Results and Lessons Learned from Phase 1 of the Air Quality Model Evaluation International Initiative (AQMEII)

Kenneth Schere, Robert Vautard, Efisio Solazzo, Christian Hogrefe, and Stefano Galmarini

Kenneth Schere and Christian Hogrefe are both with the Atmospheric Modeling and Analysis Division of the U.S. Environmental Protection Agency. Robert Vautard is with the Laboratoire des Sciences du Climat et de l'Environnement of the Pierre Simon Laplace Institute in France. Efisio Solazzo and Stefano Galmarini are both with the Institute for Environment and Sustainability of the European Commission's Joint Research Center in Italy. E-mail: hogrefe.christian@epa.gov.

AQMEII Goals and Initial Implementation

The objective of the Air Quality Model Evaluation International Initiative (AQMEII)[1,2] is to help build a coordinated international effort on regional air quality model evaluation methodologies. It evolved from several regional air quality model evaluation workshops held in the U.S. and in Europe during 2007-2009. A new international framework for model evaluation emerged from these workshop discussions and subsequent publications that included aspects of operational, diagnostic, dynamic, and probabilistic evaluation of the regional air quality models[3]. The organizers and sponsors of the workshops began a collaboration designed to more objectively and methodically develop and demonstrate insightful performance evaluation techniques for these modeling systems.

The overall purpose of AQMEII is to coordinate international efforts in air quality modeling research and evaluation in North America (NA) and Europe (EU) to help achieve the following objectives:

- exchange expert knowledge in regional-scale air quality modeling
- identify knowledge gaps in air quality science
- develop innovative methodologies to evaluate uncertainties in air quality modeling
- build a common strategy for model development and future research priorities

- establish methodologies for model evaluation to increase knowledge on processes and to support the use of models for policy development
- initiate coordinated research projects and perform rigorous model intercomparisons

The specific goal of the first phase of AQMEII (AQMEII-1) is to perform an initial set of model evaluations and intercomparisons on existing regional air quality model systems in NA and EU, and to illustrate the use of the new model evaluation framework. To accomplish this goal, model simulations were conducted for the year 2006 for NA and EU model domains by eighteen independent groups spanning both continents, using eleven state-of-the-science regional air quality modeling systems, including various versions of individual models. The modeling systems used in AQMEII-1 are primarily off-line (i.e. uncoupled) systems in which meteorological modeling is performed upstream of the chemical-transport modeling with one-way information flow from meteorology to chemical-transport models. Performance evaluations were conducted using air quality and meteorological surface monitoring networks as well as a set of upper-air measurements obtained by commercial aircraft platforms[4]. A set of chemical boundary concentrations and source emissions for each continent was provided to all modeling groups for use as common input datasets for their air quality simulations. Meteorological fields were also made available to the modelers, although the groups were invited to generate their own meteorological inputs if they were part of their overall modeling system.

This paper provides an overview of the key model evaluation results from AQMEII-1 for meteorology, ozone and particulate matter (PM). Also, model sensitivity to input data and process parameterizations is discussed. Further, we ask the question to what extent AQMEII-1 was able to accomplish its stated goals and what lessons can be learned for future AQMEII activities.

Model Performance for Meteorological Driver Models

Air pollutant concentrations are very sensitive to the prevailing weather conditions. Thus, uncertainties in simulated meteorology driving the chemistry-transport models (CTMs) may induce large error in air quality prediction predictions. In AQMEII-1, a study[5] was dedicated to quantifying these uncertainties as simulated by the various participating meteorological models or model versions for NA and EU, and the key findings from this study are summarized in this

section. Surface observations used for comparison were extracted from the Meteorological Assimilation Data Ingest System (MADIS: http://madis.noaa.gov/) database and the National Center for Atmospheric Research (NCAR) global synoptic surface data archive (http://dss.ucar.edu/datasets/ds464.0). Simulations were also compared with profile measurements taken along ozone soundings, downloaded from the WMO World Ozone and Ultraviolet Radiation Centre (www.woudc.org). The overall focus was on variables most influencing transport and dispersion (e.g., winds, boundary layer height), chemistry (e.g., temperature, humidity) and deposition/biogenic emissions.

As illustrated in Figure 1, a comparison of modeled values with observations reveals that most models exhibit a persistent overestimation of surface wind speed. These overestimations are particularly pronounced during stable atmospheric conditions (i.e. night time and winter), and are expected to inhibit the skill of CTMs in producing large enough accumulation of pollutants, leading to the underestimation of primary pollutant levels.

