.

CTMs are sensitive to grid resolution as the models work with mixing ratios, which are generally assumed constant in a model grid-box. The nonlinearity of photochemistry and aerosol formation raises the possibility that, because variations on the sub-grid scale are not represented, systematic errors occur in CTM chemical budgets (Pyle and Zavody, 1990). Several studies have evaluated the impact of model resolution on ozone production (Jang et al., 1995a,b; Esler et al., 2004; Arunachalam et al., 2006; Cohan et al., 2006; Wild and Prather, 2006; Queen and Zhang, 2008; Valari and Menut, 2008; Tie et al., 2010; Hodnebrog et al., 2011; Lauwaet et al., 2013;), PM and its components (Queen and Zhang, 2008; Stroud et al., 2011; Wolke et al., 2012; Fountoukis et al., 2013), ozone and PM precursors (Valin et al., 2011; Kryza et al., 2012), as well as wet deposition fluxes (Queen and Zhang, 2008; Appel et al., 2011). Hence, most studies in literature have focused on ozone formation. High resolution simulations may provide a much better separation between regions defined by high concentrations of biogenic volatile organic carbon (BVOC) and high NO_x levels (Esler et al., 2004; Pugh et al., 2013). The degree of artificial mixing induced by the grid resolution impacts O_3 formation efficiency and night-time titration (e.g., Cohan et al., 2006; Wild and Prather, 2006). Most studies have addressed the resolution sensitivity for 36, 12 and 4 km^2 for a US state and found that the 36-km^2 resolution leads to an under-prediction of daily maximum 8-h O_3 averages, and an over-prediction of daily minimum 8-h O_3 averages (Jang et al., 1995a,b; Arunachalam et al., 2006; Tie et al., 2010). Some studies support the finding that modelled O_3 formation systematically increases with resolution for regional and global scale applications (Wild and Prather et al., 2006; Hodnebrog et al., 2011). Evaluations of global, hemispheric and regional CTMs show that regional models typically perform better (Van Loon et al., 2007; Emery et al., 2012; Simpson et al. 2014a). Both Arunachalam et al. (2006) and Cohan et al. (2006) found similar model performances for 12-km and 4-km resolutions, except for the texture of local variability. For O_3 health impact, Thompson and Selin (2012) show that there were no significant further differences between assessments at 12-, 4- and 2-km resolution. In contrast, Valari and Menut (2008) did not find positive effects on model performance for O_3 by increasing their grid resolution from 48 to 6 km, which they linked to uncertainties in the emission data. Valin et al. (2011) found that a model resolution in the range of 4-12 km was sufficient to model NO_2 concentrations. PM and its components were found to be more sensitive to resolution than O_3 (Queen and Zhang, 2008). Differences are predicted mostly for primary rather than secondary PM components (Fountoukis et al., 2013). For all species, some studies show that finer grid resolutions do not give a better

127 performance or provide a mixed picture per component, due to the complexity in chemistry and
128 meteorology and their nonlinear interactions and responses to grid resolution (Queen and Zhang,
129 2008; Valari and Menut, 2008; Wu et al., 2008; Misenis and Zhang, 2010).
130 The magnitude of the impact of resolution changes depends on the lifetime of the component under
131 investigation, variability in and correlation with precursor concentrations, inhomogeneity of land
132 use, topography of the study area and meteorology, and the quality of the input data. Moreover,
133 optimal resolution depends on the expected application (e.g., health impact assessment, impact on
134 ecosystems through deposition, compliance to legislation), scale of the assessment as well as the
135 complexity of the model system and related computational effort used for this purpose. The studies
136 above have all focused on specific episodes or one or two months with a strong focus on summer
137 periods, whereas annual sensitivity assessments are lacking. Moreover, almost every study uses a
138 one way nesting approach in which the high resolution simulation was nested in the coarser
139 resolution simulations, enabling a comparison between the highest and coarsest resolution only for
140 a specific European country or US state. Continental scale sensitivity studies do not exist. In addition,
141 most studies focus on a single component and ensemble approaches are under-represented. Thus,
142 although many studies support that it is beneficial for the performance of a CTM to move away from
143 the current 50 km to higher resolutions, a thorough evaluation of the performance of CTMs as
144 function of resolution for Europe as a whole is lacking.
145 To support the discussion on the optimal compromise between accuracy of the model results and
146 the effort needed, both in terms of computational cost and expenses to produce high resolution
147 emission inventories and meteorological data, an initiative was taken for a model inter-comparison
148 exercise aimed at analysing the model performance of different CTMs as a function of model
149 horizontal spatial resolution. Five modelling teams participated in the exercise using the following
150 CTMs: CHIMERE (INERIS), CMAQ (BSC-CNS), EMEP MSC-W (MET Norway), LOTOS-EUROS (TNO), and
151 RCG (FU Berlin). All models performed four runs for Europe on various horizontal resolutions (56, 28,
152 14, and 7 km). The specifics of the CTMs and their experimental set-up are described in section 2. A
153 statistical evaluation of the performance of the models at different resolutions is presented in
154 section 3. Section 4 presents the most important findings from this exercise. Finally, our results are
155 discussed and put in perspective in section 5.
156
157
158

2. Participating models and simulation description

The participating CTMs in this study are CHIMERE (Schmidt et al., 2001; Bessagnet et al., 2004), CMAQ (Byun and Schere, 2006; Baldasano et al., 2011), EMEP (Simpson et al., 2012), LOTOS-EUROS (Schaap et al., 2008), and RCG (Stern et al., 2003). To represent, as much as possible, the uncertainty in our current knowledge of air quality processes, we allow all models to freely and independently utilize their best estimate for most input data and parameters (e.g., meteorological input, boundary conditions, natural emissions). However, the domains and horizontal resolution as well as anthropogenic emissions (described below) are r prescribed. A detailed overview of the model characteristics is provided in the supplementary material (Table S1). All models are regional-scale, limited-area models designed for short-term and long-term simulations of oxidant and aerosol formation. The models have different degrees of complexity (Table S1). EMEP and CMAQ describe the whole tropospheric column with 15-20 vertical layers, while LOTOS-EUROS, RCG and CHIMERE describe only the lower troposphere. LOTOS-EUROS has varying vertical layers, as it incorporates the dynamic mixing layer approach to determine the model vertical structure.

The simulations were performed for the year 2009 over a domain encompassing Europe (Figure 1). With each model four simulations were carried out doubling the resolution between each simulation. The horizontal spatial resolution ranges from 1.0 x 0.5 degrees to 0.125 x 0.0625 degrees (Table 1), which corresponds to 56 km x 56 km and 7 km x 7 km in the northern part of the domain and to 88 km x 56 km and 11 km x 7 km in the southern part, respectively. The output required for the exercise contains hourly as well as daily concentration distributions across Europe for oxidants as well as PM, its components and precursor species.

A common emission dataset was delivered by INERIS for all model resolutions separately. To be consistent with ongoing policy support simulations for the revision of the European air quality directive, we have used the emission database INERIS developed for the EC4MACS project for 2009 (EC4MACS, 2012). The basis for the allocation was the spatial differentiated EMEP 0.5° × 0.5° emission inventory for 2009 (Vestreng et al., 2007). The emission totals per sector were kept except for primary $PM_{2.5}$, for which GAINS (Amann et al., 2011) national emissions totals were considered more reliable for Belarus, Belgium, Bosnia and Herzegovina, Croatia, Czech Republic, Cyprus, France, Ireland, Lithuania, Luxembourg, Moldova, Macedonia, the Netherlands, Spain and Turkey (Klimont, International Institute for Applied Systems Analysis (IIASA), personal communication, 2011). Additional multiplication factors were applied on Polish emissions from domestic combustion (Standardized Nomenclature for Air Pollutants sector 2, or SNAP2) for PPM_{25} (4x) and PPM_{10} (x8), based on expert judgment (Klimont, International Institute for Applied Systems Analysis (IIASA),

192 personal communication, 2012). For the spatial allocation for SNAP 3, 7, 8, 9 and 10 (industrial
193 combustion, road transport, other transport, waste treatment and agriculture, respectively) the
194 TNO-MACC emissions (Kuenen et al., 2011) were used as proxy to increase the resolution of the
195 EMEP 0.5° x 0.5° annual totals. For SNAP 1, 4, 5 and 6 (energy production, industrial production
196 processes, extraction/distribution of fossil fuels and geoenergy and solvent use, respectively) the
197 EMEP 0.5° x 0.5° emissions were downscaled by land use weighing and E-PRTR (European Pollutant
198 Release and Transfer Register) data for industries from (http://prtr.ec.europa.eu/). For residential
199 combustion (SNAP 2) the country emissions were re-gridded with a methodology relating population
200 density and wood use as derived for France (Terrenoire et al., 2013), which has been applied to the
201 whole of Europe.

202 The database provides annual total 2D data, whereas CTMs require hourly, 3D emission
203 distributions. Except for SNAP 2, prescribed time profiles and height distributions were used
204 following the protocol as used in EURODELTA and CityDelta (Cuvelier et al., 2007; van Loon et al.,
205 2007). For SNAP 2 daily gridded modulation factors were calculated based on the degree day
206 concept with a reference temperature of 20 °C and a non-heating emission fraction of 0.1 (Cuvelier
207 et al., 2013), while the SNAP2 hourly variations were based on the EURODELTA hour-of-the-day
208 profile.

209 Natural emissions including biogenic VOCs, sea salt, soil NO_x and forest fires were described as
210 standard available within the individual models (Table S1). Note that only LOTOS-EUROS
211 incorporated a dust resuspension module for traffic following Schaap et al. (2009) which is
212 connected to the natural emission module of the model as this source is not included in emission
213 inventories.

214 Driving meteorology is taken from the European Centre for Medium Range Weather Forecasting
215 (ECMWF) operational analyses (CHIMERE, LOTOS-EUROS, EMEP) or from an optimal interpolation
216 analysis based on observations (RCG). These models interpolate the input meteorological data to the
217 required model resolution. The ECMWF meteorology has a resolution of ca. 16 km, so that models
218 running at e.g. 7 km are essentially running with fine-scale emissions but somewhat smoothed
219 meteorology. An important distinction is CMAQ, for which the WRF-ARWv3.3.1 model was run at
220 each required resolution using initial and boundary conditions from the Final Analyses of the
221 National Centers of Environmental Prediction (FNL/NCEP) at 12 h UTC. Chemical boundary
222 conditions are either climatological based on observations (RCG, EMEP), (climatological) monthly
223 mean distributions obtained from global model simulations with LMDz-INCA (CHIMERE) or
224 MOZART-4/NCEP (CMAQ) or 3-hourly distributions from a global model simulation (LOTOS-EUROS).

The expected change in concentrations due to an increase in resolution is therefore due to the much sharper gradients in the emissions and the sensitivities of process descriptions to concentration differences.

A specific feature of CHIMERE worth mentioning is the representation of mixing parameters above urban areas. An adjustment is made as a function of the urban land cover fraction in a grid cell which implies that the adjustment is different in each of the simulations performed here. In short, the argument is that the mixing in the urban environment within the canopy layer is overestimated with standard similarity theory. Therefore, the turbulent diffusion coefficient (Kz) in the first CHIMERE layer above urban areas is halved. The factor of 2 (derived from the literature) is also applied to lower the wind speeds in the first CHIMERE layer so as to limit the advection and dilution of primary pollutants close to the ground (Terrenoire et al., 2013). Comparison of CHIMERE simulations with and without the urban mixing correction shows a rather strong systematic increase of air pollutant concentrations over all major European cities. For NO_2, we observe an increase in concentrations ranging from a few $\mu g/m^3$ (suburban areas) to 20-30 $\mu g/m^3$ in major cities (e.g., Paris, London, Madrid and Milan). Consequently, additional O_3 titration is observed, decreasing O_3 by 4-10 $\mu g/m^3$ over the major cities, and 2-4 $\mu g/m^3$ over medium size cities. However, the strongest impact is observed for PM_{10} with an increase from a few $\mu g/m^3$ over medium size cities to 40 $\mu g/m^3$ for cities such as Katowice, Milan or Paris. Although this feature needs to be evaluated carefully, it falls outside the scope of this study.