Vautard et al.[5] also used observations available for a site in Germany to evaluate model simulations of the planetary boundary layer (PBL) height. They found that at that site PBL heights were well captured by all models, but transitions to stable PBLs were poorly reproduced. Short-wave radiation was found to vary considerably between models (for example, mean noontime radiation predictions were found to vary by a factor of two), most likely as a result of different cloud schemes used by the different models. As a consequence of the differences in short-wave radiation, ozone (O_3) production is expected to also significantly vary in the CTMs. Light to moderate precipitation events were found to be generally overestimated, leading to a probable excess of removal of particles through wet deposition. Temperature was found to be fairly well simulated. These results provide hints for the attribution of biases and other uncertainties in atmospheric pollutant concentrations. For instance, wind overestimation was found to correlate with particulate matter bias.[6]

Air Quality Model Performance for Ozone

A number of groups participating in AQMEII-1 have compared the O_3 mixing ratios predicted by their models against available observations in NA and EU and described the results of their

analyses in a series of articles[7-13]. In addition, Solazzo et al.[14] performed a multi-model intercomparison for O_3 to assess the statistical properties of the ensemble of models relative to the individual model simulations for a range of air quality cases. To perform this multi-model analysis, eleven model simulations from different modeling groups were available for the EU domain, and eight simulations were available for the NA domain. Fourteen different air quality models, and variations of the same model, were used to perform these nineteen simulations.

Ground-level O_3 monitoring network data at regionally-representative non-urban sites were used in the operational model evaluation summarized here and discussed in more detail in Solazzo et al[14]. Urban monitoring data for O_3 were excluded from the analysis because they often demonstrate local chemical effects from reactions with NO that are not well resolved by the regional-scale model configurations used here. The number of sites used in the model evaluation analyses was 640 in EU and 568 in NA.

It is beyond the scope of this overview article to provide an in-depth discussion of the model evaluation results reported by the individual groups participating in AQMEII-1[7-13]. Instead, we highlight several key aspects emerging from the multi-model intercomparison study performed by Solazzo et al. To this end, Figure 2 presents summertime average diurnal cycles of O_3 for NA and EU, computed by taking the mean for each hour across all sites and all days between May and September. Several features are apparent in these figures. First, there is substantial divergence between the model results on both continents; this divergence is more pronounced over the EU domain which had a greater number of participating groups than the NA domain. Among the factors influencing the bias and variability in the model O_3 estimates are (1) variability in biogenic O_3 precursor emissions estimated by different algorithms within the air quality models themselves, and (2) differences in the driving meteorological fields among the different model simulations, especially overestimates of surface and boundary layer wind speeds, most notably in EU simulations. Second, the simulations for the EU domain tend to underestimate peak daytime O_3 mixing ratios and overestimate nighttime mixing ratios. For the NA domain, most models tend to overestimate observed O_3 mixing ratios throughout the entire diurnal cycle. Third, the diurnal range (i.e. the difference between daily minimum and maximum) tends to be underestimated by most models. Since the diurnal range is indicative of

the relative strengths and interactions of the processes that control the formation, destruction, and transport of O_3, a closer analysis of the diurnal O_3 cycle can potentially provide insights into how well the individual models treat specific atmospheric processes. For example, the overestimation of nighttime O_3 mixing ratios by most models may indicate that these models have difficulties in capturing the effects of a stable nighttime boundary layer that leads to the observed depletion of surface O_3 through titration with NO and dry deposition. The poor representation of the stable nighttime boundary layer might be due to either the vertical resolution used by the modeling systems and/or the parameterization of vertical mixing. Similarly, most models underestimate the increase of O_3 mixing ratios from its value at 10 am local time to its peak. Since this is the time period when photochemical production of O_3 it at its highest, this underestimation may be indicative of shortcomings in the treatment of atmospheric chemistry by these models or errors in timing of emissions injection.

The multi-model ensemble analysis performed for O_3 by Solazzo et al.[14] showed that using the mean or median of the ensemble distribution improved model performance relative to most individual model simulations. However, it was also found that the most skillful ensemble is not necessarily generated by including all available model results, but rather by selecting models that result in a minimization of ensemble error. In addition, an ensemble of only top-ranking model results may perform worse than an ensemble comprising both top-ranking and low-ranking model results, as the inclusion of outliers can enhance overall performance of the ensemble and decrease the level of inter-dependence of model results in the ensemble. Finally, Solazzo et al.[14] showed that the objective selection of ensemble members based on a clustering analysis using a statistical metric, such as correlation among model results, helps to optimize the construction of the best performing ensemble.

Air Quality Model Performance for Particulate Matter

Similar to O_3, operational model evaluation for ground-level PM (PM_{10} and $PM_{2.5}$) concentrations was carried out for the 2006 simulations by many of the participating modeling groups and was reported in a series of journal articles[7-12]. In addition, a multi-model comparison using ten air quality simulation models, and variants of these models, over EU and NA was also performed[6] and a summary of the key findings from that study is presented in this section. For

this analysis, ten model simulations were available for the EU domain and seven model simulations were available for the NA domain. Observations of PM_{10} and $PM_{2.5}$ mass for 2006 were available from as many as 863 surface monitoring stations in EU, and 1902 stations in the U.S. and Canada. In addition, chemically-speciated $PM_{2.5}$ data from eighteen non-urban NA monitoring stations were used for more detailed analyses.