3. Model performance evaluation

The impact of the increase of the model resolution on its performance needs to be quantified. As an important feature of the increased resolution should be to better describe horizontal gradients, and specifically to better resolve the gradients between source regions and the regional background, the quantification of the spatial correlation coefficients between modelled concentrations for all resolutions and observed concentrations at air quality monitoring stations is the first analysis performed here. The higher resolution of the emissions and the model may separate rural and urban monitoring sites that at lower resolution were in the same grid cell. To analyse the improvement in the spatial gradients, the spatial correlation coefficient, the slope of the best fit between modelled and observed annual mean concentrations and the bias are used. The second part of the evaluation focuses on the temporal variability of the concentrations looking at correlation coefficients and Root Mean Square Errors (RMSE). The DeltaTool (Thunis et al., 2011) developed in the framework of the

Forum for Air Quality Modelling in Europe (FAIRMODE, http://fairmode.jrc.ec.europa.eu/) activity has been used for the model evaluation.

In this exercise monitoring data from two air quality monitoring networks were used. Firstly, the EMEP network (Törseth et al., 2012), which was designed to support the implementation of the United Nation Convention on Long Range Transboundary Air Pollution (http://www.unece.org/env/lrtap/). As it is fitted to catch background air pollution patterns with stations at a considerable distance from source areas in rural or remote regions, this network is appropriate to evaluate regional scale models performance with coarse resolutions (50 – 150 km^2). Hence, the EMEP network may not be the most suitable network to investigate the impact of high resolution modelling. Therefore, we also use the AIRBASE database (http://acm.eionet.europa.eu/databases/airbase/) hosted by the European Environment Agency, which gathers observations from regulatory air quality networks implemented in the European Union. All station typologies (rural, suburban, urban, traffic, industrial) are available. To focus on the representation of gradients around large agglomerations a specific selection of stations was performed as presented below. Note that the AIRBASE network contains only regulatory pollutants, which include NO_2, O_3 and PM_{10}. Hence, concentrations of PM components are addressed using the EMEP network.

To highlight the differences between the models' results obtained with the four spatial resolutions, we focus the analysis on 30 European urban areas (see Figure 2). Within a radius of 30 km around each city (regardless of the city size) all AIRBASE urban background stations are used to evaluate the model results and assess the impact of increased resolution. This 30 km radius is chosen because it leads to a surface around each city approximately equal to the area of a 50x50 km EMEP grid cell. The number of stations available within each city area differs from city to city as does the split in terms of station types (Rural, Urban, EMEP, later referred to as R, U, E, respectively). An overview of the stations used per city for the evaluation is provided in Table 2. Within a radius of 30 km around each city not many rural stations are included in the analysis. To provide some large-scale evaluation in terms of comparison with rural background stations a comparison is also made with all EMEP and AIRBASE stations within a radius of 200 km around each city. In order to better highlight model differences in terms of urban areas and station types, groups of stations are generated in which statistical performance indicators are averaged. It is important to remember that the number of stations included in each of the city groups is different. While the total number of rural stations (R) used to produce the average for PM_{10} is 151 it reaches 98 for urban stations (Table 2).

289 The analysis is performed for daily mean PM_{10} concentrations, hourly NO_2 concentrations, and daily
290 maximum of the running 8 hour mean O_3 concentration. It mostly focuses on urban background
291 station types since the increased resolution is expected to have its maximum gain at those stations.

292

4. Results

4.1 Modelled distributions and sensitivity to emission changes

Figure 3 shows the annual mean concentrations of NO_2 as calculated for the four different horizontal grid resolutions. For NO_2 the concentration pattern going from the coarse grid EC4M1 (56 km) to the fine grid EC4M4 (7 km) increasingly reflects the underlying emission pattern. Overall, the increase in structure is tremendous. Especially going from 56 to 14 km adds a lot of detail in regions with high emission density, whereas the step from 14 km to 7 km does not show such a large change in the structure. The calculated concentrations increase in particular in the high emission density areas. The findings are consistent across all models. This is illustrated for PM_{10} distributions in Figure 4 which shows the 56 km and 7 km simulation for all models. As for NO_2, the differences across high emission density areas are noticeable. However, the increase in concentration for these areas is moderated because a large part of the modelled total PM_{10} is provided by secondary aerosols. Figure 5 shows that for the primary components of PM_{10}, the impact of increasing model resolution for urban stations is much larger than for the secondary inorganic aerosols (SIA) NO_3, SO_4 and NH_4. PPM_{10} concentrations display a stronger gradient over source regions than SIA because of the limited lifetime and because secondary components are not formed instantaneously from their precursors, allowing for transport of the latter away from the source regions before the precursor gases are transformed into secondary aerosols. Therefore, modelled SIA concentrations are much less affected by the grid size than PPM_{10}. Note that the absolute PM_{10} concentrations between the models differ as the modelling teams simulated the PM composition differently (e.g., including secondary organic aerosols and mineral dust or not).

Compared to NO_2 and PM_{10}, the effect of a decreasing grid size is smaller for O_3 (not shown). In general, there are only small changes in rural areas, which can be seen from the evaluation results in Figures 7 and 10. In urban areas, a decrease of the grid size leads also to a decrease of the calculated O_3 concentrations as titration by local NO_x sources is enhanced. The resolution effect is larger for annual mean O_3 concentrations than the average O_3 daily maximum or maximum running 8 hour mean concentrations. The latter is explained by the large impact of titration during stable conditions at night, for which daytime concentrations are less sensitive.

The observed concentration increments for higher model resolution is in part due to the resolution increase of the emission database, and in particular the differences between emission densities at fine and coarse resolutions, but also to the decrease of the artificial dilution of emissions compared to the larger grid area (Gego et al., 2005; Pay et al., 2014). Figure 6 illustrates the relation between

concentration deltas (Concentration [7 km] – Concentration [56 km]) and emission deltas (Emission [7 km] – Emission [56 km]) for NO_2 and PM_{10} for RCG and CMAQ. Most of the concentration increments (deltas) can be explained by the emission deltas. The spread around the trend line provides some information on the importance of other factors such as the local meteorology or the role of chemistry. Most of this additional spread happens in stations belonging to countries like Italy, Greece, Portugal and Spain. The fit parameters for all models are summarized in Table 2. From the correlation coefficient we derive that for EMEP, CHIMERE, RCG and LOTOS-EUROS about 70% of the model response to grid resolution is explained by the higher resolution of the emission data. CMAQ shows a lower sensitivity to the difference in emission strength (46%), which we explain by the impact of using high resolution meteorological data. Also, for PM_{10} a significant part (~70%) of the concentration increment (7-56 km) can be explained by the emission density increment for EMEP, CHIMERE, and RCG as demonstrated by the high values of the coefficient of determination (R^2). Again, CMAQ provides an exception with a much larger spread (R^2 = 23 %). For PM_{10} also LOTOS-EUROS shows somewhat more spread (R^2 = 54%). In both models most of this additional variance happens in stations belonging to countries like Italy, Greece, Portugal and Spain. The additional variance by LOTOS-EUROS is explained by a modelled contribution of road dust emissions. CHIMERE and RCGC clearly show the highest response followed by LOTOS-EUROS whereas EMEP and CMAQ exhibits significantly lower increments. The absolute resolution effect for PM_{10} is lower than for NO_2, as explained by the higher importance of the rural background concentrations containing secondary material for PM_{10}. On the other hand, the slopes for PM_{10} expressing the concentration increase per unit emission are steeper than for NO_2. Possible explanation for the steeper slopes and slightly different behaviour for PM_{10} in comparison to NO_2 may lie in the fact that NO_2 increments may be limited due to the availability of O_3 as NO_2 is formed through titration of O_3.

4.2 Model performance

4.2.1. Spatial analysis

Nitrogen dioxide (NO_2)

For the models' evaluation at the different resolutions, the annual average NO_2 concentrations are grouped by three station categories: EMEP stations (E), urban stations (U) and rural stations (R). As illustrated for the EMEP model in Figure 7, the grid resolution increment is as expected very weak at rural and EMEP stations but clearly visible for urban stations. It is interesting to note the small decrease of modelled NO_2 concentrations at the EMEP stations. All models agree on a negligible impact at rural and EMEP stations and on the trends at urban stations. Although the NO_2

concentration differs between models, the gain resulting from an increased resolution is significant for urban stations (from 10 to 20 μg/m^3 between 56 km and 7 km). The largest increments are seen for CHIMERE and RCG and slightly lower responses for the other models. For EMEP, LOTOS-EUROS and CMAQ, the largest gains are seen going from 56 to 28 to 14 km resolution with a kind of saturation to the higher resolutions. CHIMERE and RCGC react more evenly to each resolution change than the others.

A summary of the spatial statistical analysis for hourly NO_2 is given in Figure 8. As expected the performance at rural sites is stable with resolution for all models but one. CMAQ shows a significant increase in the slope and explained variability between observed and modelled concentrations which is attributed to the impact of the scale dependent meteorology. For urban stations there is a strong dependence on the resolutions, with an increase in slopes and correlations, and a reduction of the bias for all the models. Also for the statistical parameters, the largest gains appear to occur between 56 km and 14 km resolution, with a saturation between 14 and 7 km resolution.

Particulate Matter (PM_{10})

For PM_{10}, the increased grid resolution yields a very slight decrease in concentration at rural and EMEP stations in all models. While the absolute concentration modelled differs between the models, the change in modelled concentrations because of the resolution change is similar. Urban increments are similar for the CHIMERE, LOTOS-EUROS, and RCGC, whereas for EMEP and CMAQ they are significantly lower. For EMEP, and to a smaller extent for CMAQ, we explain this by a higher surface layer depth (~90 m for EMEP, ~39 m for CMAQ) than in other models (~20 m). As for NO_2, the impact of the resolution is variable among the agglomerations studied due to variable PM emission strengths, station number and locations and grid impacts. At urban locations and comparing 56 km to 7 km simulations, the spatial bias for PM_{10} reduces by 6.5 μg/m^3 for CHIMERE, by 5.5 μg/m^3 for RCGC and LOTOS-EUROS, whereas the EMEP bias reduces about 3.3 μg/m^3 and only 1.2 μg/m^3 for CMAQ. A summary of the spatial statistical analysis for daily PM10 is given in Figure 9 and shows that the slope is significantly improved with higher resolution as a result of a lower bias at urban stations. As shown in Figure 5 the modelled PM component concentrations indicate that the concentration change is completely induced by the primary components and that only a slight sensitivity to model resolution is observed for the secondary components. However, the reduction of the explained variability for PM_{10} between urban areas shows that increasing the resolution of the emissions (without adapting gridding approaches) and models is not sufficient to improve the assessment of exposure of European urban populations to PM.

389 *Ozone (O_3)*

390 For O_3 we focus our analysis on the models' performance for the daily maximum of the running 8
391 hour mean (O3Max8Hr). The reason is that the analysis is then focussed on the high O_3 regime
392 during daytime, and is less sensitive to the impact of differences between models on night time
393 mixing and titration. First, the annual average O3Max8Hr concentrations show a different behaviour
394 compared to PM_{10} and NO_2. Due to the titration impact of NO_x emissions near sources, the
395 concentrations are lower inside a city than outside (Figure 10). The average pattern as a function of
396 station type and the response to a resolution change between all models is very similar. O3Max8h
397 annual biases reduce for all models over urban stations. At rural stations the impact of resolution is
398 rather small, while an increase in resolution has a significant effect for the urban locations. At these
399 stations the slopes decrease for EMEP, LOTOS-EUROS and RCGC, while CHIMERE and CMAQ show a
400 minimum at the 28 km resolution model run. Spatial correlation coefficients decrease slightly with
401 increasing resolution. Also the urban bias improves, except for LOTOS-EUROS. However, increasing
402 resolution and adding more local variability decreases the representation of the spatial contrasts,
403 although for NO_2 the spatial patterns become better between cities. This may mean that it is not the
404 variability in NO_x emission source strengths between the urban regions that is the most important
405 source of uncertainty for O_3 gradients during the day at these scales. Instead, differences in mixing
406 regimes, chemical regimes and uncertainties in NMVOC speciation could be more important.