Figures 3a and 3b present the observed and simulated annual average PM_{10} concentrations over NA and EU. A pattern common to PM_{10} model performance on both continents is a general underestimation by all models, with model predictions often exceeding a mean fractional error of 75%. This underestimation is particularly pronounced for NA and may be due to the lack of source emissions information on wind-blown dust. Support for this hypothesis comes from the fact that the most severe underprediction of PM_{10} in NA was found over the Southwestern U.S.[6] where dust is a major component of PM_{10} mass. Large underpredictions of PM_{10}, although not as large as those for NA, are also seen over the EU domain where wind-blown dust was included in the AQMEII emissions data. Seasonal differences in the magnitude of the bias were also observed, with the wintertime showing larger underpredictions. High concentrations of PM_{10} in winter often result from a strongly stabilized atmosphere that is not well simulated by the meteorological models. As discussed earlier and analyzed in greater detail in Vautard et al.[5], the meteorological models used in AQMEII-1 have a tendency to overestimate the 10 m wind speed (especially for EU), likely contributing to the negative bias for PM_{10} concentration predictions. In support of this hypothesis, Solazzo et al.[6] showed that the PM_{10} bias for the EU simulations had a marked dependence on wind speed. Moreover, Figures 3a and 3b illustrate that there is a significant variability between the models in terms of predicted annual average PM_{10} concentrations for both domains. This can be explained in part by differences in the components of PM_{10} included in the models' emissions and chemistry modules. For example, not all models included sea salt emissions or wind-blown dust, while others had no production of secondary organic aerosols.

Model performance for annual average $PM_{2.5}$ is illustrated in Figures 3c and 3d. These results, based on averaging 24-h data, show biases that are much less than PM_{10} for both continents, demonstrating an enhanced capability for the models to simulate $PM_{2.5}$. For EU, most models

underestimate annually-averaged daily $PM_{2.5}$ concentrations; this is also true when performing the analysis for specific sub-regions[6]. For NA, most models also tend to underestimate $PM_{2.5}$ concentrations. Examining the diurnal cycle of $PM_{2.5}$, the amplitude of the daily variability is generally underestimated by most models in the EU simulations. The daily variability is better reproduced in the NA model simulations and correlation with observations is higher, most likely indicating better PM emission datasets in NA than EU[6].

Shorter episodic analyses for $PM_{2.5}$ concentrations are also possible on both continents. Results show that the models seem to be able to simulate the episode peaks and the sharp oscillations around them driven by transport phenomena[6]. Investigation of the chemical components shows that the sum of the inorganic species was generally better reproduced by the models than total $PM_{2.5}$. Hence, Solazzo et al.[6] concluded that other $PM_{2.5}$ components, such as organic aerosols, can be modeled with less accuracy than inorganic ones.

Model Sensitivity to Input Data and Process Parameterizations

Reference or common datasets containing source emissions and chemical boundary concentrations were provided to all modeling groups in AQMEII-1. While the majority of groups did make use of these reference datasets, not all participants did. The motivation for harmonizing on these model input parameters was to facilitate the interpretation of differences in model results in the intercomparison exercise. In concept, differences in the results could then be attributable to other data influences or different model process representations. In practice however, this proved to be an elusive goal. Even for modeling groups that made use of these reference datasets, there were differences in the biogenic and dust emissions among the models, as these data are generally calculated dynamically within the modeling system during the simulation. For those models that seemingly used the same meteorological inputs, each model's preprocessing of these data made small perturbations in the parameter fields that the air quality model actually used. Thus, while harmonization of model inputs can help with the interpretive analysis of such an intercomparison exercise, the implementation of such harmonization in complex modeling systems is very difficult. The resulting freedom of practice of each modeling group enabled them to produce the best possible results from their models based on native model system protocols, but the confounding influence of different model inputs as well as different

model process representations hampered the diagnosis of causes for differences among the model results.

For those models using the prescribed emissions and boundary concentrations, it is possible to assess the sensitivity to these inputs. Schere et al.[8] found that the boundary concentration profile specification for O_3 was very influential on model results far into the interior of the model simulation domain. This was especially true for the winter months and for rural areas away from major emission sources. Systematic underestimates of tropospheric O_3 by the global modeling system that was used to derive the boundary concentrations for AQMEII-1 modeling caused the regional-scale models to often underestimate NA and EU near-surface O_3 concentrations. On the other hand, the global modeling system was informed by satellite data assimilation for some trace gases and aerosols, helping to accurately estimate the stratospheric O_3 profile as well as the plumes of large wildfires and transported dust on the regional air quality model boundaries. Figure 4 illustrates the aerosol plume of a large Canadian wildfire, detected in the global model data assimilation, transported into an AQMEII regional model simulation domain.

Large differences were observed among the models in the dry deposition of $PM_{2.5}$ component species on both continents, despite the general use of similar deposition parameterization schemes. However, dry deposition is very sensitive to local land-cover and topographical conditions as well as to near-surface meteorology and turbulence, which can vary among models. The wet removal of soluble ions (sulfate, nitrate, ammonium) depends strongly on the characterization of precipitation. Since the meteorological models tended to overestimate seasonal precipitation, especially for convective conditions[5], it is reasonable to conclude that wet deposition of these species is also overestimated.

<u>Did AQMEII-1 Achieve Its Goals?</u>

Many of the overarching goals of AQMEII, as listed earlier, were indeed achieved in this first phase of the program. AQMEII-1 resulted in an unprecedented collaboration among regional air quality modelers in NA and EU. A comprehensive database of surface and aloft meteorological and air quality observations over both continents was assembled for the year 2006. The diversity of the information included and the high level of data harmonization and accessibility makes

such a database unique in nature. Common sets of source emissions and boundary concentrations were prepared for all modelers through the efforts of teams of scientists at EPA (U.S.), TNO (Netherlands), and ECMWF (EU). Model simulations for 2006 were performed by eighteen independent modeling groups and all model and observational data were archived for analysis at the European Commission's Joint Research Centre ENSEMBLE model evaluation platform[9] in Italy.

Model evaluations of individual models and cross-model comparisons have been carried out and analyses documented through AQMEII workshops and peer-reviewed publications. These collaborative efforts by many scientists and support staff across NA and EU resulted in a highly successful and productive AQMEII-1 program over a relatively short two-year period. The goals of exchanging expert knowledge among the modeling groups and identifying some of the significant modeling science gaps were achieved. A very efficient methodology has been established that allows the set up of a case study, collection of input information, collection of model results and analysis to be performed in a relatively short time. The methodology can be replicated and will be used again for future AQMEII work. AQMEII-1 made a strong start on demonstrating the operational component of the motivating model evaluation framework. The ensemble model analyses for O_3 also provided a glimpse of a probabilistic model evaluation technique. However, if there are areas that fell short of achieving the stated goals they are in the complementary framework components of diagnostic and dynamic model evaluation.

The initial modeling protocol for AQMEII-1 included two full years of model simulations spanning the time of significant emissions change for NA and EU caused by regulatory control programs. Comparison of two such years would enable assessing the models' ability to detect a change in the air quality signals from these emissions changes. A subsequent acceleration in the timetable resulted in eliminating the second year of simulation. This meant that only limited aspects of dynamic evaluation (such as comparing the observed and modeled effects of weekday/weekend emission differences or changing weather patterns on ambient pollutant concentrations) can be addressed with the dataset generated by AQMEII-1; such analyses are currently on-going. The accelerated schedule also made it difficult to complete more in-depth diagnostic analyses of the model results beyond basic operational statistical measures and

graphical analysis. Such analyses are continuing and will be reported by individual modeling groups as the results emerge. Diagnostic testing across models was hampered, as stated earlier, by the difficulty in harmonizing model data and systems to a sufficient degree. The model inter-comparison, however, has provided feedback to model developers regarding some aspects of their modeling systems that need improvement. Diagnostic testing by individual modelers within AQMEII on their own modeling systems is continuing and results are being reported in journals. Also, the database created from the AQMEII-1 effort is being made available to the scientific community to help develop innovated model evaluation techniques[17].

In the next phase of AQMEII (AQMEII-2), attention will focus on fully integrated meteorological and air quality modeling systems[18]. It is instructive to learn from the strengths and weaknesses discovered in AQMEII-1 to improve the outcomes of the second phase. Key lessons to consider for the path forward include:

- *Adequate time must be allowed* for the participating modeling groups to not only perform their model simulations but also to carry out full and comprehensive analyses of results.
- More focus must be given to the aspects of the model evaluation framework that were not adequately demonstrated in AQMEII-1, namely, *diagnostic and dynamic evaluation*.
- A careful analysis on the *degree of input data harmonization* should be performed based on the expected types of analyses for model inter-comparisons in AQMEII-2.
- *International collaborations and effective communications* are essential for the rapid advancement of the science of air quality model development and evaluation.

Based upon the significant accomplishments of the first phase of AQMEII, much is anticipated from the next phase that will continue to advance the state of air quality modeling science in NA and EU.

Acknowledgments

We gratefully acknowledge the contributions of various groups to the first Air Quality Model Evaluation International Initiative (AQMEII) activity. The following agencies have prepared the datasets used in the preparation phase of this study: U.S. EPA (North American emissions processing and gridded meteorology); U.S. EPA, Environment Canada, Mexican Secretariat of the Environment and Natural Resources (Secretaría de Medio Ambiente y Recursos Naturales-

SEMARNAT), and National Institute of Ecology (Instituto Nacional de Ecología-INE) (North American national emissions inventories); TNO (European emissions processing); Laboratoire des Sciences du Climat et de l'Environnement, IPSL, CEA/CNRS/UVSQ (gridded meteorology for Europe); and ECMWF/GEMS project and Météo-France/CNRM-GAME (Chemical boundary conditions). Ambient North American concentration measurements were extracted from Environment Canada's National Atmospheric Chemistry Database (NAtChem) PM database and provided by several U.S. and Canadian agencies (AQS, CAPMoN, CASTNet, IMPROVE, NAPS, SEARCH, and STN networks); North American precipitation-chemistry measurements were extracted from NAtChem's precipitation-chemistry database and were provided by several U.S. and Canadian agencies (CAPMoN, NADP, NBPMN, NSPSN, and REPQ networks); the WMO World Ozone and Ultraviolet Data Centre (WOUDC) and its data contributing agencies provided North American and European ozonesonde profiles; NASA's Aerosol Robotic Network (AeroNet) and its data-contributing agencies provided North American and European AOD measurements; the MOZAIC Data Centre and its contributing airlines provided North American and European aircraft takeoff and landing vertical profiles. For European air quality data, the EMEP European Environment Agency/European Topic Center on Air and Climate Change/AirBase provided European air- and precipitation-chemistry data. Data from meteorological station monitoring networks were provided by NOAA and Environment Canada (for the U.S. and Canadian meteorological network data) and the National Center for Atmospheric Research (NCAR) data support section. Joint Research Center Ispra/Institute for Environment and Sustainability provided its ENSEMBLE system for model-output harmonization and analyses and evaluation. Note that not all the networks listed above were used for the results presented in this article.