407 *4.2.2. Temporal analysis*

408 To investigate the impact of resolution on modelled time series and the models' temporal
409 performance, Table 4 shows an overview of temporal correlation coefficients, biases and RMSE
410 values for the 56 km and 7 km resolutions. In general, for urban and suburban stations the models'
411 performance is better (i.e., lower RMSE, lower bias) for the 7 km simulations as the urban increment
412 is better represented than for the coarser resolution simulations. The temporal correlation
413 coefficient between modelled and measured time series shows a mixed picture with small deviations
414 between the different resolutions. This shows that the increased resolution better reproduces the
415 gradients in pollution regimes, but does not help to improve significantly the performance in time.
416 The latter was expected as meteorological conditions and pollution levels upwind have a key impact
417 on the temporal variability of air pollutants.

418

5. Discussion and conclusions

In this study we have investigated the impact of increasing horizontal resolution on the performance of five regional chemistry transport models. The exercise indicates that the model responses to an increase in resolution show a broadly consistent picture among all models. The analysis showed that the grid size does not play a major role for air quality model calculations, which are targeted on the determination of the background (non-urban) air quality. Downscaling model resolution does not change concentration estimations and model performance at rural and EMEP sites, which suggests that long-range transport and secondary compounds dominate the concentration fields at these locations. The grid resolution plays an important role in agglomerations characterized by high emission rates. The urban signal, i.e. the concentration difference between high emission areas and their surroundings, usually increases with decreasing grid size. This grid effect is more pronounced for NO_2 than for PM10, because a large part of the urban PM_{10} mass consists of secondary components not affected by resolution. The degree of artificial mixing induced by the grid resolution impacts O_3 titration as already identified in many studies (e.g., Cohan et al., 2006; Wild and Prather, 2006; Tie et al., 2010). O_3 is less sensitive to model resolution than PM or NO_2 (Queen and Zhang, 2008). The grid effect differs between urban regions. The strength of the urban signal is a function of local emissions conditions (e.g., extension of the emission areas, emission density, emission gradient) and meteorological conditions determining ventilation efficiency. For all models, increasing model resolution improves the model performance at stations near large conglomerations as reflected by lower biases for PM_{10}, NO_2, and O_3 and an increased spatial correlation coefficient for NO_2. On the other hand, for PM_{10} and O_3 increasing resolution decreases the representation of the spatial contrasts between agglomerations. The reduction of the explained concentration variability between urban areas shows that increasing the resolution of the models without improving the input data is not the holy grail to improve the assessment of exposure of European urban populations to PM_{10} and O_3.

In this exercise four models used meteorology at a single resolution that was interpolated to the model grid. For this purpose, meteorological data at a resolution of about 20 km was used. Degrading the model resolution based on high resolution meteorology shows no impact on the average transport of non-reactive pollutants, whereas reactive species are somewhat sensitive, illustrating the nonlinear relationship between chemistry and horizontal grid resolution as already shown by Jang et al. (1995a,b). In case of CMAQ simulations the meteorological WRF-ARW model was run at the four required resolutions. The much stronger dependence of the model performance for this model indicates that the availability of meteorological data with increased resolution has

benefitted the performance of CTMs in the past. The importance of the impact of the resolution at which the meteorological data was obtained is also illustrated for PM by Wolke et al. (2012) and for O_3 by Tie et al. (2010). It should be noted that more significant response for CHIMERE simulations compared to the other participating models seems also to be related to a specific treatment of the mixing parameterization over urban areas. As non-hydrostatic meteorological data at around 1-5 km resolution are available nowadays, studies using high resolution meteorology indicate that including the urban impacts on ventilation and subgrid emission variability is needed to study the variability of short-lived pollutants across large urban agglomerations (Pay et al., 2014).

As about 70% of the model response to grid resolution is determined by the difference in emission strength, improved knowledge on spatial variation in emission at high resolution is necessary for the improvement of modelled urban increments . Timmermans et al. (2013) compared urban increment calculations using the downscaled TNO-MACC emission database against using a database that contains bottom up emissions for a number of large conglomerations in Europe. It was shown that the air pollutant emissions in these large agglomerations can be significantly lower than those in the down scaled emission database, mostly due to a more efficient energy use (economy of scale) and the use of cleaner fuels in the urban areas. This was especially the case for PM10, whereas NO_2 concentrations were hardly effected. These findings may explain why for NO_2 the increased resolution provides a larger increase in performance than for PM_{10}. Moreover, as a top down inventory was used here, part of the lower bias for PM_{10} at the higher resolution may be attributed to the lack of detail in the gridding of the emission database. Hence, to model urban increments improved spatial allocation algorithms and proxy data including local information are needed. In addition, parameterizations for missing (urban) sources such as road dust suspension should be developed (Pay et al., 2011; Amato et al., 2013). Note that also the temporal variability of emissions is poorly represented and could be improved (Kukkonen et al., 2012; Menut et al., 2012; Mues et al., 2013).

It is difficult to define a grid size that is adequate to resolve the urban signal under all conditions occurring in a European-wide modelling area, and for all pollutants. Ideally, a grid size in the range of a few km down to 1 km should be chosen. Such a small grid size is not feasible for operational regional model applications because the data demands and operating requirements are too large (Colette et al., 2014). If the main emphasis of a model application is targeted on the determination of background air quality for rural areas and large agglomerations, the 28 km grid scale or, if the data and operational requirements can be fulfilled, the 14 km grid scale currently seems to be a reasonable compromise between a pure background application and an application which

reproduces most of the urban signals (7 km or even higher resolution). Several other authors found little additional value of resolutions higher than those proposed here (Arunachalamet al., 2006; Cohan et al., 2006; Thompson and Selin, 2012). Not surprisingly, a resolution in the range of 14-28 km coincides with the resolution for which the meteorological and emission data are representative.

This study clearly shows that increasing model resolution is advantageous and that leaving a resolution of 50 km in favour of a resolution between 10 and 20 km is practical and worthwhile. Other input data, notably emissions and meteorological data, are also available at this scale, but become more problematic at finer resolutions. Further improvements of resolution should go hand in hand with improved resolution of meteorological models, the improved representation of spatial and temporal variability in emission data as well as the adjustment of CTM process descriptions and parameterizations to higher resolution.

Acknowledgement

The EMEP MSC-W team received funding through EMEP under UN-ECE, as well as through the EU FP7 ECLAIRE project (Project No. 282910). INERIS was funded by the French Ministry in charge of Ecology the European project EC4MACS (EU LIFE, www.ec4macs.eu). The LOTOS-EUROS modelling team was funded by the Dutch Ministry for Infrastructure and Environment. The RCGC modelling team was funded by Umweltbundesamt Germany, project number 21981. The CMAQ modelling team received funding from the the Severo Ochoa Program awarded by the Spanish Government (grant number SEV-2011-00067) and the postdoctoral grant in the Beatriu de Pinós programme (file number 2011 BP-A 00427).

507 **References**

508 Adams, R.M., Crocker, T.D., Thanvaibulchai, N., 1982. An economic assessment of air pollution
509 damages to selected annual crops in southern California. Journal of Environmental Economics and
510 Management 9, 42-58

511 Ainsworth, E. A.; Yendrek, C. R.; Sitch, S.; Collins, W. J. & Emberson, L. D. The Effects of Tropospheric
512 Ozone on Net Primary Productivity and Implications for Climate Change Ann. Rev. Plant Biol, 2012,
513 63, {637-661}

514 Amann, M. , Bertok, I., Borken-Kleefeld, J., Cofala, J., Heyes, C., Höglund-Isaksson, L., Klimont, Z.,
515 Nguyen, B., Posch, M., Rafaj, P., Sandler, R., Schöpp, W., Wagner, F., Winiwarter, W., 2011. Cost-
516 effective control of air quality and greenhouse gases in Europe: Modeling and policy applications.
517 Environmental Modelling and Software, 26, pp. 1489-1501

518 Amato, F., Schaap, M., Denier van der Gon, H.A.C., Pandolfi, M., Alastuey, A., Keuken, M., Querol, X.
519 2013. Short-term variability of mineral dust, metals and carbon emission from road dust
520 resuspension, Atmospheric Environment, 74, pp. 134-140.

521 Appel, K.W. , Foley, K.M., Bash, J.O., Pinder, R.W., Dennis, R.L., Allen, D.J., Pickering, K., 2011. A
522 multi-resolution assessment of the Community Multiscale Air Quality (CMAQ) model v4.7 wet
523 deposition estimates for 2002-2006, Geoscientific Model Development **4** (2), pp 357-371

524 Arunachalam, S., Holland, A., Do, B., Abraczinskas, M., 2006. A quantitative assessment of the
525 influence of grid resolution on predictions of future-year air quality in North Carolina, USA,
526 Atmospheric Environment, 40 (26), 5010-5026.

527 Baldasano JM, Pay MT, Jorba O, Gassó S, Jiménez-Guerrero P, 2011. An annual assessment of air
528 quality with the CALIOPE modelling system over Spain. Sci Total Environ, 409, 2163-2178.
529 doi:10.1016/j.scitotenv.2011.01.041.

530 Bell, M.L., McDermott, A., Zeger, S.L., Samet, J.M., Dominici, F., 2004. Ozone and short-term
531 mortality in 95 US urban communities, 1987e2000. Journal of the American Medical Association 292
532 (19), 2372-2378. http://dx.doi.org/10.1001/jama.292.19.2372.

533 Berge, E. and Jakobsen, H. A., 1998. A regional scale multi-layer model for the calculation of long-
534 term transport and deposition of air pollution in Europe, Tellus, 50, 205–223.

535 Bessagnet, B., A. Hodzic, R. Vautard, M. Beekmann, S. Cheinet, C. Honoré, C. Liousse, L. Rouil, 2004.
536 Aerosol modelling with CHIMERE—preliminary evaluation at the continental scale, Atmospheric
537 Environment, 38 (18), 2803-2817, ISSN 1352-2310, 10.1016/j.atmosenv.2004.02.034.

538 Bobbink, R., Hornung, M., and Roelofs, J. G. M., 1998. The effects of air-borne nitrogen pollutants on
539 species diversity in natural and semi-natural European vegetation, J. Ecol., 86, 717–738.

540 Byun, D. and Schere, K. L., 2006. Review of the Governing Equations, Computational Algorithms, and
541 Other Components of the Models-3 Community Multiscale Air Quality (CMAQ) Modelling System,
542 Appl. Mech. Rev., 59, 51–77.

Cohan, D.S., Hu, Y., Russell, A.G., 2006. Dependence of ozone sensitivity analysis on grid resolution, Atmospheric Environment, 40 (1), 126-135.

Colette, A., Bessagnet, B., Meleux, F., Terrenoire, E., and Rouïl, L.: Frontiers in air quality modelling, Geosci. Model Dev., 7, 203-210, 2014.

Cuvelier, C., Thunis, P., Vautard, R., Amann, M., Bessagnet, B., Bedogni, M., Berkowicz, R., Brandt, J., Brocheton, F., Builtjes, P., Coppalle, A., Denby, B., Douros G., Graf, A., Hellmuth, O., Honore´ , C., Hodzic, A., Jonson, J., Kerschbaumer, A., de Leeuw, F., Minguzzi, E., Moussiopoulos, N., Pertot, C., Pirovano, G., Rouil, L., Schaap, M., Stern, R., Tarrason, L., Vignati, E., Volta, M., White, L., Wind, P., Zuber, A., 2007. CityDelta: a model intercomparison to explore the impact of emission reductions in 2010 in European cities in 2010, Atmospheric Environment 41, 189–207.

Cuvelier, C., Thunis, P., Karam, D., Schaap, M., Hendriks, C., Kranenburg., R., Fagerli, H., Nyíri, Á., Simpson, D., Wind, P., Schulz, M., Bessagnet, B., Colette, A., Terrenoire, E., Rouïl, L., Stern, R., Graff, A., Baldasano, J.M., Pay, M.T., 2013. ScaleDep: Performance of European chemistry-transport models as function of horizontal spatial resolution. EMEP Technical Report 1/2013.