The views expressed here are those of the authors and do not necessarily reflect the views and policies of the U.S. Environmental Protection Agency or any other organization participating in the AQMEII project. This manuscript has been subjected to U.S. EPA review and approved for publication.

References

1. Rao, S.T.; Galmarini, S.; Puckett, K. Air Quality Model Evaluation International Initiative (AQMEII); *Bulletin of the American Meteorological Society* **2011**, *92*, 23-30.

2. Galmarini, S.; Steyn, D.G.; Schere, K.L.; Moran, M.D. Advancing the Evaluation of Regional-Scale Air Quality Models; JRC European Commission Report (EUR24245EN) **2010**, Available from http://bookshop.europa.eu.
3. Dennis, R.; Fox, T.; Fuentes, M.; Gilliland, A.; Hanna, S.; Hogrefe, C.; Irwin, J.; Rao, S.T.; Scheffe, R.; Schere, K.; Steyn, D.; Venkatram, A. A Framework for Evaluating Regional-Scale Numerical Photochemical Modeling Systems; *Environmental Fluid Mechanics* **2010**, *10*, 471-489.
4. Volz-Thomas, A.; Cammas, J.-P.; Brenninkmeijer, C.; Machida, T.; Cooper, O.; Sweeney, C.; Waibel, A. Civil Aviation Monitors Air Quality and Climate; *Environmental Manager* **2009**, *October*, 16-19.
5. Vautard, R.; et al. Evaluation of the Meteorological Forcing Used for the Air Quality Model Evaluation International Initiative (AQMEII) Air Quality Simulations; *Atmospheric Environment* **2012**, in press
6. Solazzo, E.; et al. Operational Model Evaluation for Particulate Matter in Europe and North America in the Context of the AQMEII Project; *Atmospheric Environment* **2012**, in press
7. Appel, K.W., Chemel, C., et al., Examination of the Community Multiscale Air Quality (CMAQ) model performance for North America and Europe for the AQMEII project. Atmospheric Environment. **2012**, doi:10.1016/j.atmosenv.2011.11.016.
8. Brandt, J., Silver, J.D., et al., 2012. An integrated model study for Europe and North America using the Danish Eulerian hemispheric model with focus on intercontinental transport of air pollution. Atmospheric Environment, **2012**, in press
9. Nopmongcol, U., Koo, B., Tai, E., Jung, J., Piyachaturawat, P. Modeling Europe with CAMx for the Air Quality Model Evaluation International Initiative (AQMEII). Atmospheric Environment. **2012**, doi:10.1016/j.atmosenv.2011.11.023.
10. Sartelet, K., Couvidat, F., Seigneur, C., Roustan, Y., 2012. Impact of biogenic emissions on air quality over Europe and North America. Atmospheric Environment, **2012**, doi:10.1016/j.atmosenv.2011.10.046.
11. Forkel, R., J. Werhahn, A. B. Hansen, S. McKeen, S. Peckham, G. Grell, P. Suppan, Effect of aerosol-radiation feedback on regional air quality – A case study with WRF/Chem, Atmospheric Environment, **2012**, 10.1016/j.atmosenv.2011.10.009.

12. Pirovano, G., A. Balzarini, B. Bessagnet, C. Emery, G. Kallos, F. Meleux, C. Mitsakou, U. Nopmongcol, G.M. Riva, G. Yarwood, Investigating impacts of chemistry and transport model formulation on model performance at European scale, Atmospheric Environment, **2012**, 10.1016/j.atmosenv.2011.12.052.
13. Ferreira, J. A. Rodriguez, A. Monteiro, A.I. Miranda, M. Dios, J.A. Souto, G. Yarwood, U. Nopmongcol, C. Borrego, Air quality simulations for North America - MM5–CAMx modelling performance for main gaseous pollutants, Atmospheric Environment, **2012**, 10.1016/j.atmosenv.2011.10.020.
14. Solazzo, E.; et al. Model Evaluation and Ensemble Modelling of Surface-Level Ozone in Europe and North America in the Context of AQMEII; *Atmospheric Environment* **2012**, in press
15. Schere, K.; et al. Trace Gas/Aerosol Boundary Concentrations and Their Impacts on Continental-Scale AQMEII Modeling Domains; *Atmospheric Environment* **2012**, in press
16. Galmarini, S.; Bianconi, R.; Appel, W.; Solazzo, E.; Mosca, S.; Grossi, P.; Moran, M.; Schere, K.; Rao, S.T. ENSEMBLE and AMET: Two Systems and Approaches to a Harmonized, Simplified and Efficient Facility for Air Quality Models Development and Evaluation; *Atmospheric Environment* **2012**, in press
17. Galmarini, S., and S.T. Rao. The AQMEII Two-Continent Regional Air Quality Model Evaluation Study: Fueling Ideas with Unprecedented Data. *Atmospheric Environment*, 2464, **2011**, 2464
18. Alapaty, K., Mathur, R., Pleim, J., Hogrefe, Ch., Rao, S. T., Ramaswamy, V., Galmarini, S., Schaap, M., Vautard, R., Makar, P., Baklanov, A., Kallos, G., Vogel, B., and Sokhi, R. New Directions: Understanding Interactions of Air Quality and Climate Change at Regional Scales, *Atmospheric Environment*, doi:10.1016/j.atmosenv.2011.12.016, **2012**, 419–421.