Dockery, D.W., Pope, C.A., Xu, X., Spengler, J.D. Ware, J.H. Fay, M.E., Ferris, B.G., Speizer, F.E., 1993. Ann association between air pollution and mortality in six US cities. The New England Journal of Medicine 329, 1753-1759

EC4MACS, 2012, The GAINS integrated assessment model, EC4MACS report, March 2012. http://www.ec4macs.eu/content/report/EC4MACS_Publications/MR_Final%20in%20pdf/GAINS_Methodologies_Final.pdf

Eliassen, A. and Saltbones, J., Modelling of long range transport of sulphur over Europe: a two year model run and some experiments, Atmos. Environ., 17, 1457-1473, 1983

Emery, C., Jung, J., Downey, N., Johnson, J., Jimenez, M., Yarwood, G., Morris, R., 2012. Regional and global modeling estimates of policy relevant background ozone over the United States. Atmospheric Environment 47, 206-217.

Esler, J.G., Roelofs, G.J., Köhler, M.O., O'Connor, F.M., 2004. A quantitative analysis of grid-related systematic errors in oxidising capacity and ozone production rates in chemistry transport models, Atmospheric Chemistry and Physics, 4 (7), 1781-1795.

Fountoukis, C., Koraj, D., Denier van der Gon, H.A.C., Charalampidis, P.E., Pilinis, C., Pandis, S.N., 2013. Impact of grid resolution on the predicted fine PM by a regional 3-D chemical transport model, Atmospheric Environment, 68, 24-32.

Gego, E., Hogrefe, C., Kallos, G., Voudouri, A., Irwin, J.S., Trivikrama, R., 2005. Examination of model predictions at different horizontal grid resolutions. Environmental Fluid Mechanics, 5, 63-85.

Hendriks, C., Kranenburg, R., Kuenen, J.J.P., Van Gijlswijk, R., Wichink Kruit, R., Segers, A.J, Denier van der Gon, H.A.C, Schaap, M., 2013. The origin of ambient particulate matter concentrations in the Netherlands. Atmospheric Environment, 69, p. 289-303, doi:10.1016/j.atmosenv.2012.12.017

579 Hodnebrog, T., Stordal, F., Berntsen, T.K., 2011. Does the resolution of megacity emissions impact
580 large scale ozone?, Atmospheric Environment, 45 (38), pp. 6852-6862.

581 Jang, J.-C.C., Jeffries, H.E., Byun, D., Pleim, J.E., 1995a. Sensitivity of ozone to model grid resolution -
582 1. Application of high-resolution regional acid deposition model, Atmospheric Environment, 29 (21),
583 pp. 3085-3100.

584 Jang, J.-C.C., Jeffries, H.E., Tonnesen, S., 1995b. Sensitivity of ozone to model grid resolution - II.
585 Detailed process analysis for ozone chemistry, Atmospheric Environment, 29 (21), pp. 3101-3114.

586 Kiehl, J.T., and Briegleb, B.P., 1993. The relative roles of sulfate aerosols and greenhouse gases in
587 climate forcing, Science, 260, 311 – 314

588 Kryza, M., Dore, A.J., Werner, M., Błaś, M., 2012. Comparison and evaluation of the 1 km and 5 km
589 resolution FRAME modelled annual concentrations of nitrogen oxides, International Journal of
590 Environment and Pollution, 50 (1-4), pp. 53-63.

591 Kuenen, J., Denier van der Gon, H., Visschedijk, A., van der Brugh, H., van Gijlswijk, R., 2011. MACC
592 European Emission Inventory for the Years 2003-2007. TNO report, TNO-060-UT-2011-00588,
593 Utrecht.

594 Kukkonen J., Olsson T., Schultz D.M., Baklanov A., Klein T., Miranda A.I., Monteiro A., Hirtl M.,
595 Tarvainen V., Boy M., Peuch V.-H., Poupkou A., Kioutsioukis I., Finardi S., Sofiev M., Sokhi R.,
596 Lehtinen K.E.J., Karatzas K., San José R., Astitha M., Kallos G., Schaap M., Reimer E., Jakobs H., Eben
597 K., 2012. A review of operational, regional-scale, chemical weather forecasting models in Europe.
598 Atmospheric Chemistry and Physics, 12 (1), pp. 1-87.

599 Lauwaet, D., Viaene, P., Brisson, E., van Noije, T., Strunk, A., Van Looy, S., Maiheu, B., Veldeman, N.,
600 Blyth, L., De Ridder, K., Janssen, S., 2013. Impact of nesting resolution jump on dynamical
601 downscaling ozone concentrations over Belgium, Atmospheric Environment, 67, pp. 46-52.

602 Menut L., Goussebaile A., Bessagnet B., Khvorostiyanov D., Ung A., 2012. Impact of realistic hourly
603 emissions profiles on air pollutants concentrations modelled with CHIMERE. Atmospheric
604 Environment 49, 233-244.

605 Mills, G.; Hayes, F.; Simpson, D.; Emberson, L.; Norris, D.; Harmens, H. & Büker, P. Evidence of
606 widespread effects of ozone on crops and (semi-) natural vegetation in Europe (1990-2006) in
607 relation to AOT40- and flux-based risk maps Global Change Biology, Blackwell Publishing Ltd, 2011,
608 17, 592-613

609 Misenis, C., Zhang, Y., 2010. An examination of sensitivity of WRF/Chem predictions to physical
610 parameterizations, horizontal grid spacing, and nesting options, Atmospheric Research, 97 (3), pp.
611 315-334.

612 Mues, A., Kuenen, J.J.P., Hendriks, C., Manders, A., Segers, A., Scholz, Y., Hueglin, C., Builtjes, P., and
613 Schaap, M., 2014. *Sensitivity of air pollution simulations with LOTOS-EUROS to temporal distribution*
614 *of anthropogenic emissions.* Atmos. Chem. Phys., 14, 939-955, doi:10.5194/acpd-14-939-2014.

Pay MT, Piot M, Jorba O, Basart S, Gassó S, Jiménez-Guerrero P, Gonçalves M, Dabdub D, Baldasano JM, 2010. A full year evaluation of the CALIOPE-EU air quality system in Europe for 2004: a model study. *Atmos Environ*, 44, 3322-3342.

Pay, M.T., Jiménez-Guerrero, P., Baldasano, J.M., 2011. Implementation of resuspension from paved road from the improvement of CALIOPE air quality system in Spain. Atmos. Environ., 45, 802-207.

Pay, M.T., Martínez, F., Guevara, M., Baldasano, J.M., 2014. Air quality forecasts at kilomenter scale grid over Spanish complex terrains. Geosci. Model Dev. Discuss., 7, 2293-2334.

Pugh, T.A.M., Ashworth, K., Wild, O., Hewitt, C.N., 2013. Effects of the spatial resolution of climate data on estimates of biogenic isoprene emissions. Atmospheric Environment **70,** pp 1-6

Pyle, J. A. and Zavody, A. M., 1990. The modelling problems associated with spatial averaging, Q. J. R. Meteorol. Soc., 116, 753–766.

Queen, A., Zhang, Y., 2008. Examining the sensitivity of MM5-CMAQ predictions to explicit microphysics schemes and horizontal grid resolutions, Part III-The impact of horizontal grid resolution, Atmospheric Environment, 42 (16), pp. 3869-3881.

Reis, S.; Grennfelt, P.; Klimont, Z.; Amann, M.; ApSimon, H.; Hettelingh, J.-P.; Holland, M.; LeGall, A.-C.; Maas, R.; Posch, M.; Spranger, T.; Sutton, M. A. & Williams, M. From Acid Rain to Climate Change Science, 2012, 338, 1153-1154

Schaap, M., van Loon, M., ten Brink, H.M., Dentener, F.D., Builtjes, P.J.H., 2004. Secondary inorganic aerosol simulations for Europe with special attention to nitrate. Atmospheric Chemistry and Physics 4, 857-874.

Schaap, M., Sauter, F., Timmermans, R.M.A., Roemer, M., Velders, G., Beck J., Builtjes, P.J.H., 2008. The LOTOS-EUROS model: description, validation and latest developments, Int. J. Environment and Pollution, Vol. 32, No. 2, pp.270–290

Schaap, M., Manders, A. A. M., Hendriks, E. C. J., Cnossen, J.M., Segers, A. J., Denier van der Gon, H. A. C., Jozwicka, M., Sauter, F. J., Velders, G. J. M., Matthijsen, J., and Builtjes, P. J. H., 2009. Regional Modelling of Particulate Matter for the Netherlands, PBL report 500099008, Bilthoven, The Netherlands, http://www.rivm.nl/bibliotheek/rapporten/500099008.pdf (last access: 31 May 2013).

Schmidt, H., Derognat, C., Vautard, R., Beekmann, M., 2001. A comparison of simulated and observed ozone mixing ratios for the summer of 1998 in Western Europe, Atmospheric Environment **35**, pp 6277–6297

Schöpp, W., Amann, M., Cofala, J., Heyes, C., Klimont, Z., 1999. Integrated assessment of European air pollution emission control strategies, Environmental Modelling & Software **14**, 1–9

Simpson, D., Benedictow, A., Berge, H., Bergström, R., Emberson, L. D., Fagerli, H., Flechard, C: Hayman, G. D., Gauss, M., Jonson, J. E., Jenkin, M. E., Nyíri, Á., Richter, C., Semeena, V. S., Tsyro, S., Tuovinen, J.-P., Valdebenito, A., and Wind, P., 2012. The EMEP MSC-W chemical transport model – technical description, Atmos. Chem. Phys. 12, 7825-7865, doi:10.5194/acp-12-7825-2012.

651 Simpson, D., Christensen, J., Engardt, M., Geels, C., Nyiri, A., Soares, J., Sofiev, M., Wind, P., Langner,
652 J., 2014a. Impacts of climate and emission changes on nitrogen deposition in Europe: a multi-model
653 study Atmos. Chem. Physics, 14, 6995-7017

654 Simpson, D., Arneth, A., Mills, G., Solberg, S., Uddling, J., 2014b. Ozone - the persistent menace;
655 interactions with the N cycle and climate change Current Op. Environ. Sust., In press

656 Stern, R., Yamartino, R., Graff, A., 2003. Dispersion modeling within the European community's air
657 quality directives: long term modelling of O3 PM10 and NO2. 26th ITM on Air Pollution Modelling
658 and its Application. 26–30 May 2003, Istanbul, Turkey.

659 Stroud, C.A., Makar, P.A., Moran, M.D., Gong, W., Gong, S., Zhang, J., Hayden, K., Mihele, C., Brook,
660 J.R., Abbatt, J.P.D., Slowik, J.G., 2011. Impact of model grid spacing on regional- and urban- scale air
661 quality predictions of organic aerosol, Atmospheric Chemistry and Physics, 11 (7), pp. 3107-3118.

662 Terrenoire, E., Bessagnet, B., Rouïl, L., Tognet, F., Pirovano, G., Létinois, L., Colette, A., Thunis, P.,
663 Amann, M., and Menut, L., 2013. High resolution air quality simulation over Europe with the
664 chemistry transport model CHIMERE, Geosci. Model Dev. Discuss., 6, 4137–4187, doi:10.5194/gmdd-
665 6-4137-2013.

666 Thompson, T.M., Selin, N.E., 2012. Influence of air quality model resolution on uncertainty
667 associated with health impacts, Atmospheric Chemistry and Physics, 12 (20), pp. 9753-9762.

668 Thunis P., E. Georgieva, A. Pederzoli, 2011. The DELTA tool and Benchmarking Report template:
669 Concepts and User's Guide Version 1., Fairmode report, http://fairmode.ew.eea.europa.eu/

670 Tie, X., Brasseur, G., Ying, Z., 2010. Impact of model resolution on chemical ozone formation in
671 Mexico City: Application of the WRF-Chem model, Atmospheric Chemistry and Physics **10** (18), pp
672 8983-8995

673 Timmermans, R.M.A., Denier van der Gon, H.A.C. , Kuenen, J.J.P., Segers, A.J., Honoré, C., Perrussel,
674 O., Builtjes, P.J.H., Schaap, M., 2013. Quantification of the urban air pollution increment and its
675 dependency on the use of down-scaled and bottom-up city emission inventories. Urban Climate, 6,
676 44-62, doi:10.1016/j.uclim.2013.10.004.