Figure Captions

Figure 1. Mean diurnal cycle of simulated (colored lines) and observed (black line) wind speed for a) North America, summertime, b) Europe, summertime, c) North America, wintertime, and d) Europe, wintertime. Please note that the numbering of the model predictions is assigned randomly for each continent and each variable (wind speed in Figure 1, ozone in Figure 2, and PM in Figure 3). In addition, the number of model simulations available for a particular continent is different for the different parameters analyzed in Figures 1 – 3.

Figure 2. Mean diurnal cycle of simulated (colored lines) and observed (black line) May – September ozone for a) North America, and b) Europe. Please note that the numbering of the model predictions is assigned randomly for each continent and each variable (wind speed in Figure 1, ozone in Figure 2, and PM in Figure 3). In addition, the number of model simulations available for a particular continent is different for the different parameters analyzed in Figures 1 – 3.

Figure 3. Annual average concentrations of simulated (colored lines) and observed (black line) a) PM_{10} over North America, b) PM_{10} over Europe, c) $PM_{2.5}$ over North America, and d) $PM_{2.5}$ over Europe. Please note that the numbering of the model predictions is assigned randomly for each continent and each variable (wind speed in Figure 1, ozone in Figure 2, and PM in Figure 3). In addition, the number of model simulations available for a particular continent is different for the different parameters analyzed in Figures 1 – 3.

Figure 4. North American regional model-predicted average primary organic aerosol on 30 June 2006 at 01 UTC using global model-derived boundary concentrations.

Figure 1.

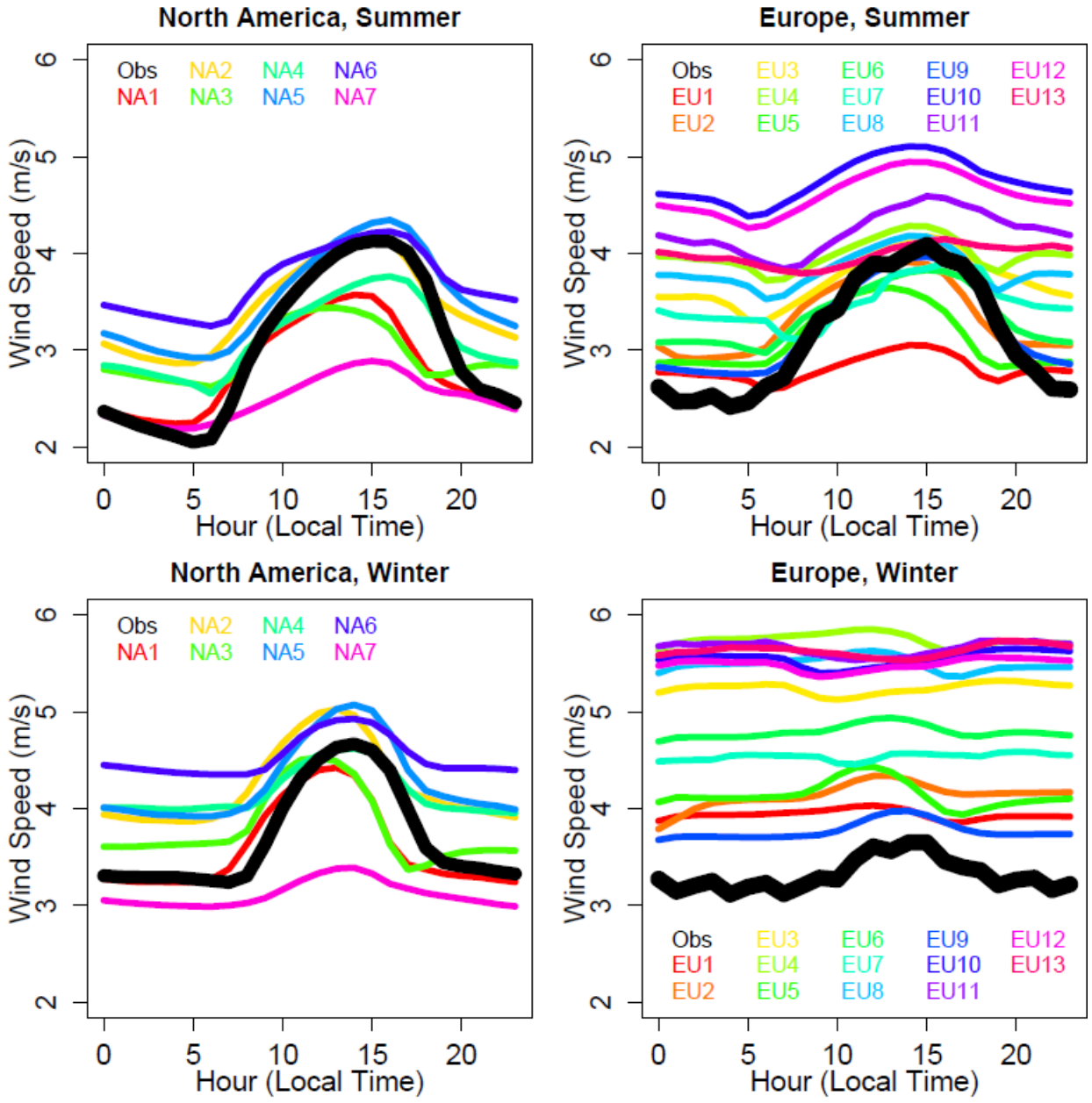
North America, Summer
Europe, Summer
North America, Winter
Europe, Winter
Obs
NA1
NA2
NA3
NA4
NA5
NA6
NA7
EU1
EU2
EU3
EU4
EU5
EU6
EU7
EU8
EU9
EU10
EU11
EU12
EU13
Wind Speed (m/s)
Hour (Local Time)
2
3
4
5
6
0
5
10
15
20

Figure 2.

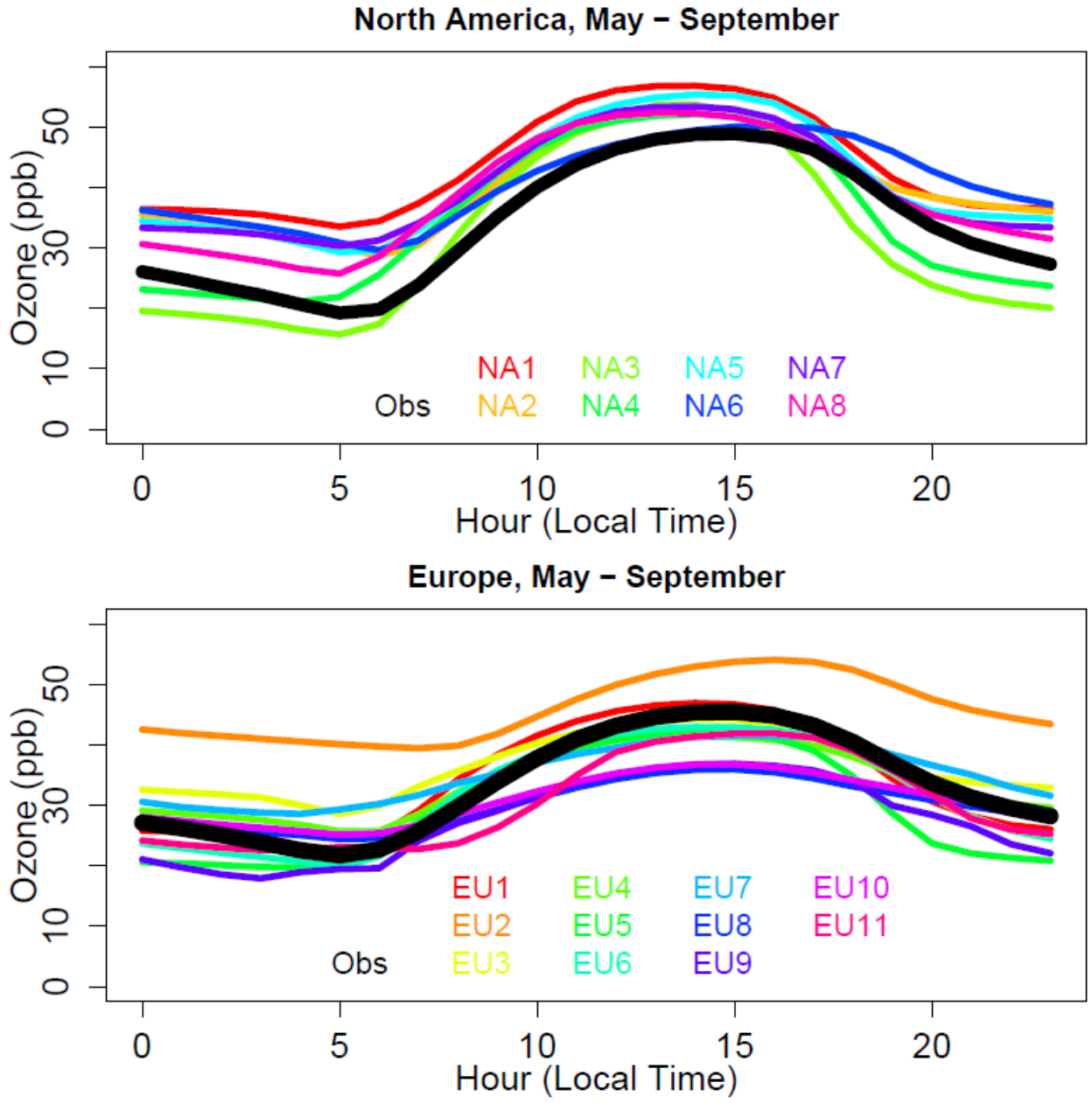
North America, May – September
Ozone (ppb)
0
10
30
50
0
5
10
15
20
Hour (Local Time)
Obs
NA1
NA2
NA3
NA4
NA5
NA6
NA7
NA8
Europe, May – September
Ozone (ppb)
0
10
30
50
0
5
10
15
20
Hour (Local Time)
Obs
EU1
EU2
EU3
EU4
EU5
EU6
EU7
EU8
EU9
EU10
EU11

Figure 3.