677 Törseth, K., Aas, W., Breivik, K, Fjæraa, A. M., Fiebig, M., Hjellbrekke, A. G., Lund Myhre, C., Solberg,
678 S., Yttri, K. E., 2012. Introduction to the European Monitoring and Evaluation Programme (EMEP) and
679 observed atmospheric composition change during 1972--2009 Atmos. Chem. Physics, 12, 5447-5481

680 Valari, Myrto, Laurent Menut, 2008. Does an Increase in Air Quality Models' Resolution Bring Surface
681 Ozone Concentrations Closer to Reality?. J. Atmos. Oceanic Technol., 25, 1955–1968. doi:
682 http://dx.doi.org/10.1175/2008JTECHA1123.1

683 Valin, L.C., Russell, A.R., Hudman, R.C., Cohen, R.C., 2011. Effects of model resolution on the
684 interpretation of satellite NO2 observations, Atmospheric Chemistry and Physics **11** (22), pp 11647-
685 11655

686 van Loon, M., Vautard, R., Schaap, M., Bergström, R., Bessagnet, B., Brandt, J., Builtjes, P.J.H.,
687 Christensen, J., Cuvelier, K., Jonson, J.E., Krol, M., Langner, J., Roberts, P., Rouil, L., Stern, R.,

Tarrasón, L., Thunis, P., Vignati, E., White, L., Wind, P., 2007. Evaluation of long-term ozone simulations from seven regional air quality models and their ensemble. Atmospheric Environment 41, 2083-2097.

Vestreng, V. , G. Myhre, H. Fagerli, S. Reis, and L. Tarrason, 2007. Twenty-five years of continuous sulphur dioxide emission reduction in Europe, Atmos. Chem. Phys., 7, 3663–3681.

Vieno, M., Dore, A., J., Stevenson, D.S., et al., Modelling surface ozone during the 2003 heat wave in the UK, Atmos. Chem. Physics, 10, 7963-7978, 2010.

Wagner, F., Amann, M., Bertok, I., Cofala, J., Heyes, C., Klimont, Z., Rafaj, P., Schöpp, W., 2000. Baseline Emission Projections and Further Cost-effective Reductions of Air Pollution Impacts in Europe - A 2010 Perspective, NEC Scenario Analysis Report Nr. 7, International Institute for Applied Systems Analysis (IIASA), Laxenburg, Austria, 27 August 2010.

Wild, O., Prather, M.J., 2006. Global tropospheric ozone modeling: quantifying errors due to grid resolution. Journal of Geophysical Research 111, d11305. http://dx.doi.org/10.1029/2005jd006605

Wolke, R., Schröder, W., Schrödner, R., Renner, E., 2012. Influence of grid resolution and meteorological forcing on simulated European air quality: A sensitivity study with the modeling system COSMO-MUSCAT, Atmospheric Environment, 53, pp. 110-130.

Wu., S.-Y., Krishnan., S., Zhang, Y., Aneja, V., 2008. Modeling atmospheric transport and fate of ammonia in North Carolina—Part I: Evaluation of meteorological and chemical predictions, Atmospheric Environment **42**, pp 3419–3436

Zhang, Y., Bocquet, M., Mallet, V., Seigneur, C., Baklanov, A., 2012. Real-time air quality forecasting, part I: History, techniques, and current status. Atmos. Environ., 60, 632-655.

Tables to manuscript "Performance of European chemistry transport models as function of horizontal resolution"

Table 1: Definition of the four resolution domains. nx and ny indicate the number of cells in longitude and latitude direction. N and S refer to cells in the northern and southern part of the domain, respectively.

Resolution domain	**nx**	**ny**	**ΔLon** (degrees)	**ΔLat** (degrees)	**ΔLon** x **ΔLat** (km x km)	**SW corner grid centre** (Lon / Lat)
EC4M1	41	52	1.0	0.5	56 x 56 (N) 88 x 56 (S)	-10.000 / 36.125
EC4M2	82	104	0.5	0.25	28 x 28 (N) 44 x 28	-10.250 / 36.000
EC4M3	164	208	0.25	0.125	14 x 14 (N) 22 x 14 (S)	-10.3750 / 35.9375
EC4M4	328	416	0.125	0.0625	7 x 7 (N) 11 x 7 (S)	-10.43750 / 35.90625

Table 2: Overview of the available measurement stations per city, pollutant and station type. For AIRBASE urban groups a radius of 30 km around each city is assumed, whereas for the AIRBASE rural background and EMEP groups, the radius is 200 km.

		PM10			NO2			O3		
		Urban	Rural	EMEP	Urban	Rural	EMEP	Urban	Rural	EMEP
Amsterdam	**AMS**	2	11		7	12		1	12	
Athens	**ATH**				2	1		2	1	
Barcelona	**BAR**	5	4		5	6		4	6	
Berlin	**BER**	5	10	1	6	9		3	9	
Bilbao	**BIL**	4	5	1	4	5	1	4	5	1
Bruxelles	**BRU**		11		1	16		1	16	
Bucarest	**BUC**	1	0			1		1	1	
Budapest	**BUD**	1	2	2	1	2		1	2	2
Cologne	**COL**		7			9			9	
Dublin	**DUB**	2	2		2	2		1	2	
Hambourg	**HAM**	6	7	1	6	10		2	10	
Krakow	**KRA**	5	3		3	3		2	3	
Leeds	**LEE**	1			2	3	2	2	3	3
Lisbon	**LIS**	8	4		11	4		10	4	
London	**LON**	3	3	1	5	5	3	5	5	3
Lyon	**LYO**	2	2		3	9		3	9	
Madrid	**MAD**	5	9	3	7	9	3	7	9	3
Marseille	**MAR**	4	1		4	10		1	10	
Milan	**MIL**	8	7		9	13		9	13	
Munich	**MUN**	1	5		1	10		1	10	
Naples	**NAP**		1			2			2	
Paris	**PAR**	7	6		19	11		12	11	
Prague	**PRA**	4	20	1	4	14		2	14	1
Rome	**ROM**	6	2	1	6	2		6	2	
Sevilla	**SEV**	2	2		5	3		4	3	
Sofia	**SOF**		1			1			1	
Stockholm	**STO**	1	2	1	1	2		1	2	1
Valencia	**VAL**	2	3	1	2	5	1	2	5	1
Vienna	**VIE**	3	21	1	4	19		3	19	
Warsaw	**WAR**	10			6	3		3	3	1
Total		**98**	**151**	**14**	**126**	**201**	**10**	**93**	**201**	**16**

Table 3: Parameters for the fit between the delta concentration ($\mu g/m^3$) per unit emission (t/km^2) ($\Delta C/\Delta E$) for all models. Fit equation for each model is illustrated in Figure 6.

	NO_2		PM_{10}	
	$\Delta C/\Delta E$	R^2	$\Delta C/\Delta E$	R^2
EMEP	0.26	0.75	0.35	0.70
CHIMERE	0.43	0.74	0.76	0.70
LOTOS-EUROS	0.22	0.65	0.57	0.54
RCG	0.34	0.73	0.75	0.79
CMAQ	0.24	0.46	0.30	0.23

Table 4: Models' performance for NO_2, O_3 and PM_{10} showing mean bias, root mean square error (RMSE) and temporal correlation coefficient (Corr) for 641 stations. CHIM = CHIMERE, LOTO = LOTOS-EUROS.

	model	res	Rural			Urban			Suburban		
			RMSE	bias	Corr	RMSE	bias	Corr	RMSE	bias	Corr
NO2	CHIM	EC4M1	11.79	-7.39	0.28	27.33	-21.60	0.28	23.80	-18.22	0.29
	CHIM	EC4M4	10.81	-3.15	0.31	20.59	-5.76	0.32	18.67	-4.47	0.32
	CMAQ	EC4M1	11.86	1.25	0.28	22.58	-14.36	0.27	19.43	-10.46	0.28
	CMAQ	EC4M4	11.24	-0.37	0.30	20.00	-8.73	0.30	18.47	-5.84	0.30
	EMEP	EC4M1	9.75	-2.11	0.31	23.08	-15.89	0.26	19.78	-12.31	0.28
	EMEP	EC4M4	10.53	-2.83	0.28	20.23	-10.37	0.28	18.45	-8.41	0.28
	LOTO	EC4M1	10.07	-2.44	0.27	23.03	-16.02	0.29	19.70	-12.28	0.29
	LOTO	EC4M4	10.59	-2.94	0.25	19.98	-10.06	0.27	18.12	-7.70	0.26
	RCGC	EC4M1	10.40	-2.93	0.25	23.96	-16.80	0.24	20.58	-13.10	0.24
	RCGC	EC4M4	10.97	-2.51	0.23	20.37	-7.53	0.27	18.84	-6.00	0.27
O3	CHIM	EC4M1	18.56	1.11	0.66	22.30	9.75	0.66	21.84	6.48	0.65
	CHIM	EC4M4	16.08	3.31	0.74	18.08	6.48	0.74	17.90	5.26	0.76
	CMAQ	EC4M1	18.29	5.56	0.69	23.86	14.85	0.68	22.34	11.80	0.70
	CMAQ	EC4M4	17.12	6.10	0.70	20.50	10.98	0.70	20.18	9.34	0.72
	EMEP	EC4M1	18.39	9.29	0.70	24.44	17.62	0.69	23.45	15.25	0.70
	EMEP	EC4M4	19.35	10.78	0.69	23.77	16.36	0.70	23.59	15.20	0.70
	LOTO	EC4M1	17.31	-6.30	0.69	17.21	1.46	0.69	17.51	-0.86	0.73
	LOTO	EC4M4	18.09	-7.65	0.67	17.70	-3.16	0.69	18.52	-4.84	0.72
	RCGC	EC4M1	17.92	1.18	0.67	20.84	9.85	0.67	19.58	7.14	0.70
	RCGC	EC4M4	17.60	-1.24	0.66	18.81	4.33	0.68	18.72	2.41	0.70
PM10	CHIM	EC4M1	11.93	-5.22	0.39	19.17	-12.04	0.41	17.08	-10.20	0.43
	CHIM	EC4M4	11.42	-3.31	0.38	18.17	-4.37	0.42	15.49	-5.69	0.41
	CMAQ	EC4M1	15.08	-9.93	0.31	23.03	-16.87	0.35	21.51	-16.08	0.34
	CMAQ	EC4M4	15.38	-9.78	0.32	22.53	-14.94	0.36	20.65	-14.75	0.35
	EMEP	EC4M1	13.37	-6.99	0.34	20.47	-13.30	0.37	19.30	-11.73	0.34
	EMEP	EC4M4	13.75	-7.04	0.32	20.12	-11.28	0.35	19.06	-10.65	0.32
	LOTO	EC4M1	14.81	-9.98	0.34	22.26	-15.81	0.37	20.67	-15.09	0.36
	LOTO	EC4M4	15.00	-10.14	0.32	21.38	-12.45	0.35	19.55	-13.16	0.34
	RCGC	EC4M1	15.26	-9.44	0.21	22.76	-14.30	0.24	21.54	-15.09	0.22
	RCGC	EC4M4	15.64	-10.07	0.20	23.60	-10.89	0.25	20.72	-13.85	0.23

Figures to manuscript **"Performance of European chemistry transport models as function of horizontal resolution"**

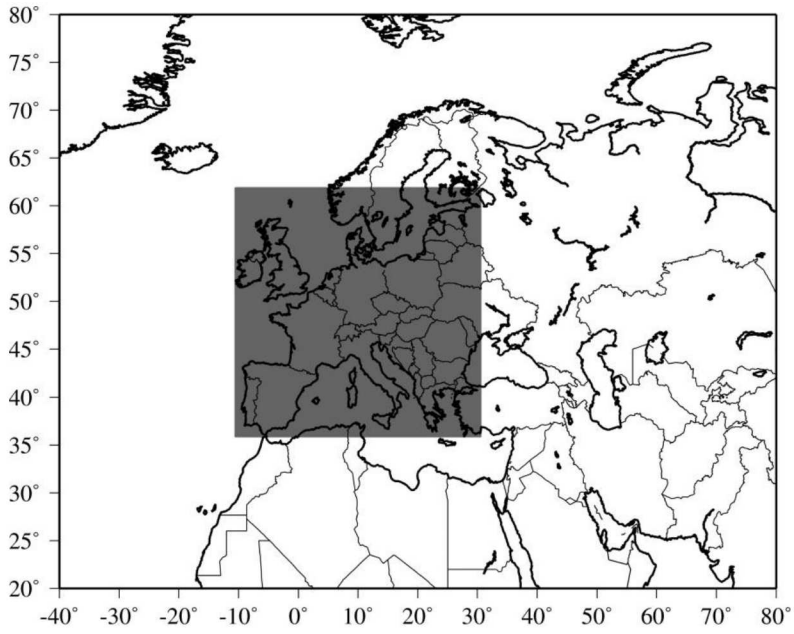


Figure 1: Common domain (grey area) used in this study at 4 resolutions.

Figure 2: Urban areas selected for the analysis. Available measurements within a radius of 30 km around each city area are selected for urban background stations. For comparison with AIRBASE rural background and EMEP stations, the selected radius around each city is 200 km (Table 2).

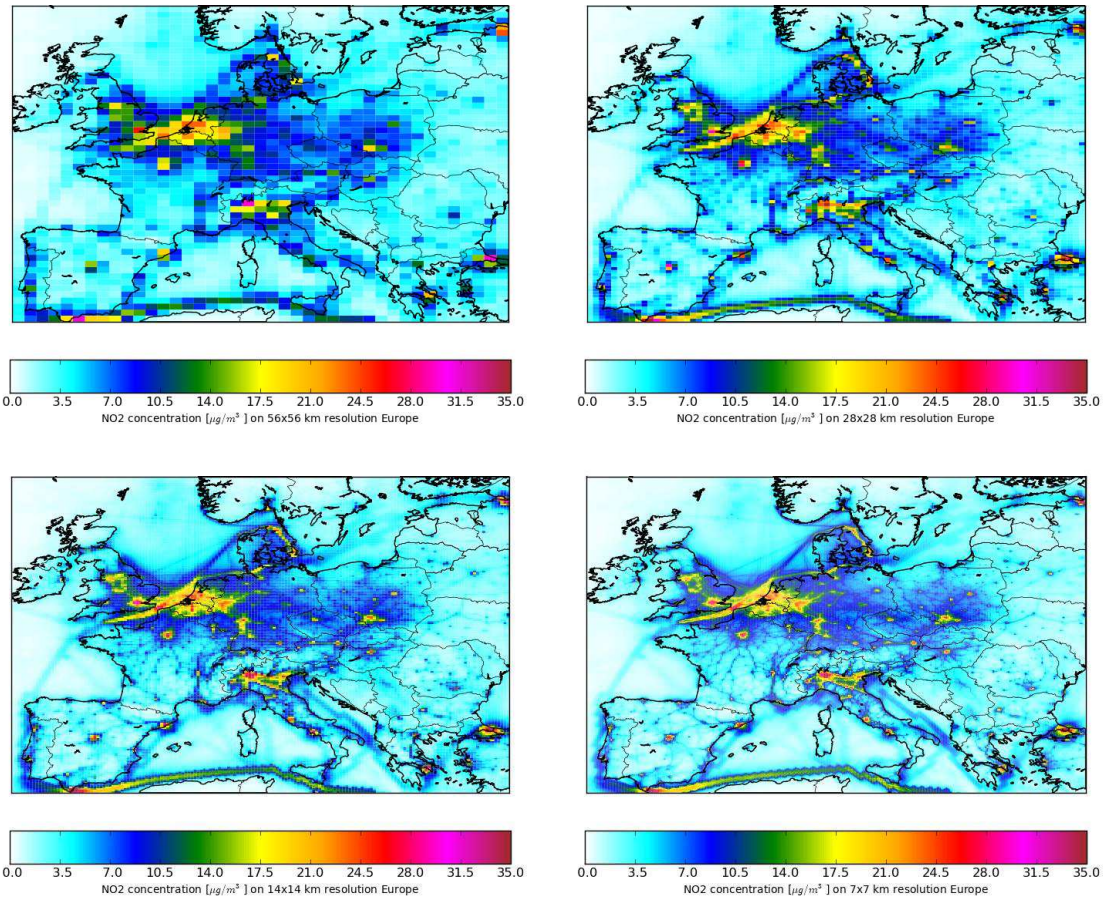


Figure 3. Modelled annual-mean NO_2 concentration distributions (µg/m^3) at 56 (top left), 28 (top right), 14 (bottom left) and 7 (bottom right) km resolution using the LOTOS-EUROS model.

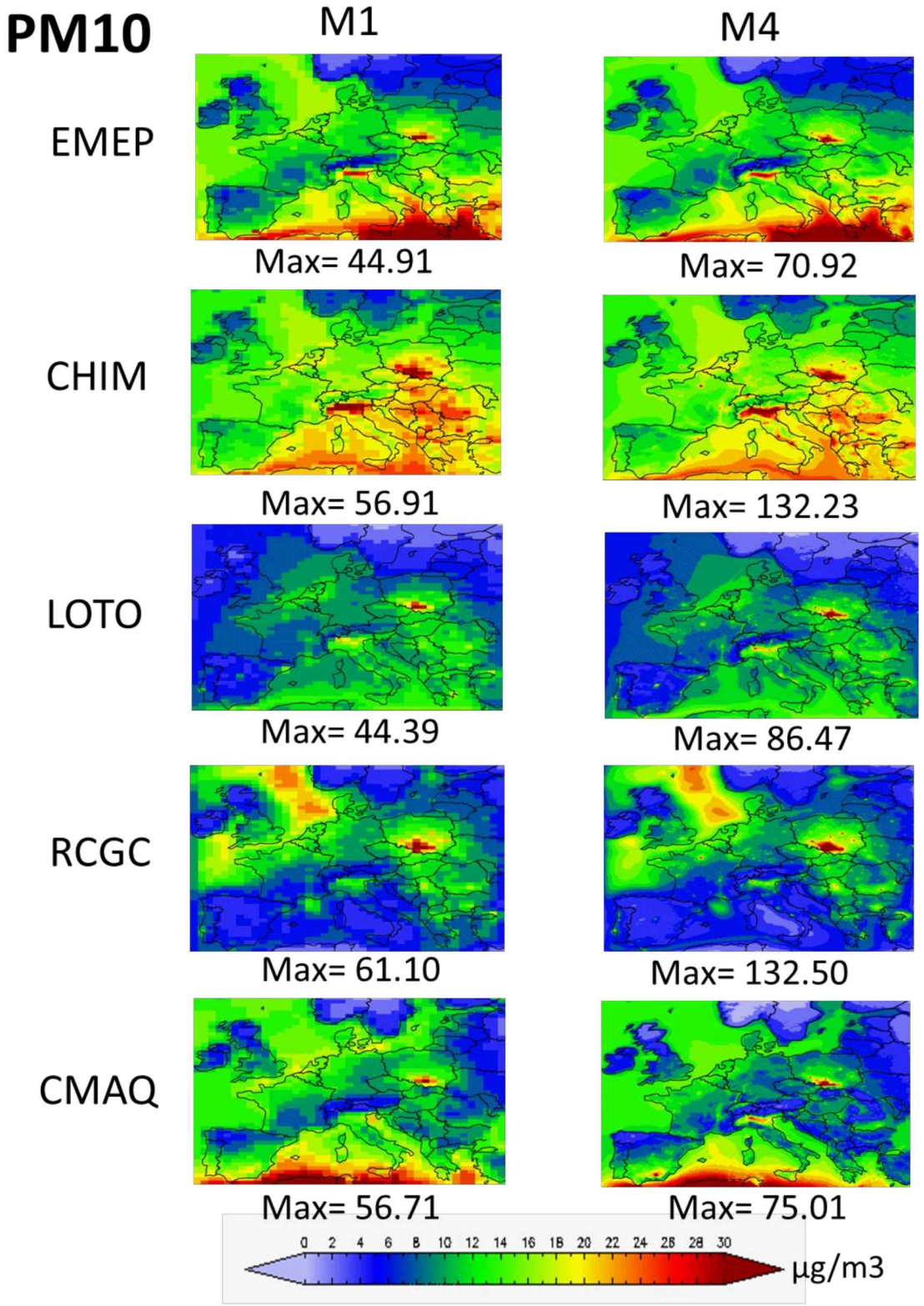
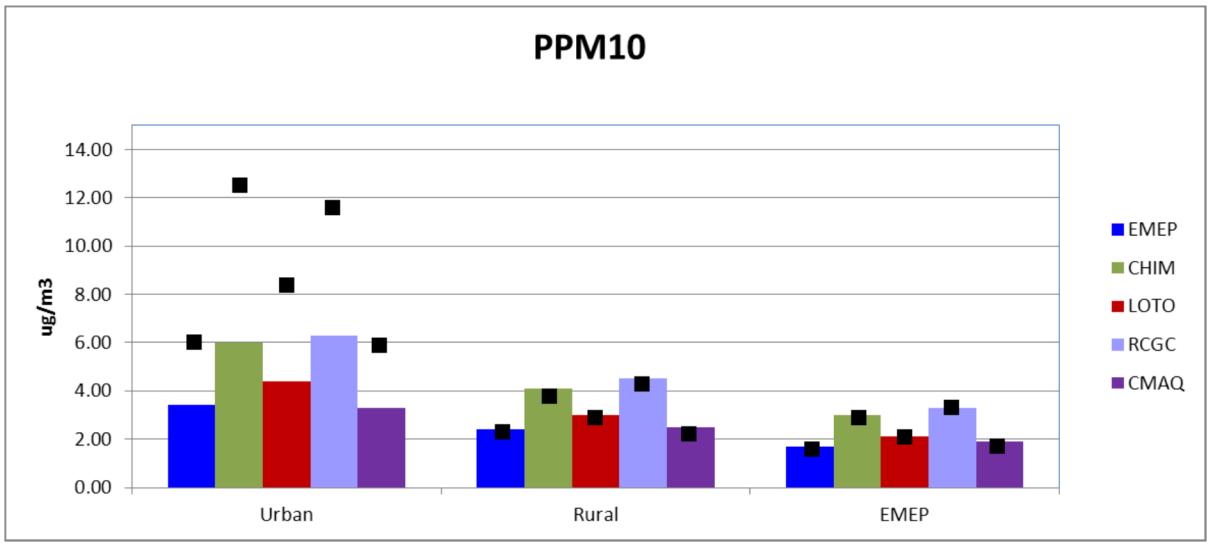
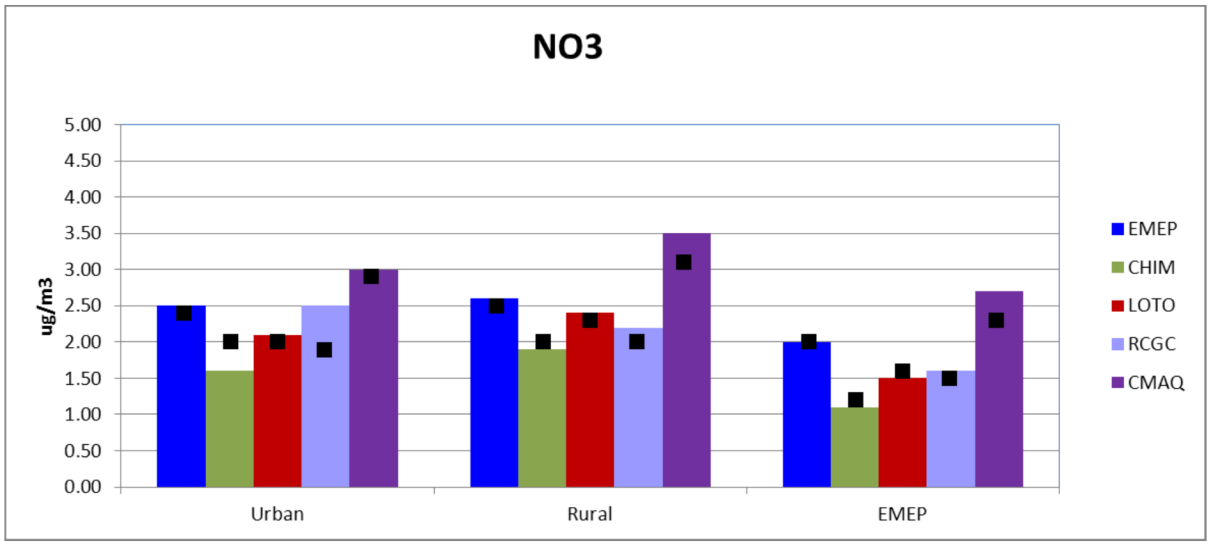