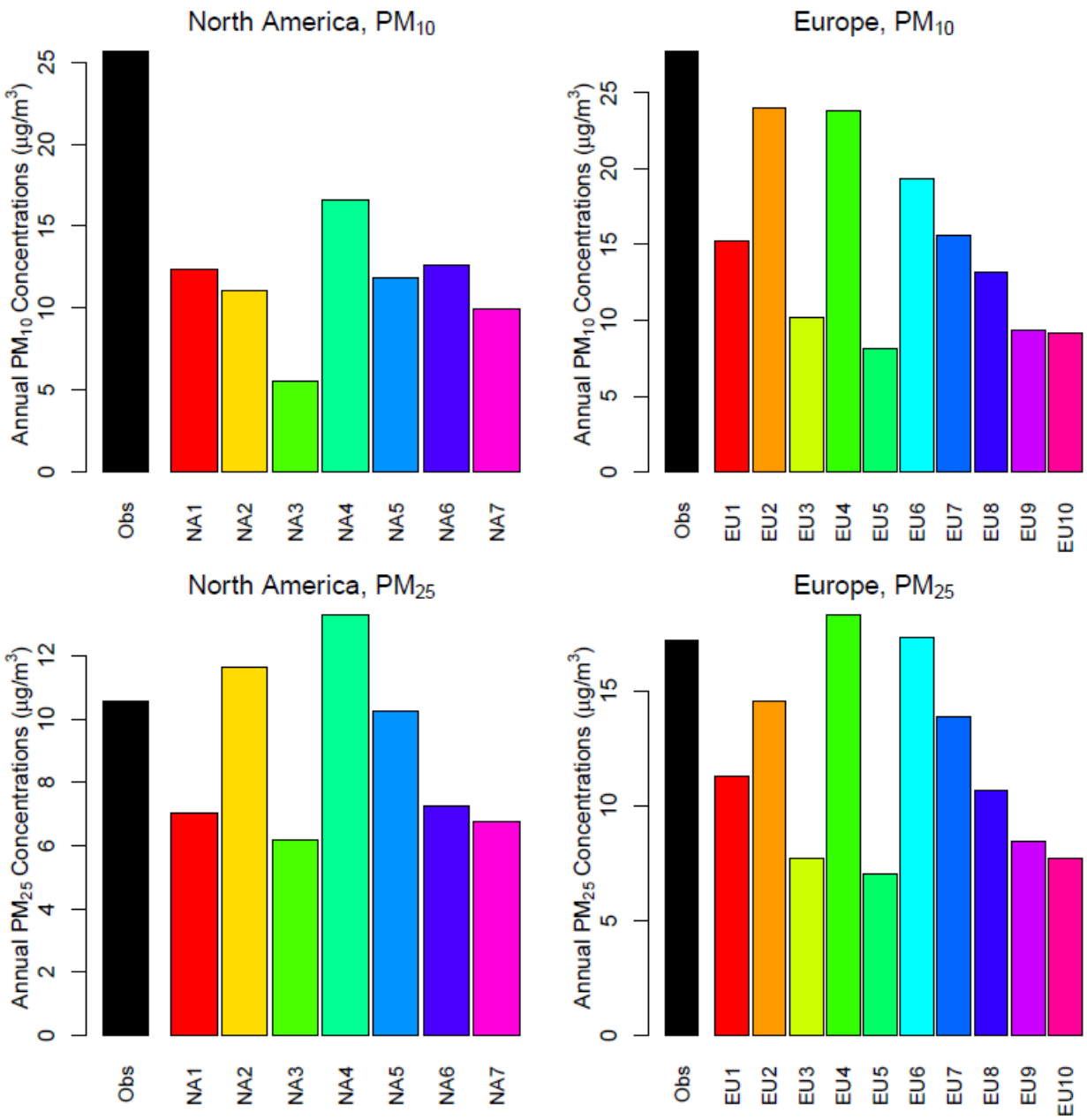
North America, PM10
Annual PM10 Concentrations (µg/m3)
0
5
10
15
20
25
Obs
NA1
NA2
NA3
NA4
NA5
NA6
NA7
Europe, PM10
Annual PM10 Concentrations (µg/m3)
0
5
10
15
20
25
Obs
EU1
EU2
EU3
EU4
EU5
EU6
EU7
EU8
EU9
EU10
North America, PM25
Annual PM25 Concentrations (µg/m3)
0
2
4
6
8
10
12
Obs
NA1
NA2
NA3
NA4
NA5
NA6
NA7
Europe, PM25
Annual PM25 Concentrations (µg/m3)
0
5
10
15
Obs
EU1
EU2
EU3
EU4
EU5
EU6
EU7
EU8
EU9
EU10

Figure 4.