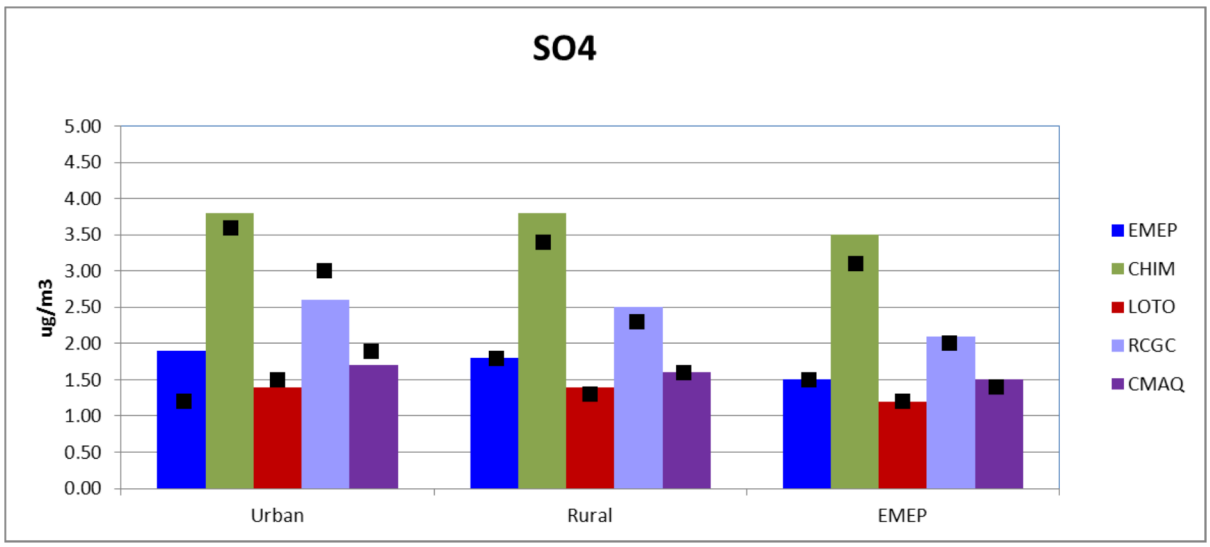
Figure 4. Modelled annual-mean PM_{10} concentration maps (µg/m^3) at ground level at 56 km and 7 km resolution for all participating models. Max indicates the maximum annual average concentration in the domain for 2009.


PPM10
14.00
12.00
10.00
8.00
6.00
4.00
2.00
0.00
ug/m3
Urban
Rural
EMEP
EMEP
CHIM
LOTO
RCGC
CMAQ

29


NO3
5.00
4.50
4.00
3.50
3.00
2.50
2.00
1.50
1.00
0.50
0.00
ug/m3
Urban
Rural
EMEP
EMEP
CHIM
LOTO
RCGC
CMAQ

30


SO4
5.00
4.50
4.00
3.50
3.00
2.50
2.00
1.50
1.00
0.50
0.00
ug/m3
Urban
Rural
EMEP
EMEP
CHIM
LOTO
RCGC
CMAQ

31

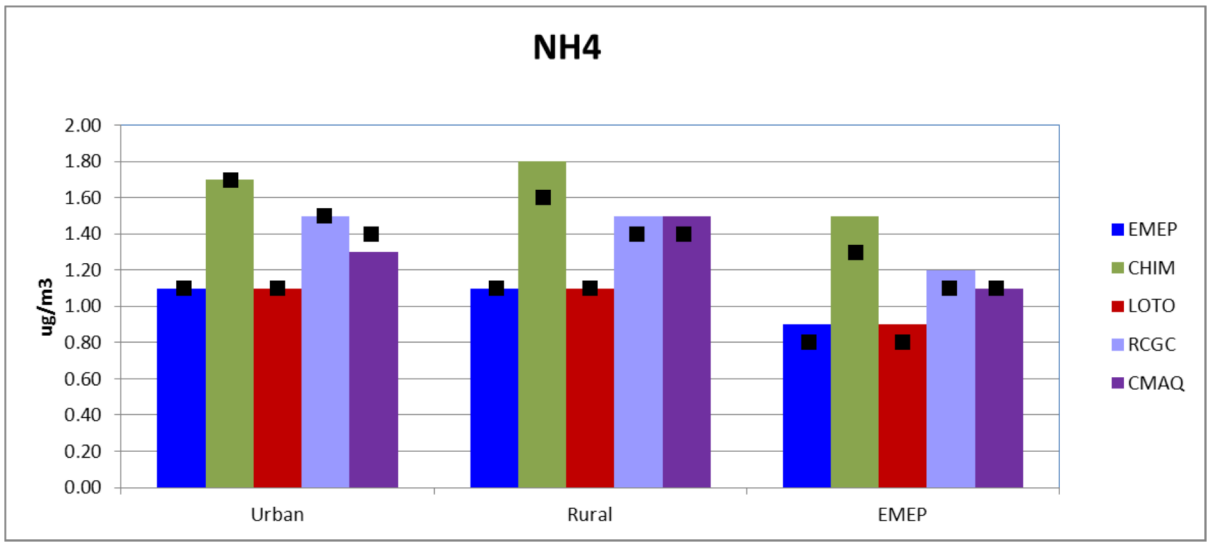


Figure 5: Scale dependency for PPM_{10} (a), nitrate (b), sulphate (c), and ammonium (d) concentration ($\mu g/m^3$) at 56-km resolution (bars) compared to 7-km resolution (black squares) for all models.

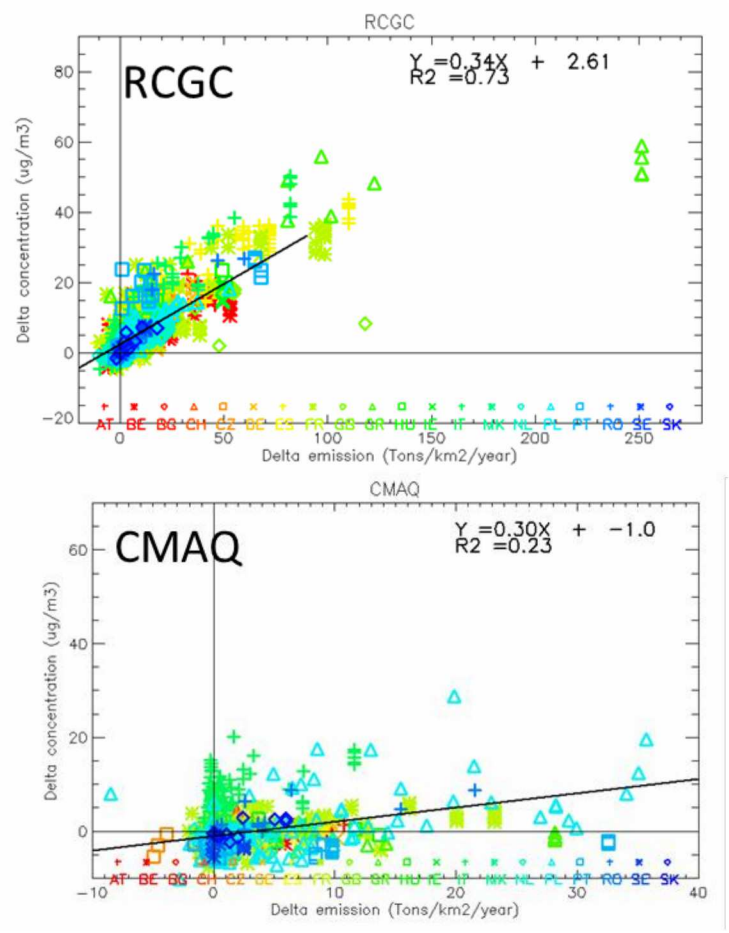


Figure 6: Concentration change (y-axis) as a function of emission delta for all cities for NO_2 by RCG (left) and PM by CMAQ (right). Each colour/symbol represents a country. AT = Austria, BE = Belgium, BG = Bulgaria, CH = Switserland, CZ = Czech Republic, DE = Germany, ES = Spain, FR = France, GB = United Kingdom, GR = Greece, HU = Hungary, IE = Ireland, IT = Italy, MK = Macedonia, NL = Netherlands, PL = Poland, PT = Portugal, RO = Romania, SE = Sweden, SK = Slovakia

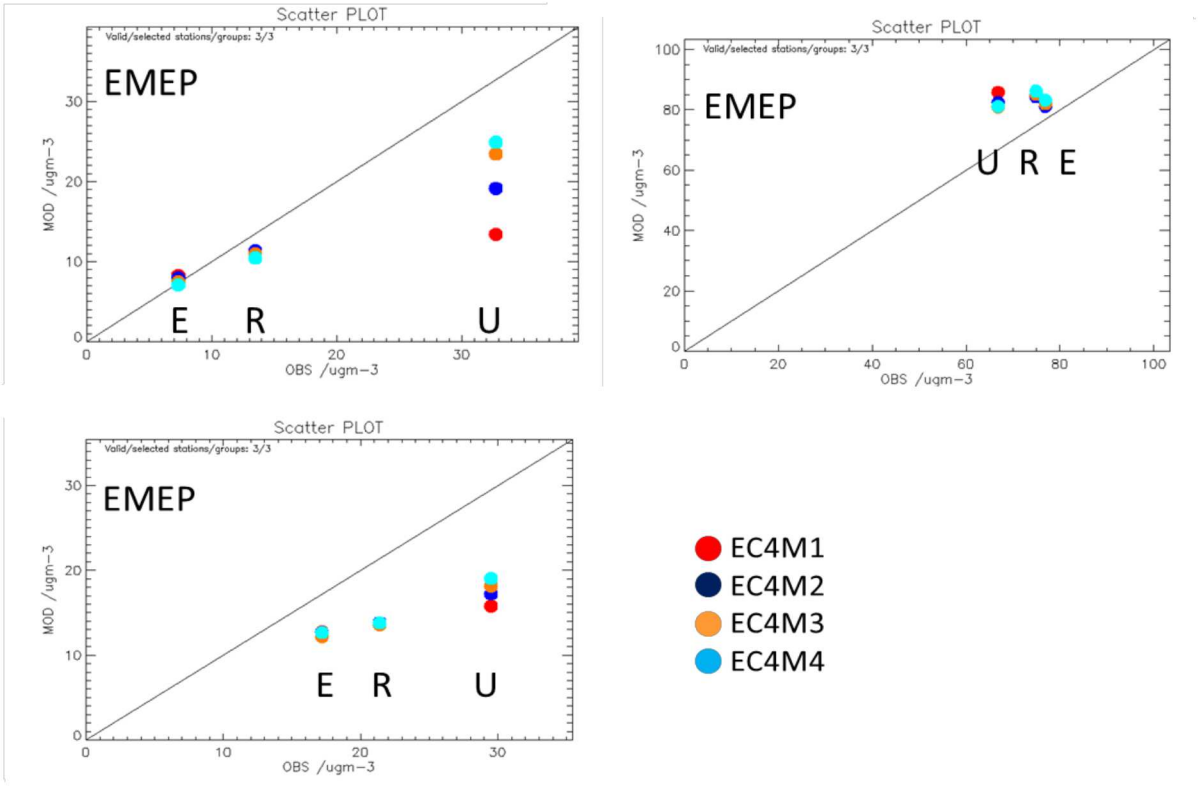


Figure 7. Comparison of annual of observed and modelled averaged NO_2 (top left), O_3 (top right) and PM_{10} (bottom left) concentrations per station type for the EMEP model at the 4 spatial resolutions. Urban (U) groups are based on the 30 km radius; EMEP (E) and Rural (R) groups are based on a radius of 200 km around the urban centers.

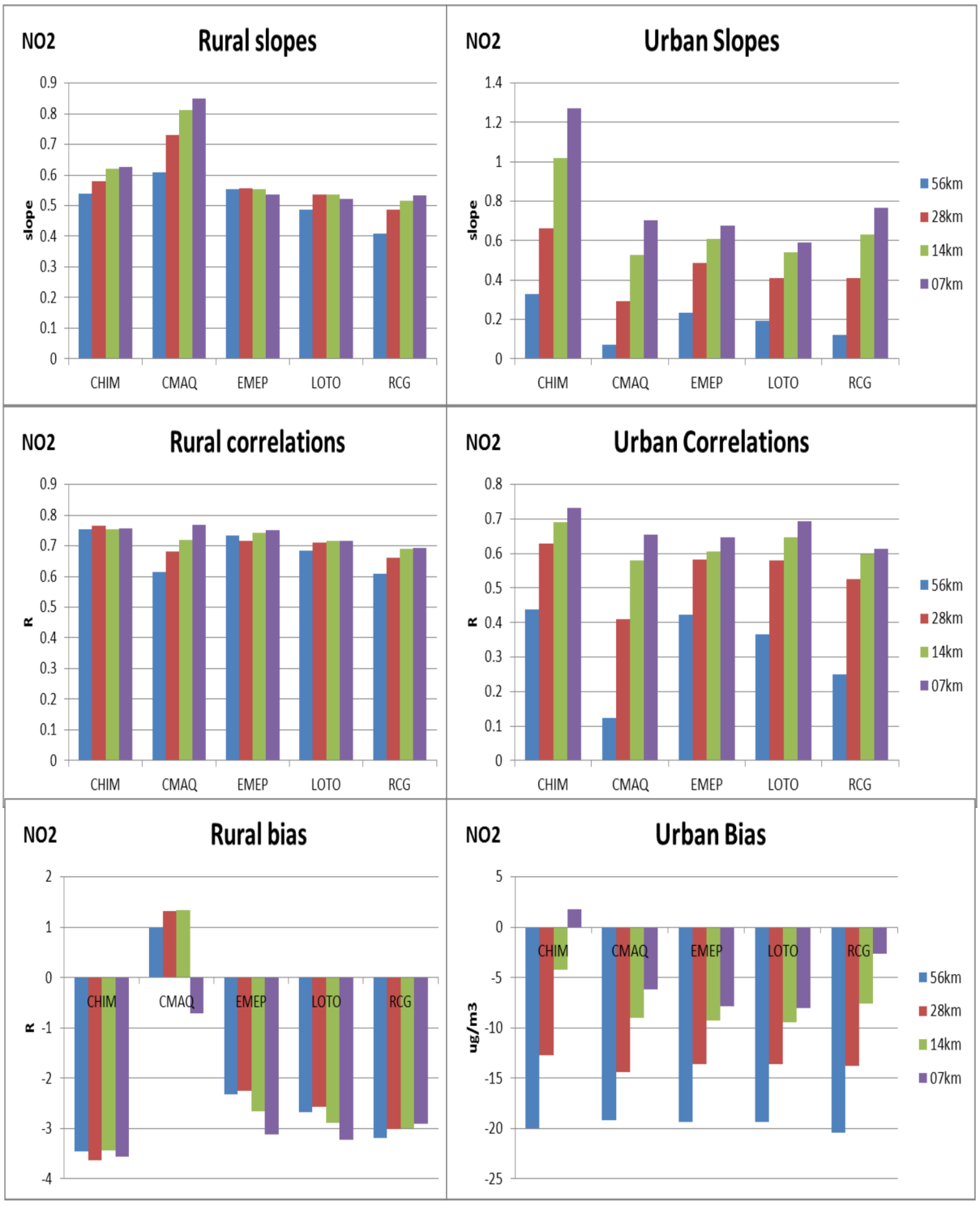


Figure 8. Summary of statistical temporal analysis for hourly NO_2 concentrations by station group, model and resolution. Note the different scales between the station groups.

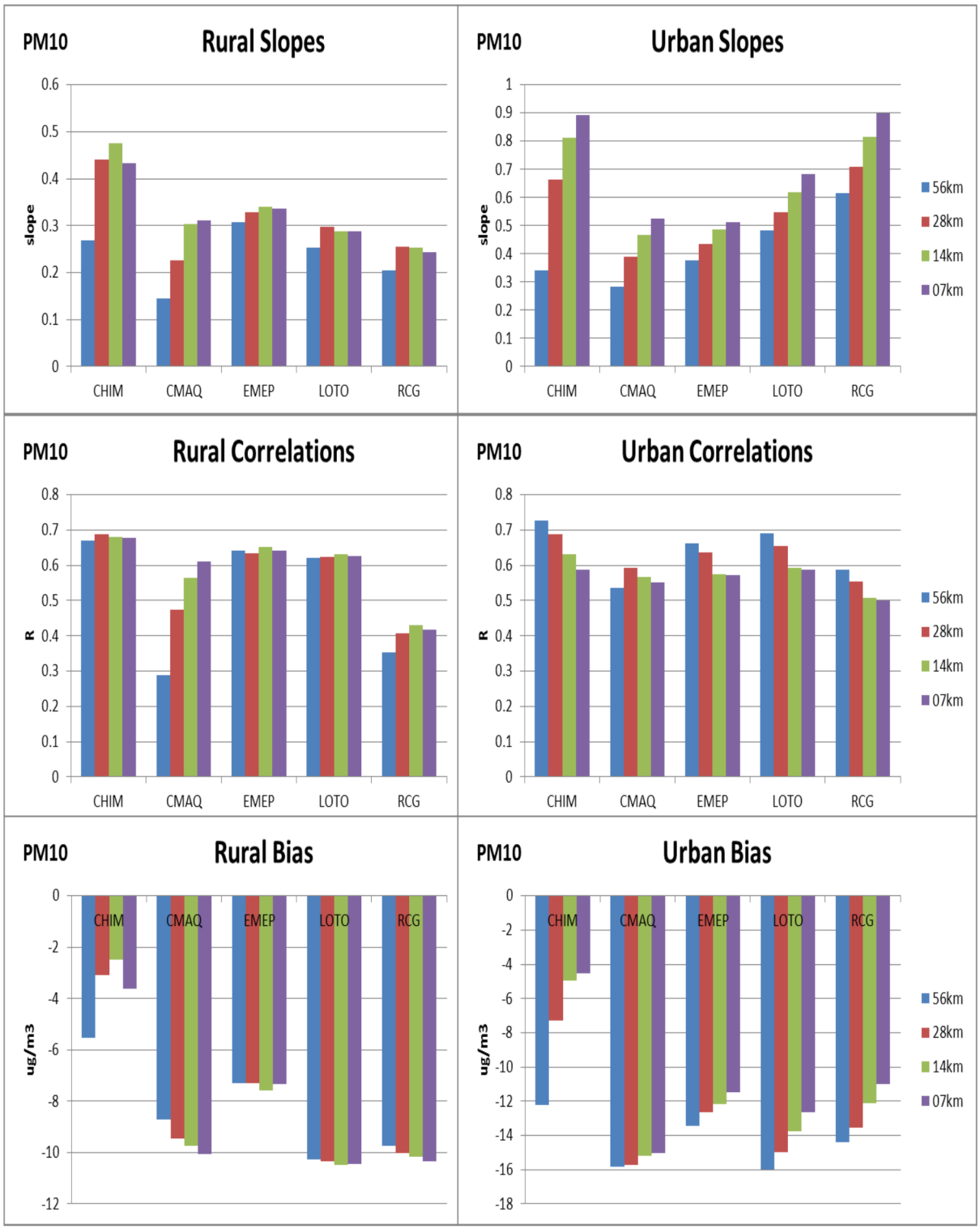


Figure 9. Summary of statistical temporal analysis for daily PM_{10} concentrations by station group, model and resolution. Note the different scales between the station groups.

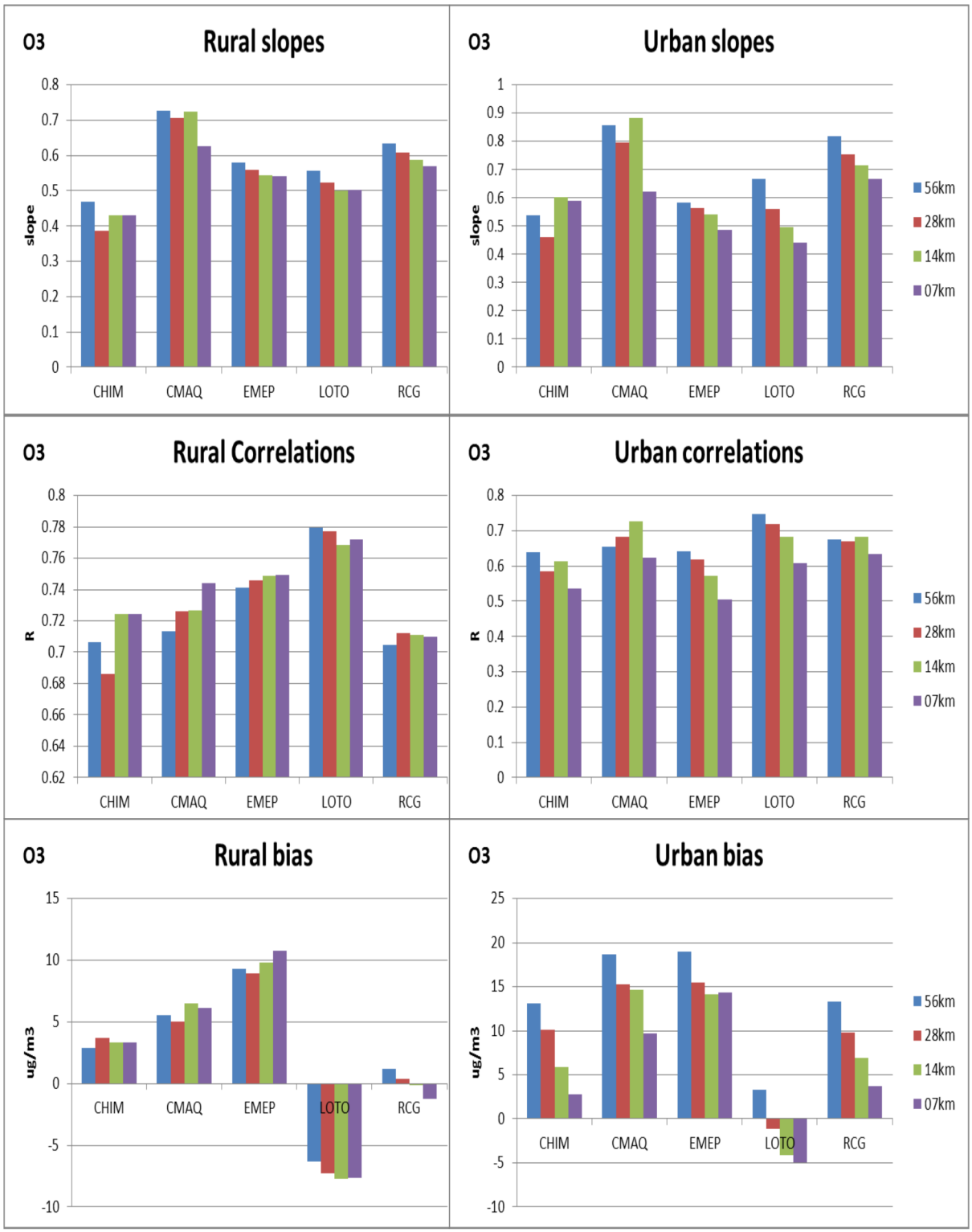


Figure 10. Summary of statistical temporal analysis for O3Max8hr concentrations by station group, model and resolution. Note the different scales between the station groups

Supplementary Material

Click here to download Supplementary Material: Supplementary material.docx