1 2	
3	Assessment of the MACC reanalysis and its influence as chemical boundary
4	conditions for regional air quality modeling in AQMEII-2
5	
6	
7	Giordano, L. ^a , Brunner, D. ^{a,*} , Flemming, J. ^b , Hogrefe, C. ^c , Im, U. ^d , Bianconi, R. ^e , Badia, A. ^f ,
8	Balzarini, A. ^g , Baro, R. ^h , Chemel, C. ⁱ , Curci, G. ^j , Forkel, R. ^k , Jimenez-Guerrero, P. ^h , Hirtl,
9	M. ¹ , Hodzic, A. ^m , Honzak, L. ⁿ , Jorba, O. ^f , Knote, C. ^m , Kuenen, J.J.P. ^o , Makar, P.A. ^p ,
10	Manders-Groot, A. ^o , Neal, L_{i}^{q} , Perez, $L_{i}L_{i}^{r}$, Pirovano, G_{i}^{g} , Pouliot, G_{i}^{s} , San Jose, R. ^r , Savage
11	N $\stackrel{\text{Schroder}}{}$ W $\stackrel{\text{Schrider}}{}$ Solve $\stackrel{\text{Schroder}}{}$ Sol
17	Di Vahua K V Žahkan D W. Zhang V V and Calmarini S d
12	R., Fanya, K., Zaokar, R. ^{***} , Znang, F. [*] , and Gaimarini, S. [*]
13	
14	^a Laboratory for Air Pollution and Environmental Technology, EMPA, Duebendorf, Switzerland.
15	^b ECMWF, Shinfield Park, RG2 9AX Reading, United Kingdom.
16	^c Atmospheric Modelling and Analysis Division, Environmental Protection Agency, Research Triangle Park, USA.
17	^d Institute for Environment and Sustainability, Joint Research Centre, European Commission, Ispra, Italy.
18	^e Enviroware srl, Concorezzo (MB), Italy.
19	^f Earth Sciences Department, Barcelona Supercomputing Center (BSC-CNS), Barcelona, Spain.
20	⁸ RSE S.p.A., Milano, Italy
21 22	^h University of Murcia, Department of Physics, Physics of the Earth. Campus de Espinardo, Ed. CIOyN, 30100 Murcia, Spain
22 72	Murcu, Span.
23 24	9AB. UK.
25	^j Department of Physical and Chemical Sciences, Center of Excellence for the forecast of Severe Weather
26	(CETEMPS), University of L'Aquila, L'Aquila, Italy.
27	^k Karlsruher Institut fu [°] r Technologie (KIT), IMK-IFU, Kreuzeckbahnstr. 19, 82467 Garmisch-Partenkirchen,
28	Germany.
29	¹ Section Environmental Meteorology, Division Customer Service, ZAMG - Zentralanstalt fuer Meteorologie und
30	Geodynamik, 1190 Wien, Austria
31	^m National Center for Atmospheric Research, Boulder, CO, US.
32	ⁿ Center of excellence SPACE-SI, Ljubljana, Slovenia
33	^o Netherlands Organization for Applied Scientific Research (TNO), Utrecht, The Netherlands.
34	^p Air Quality Research Section, Atmospheric Science and Technology Directorate, Environment Canada, 4905
35	Dufferin Street, Toronto, Ontario, Canada.

- 36 ^{*q*} Met Office, FitzRoy Road, Exeter, EX1 3PB, United Kingdom.
- 37 r Environmental Software and Modelling Group, Computer Science School Technical University of Madrid,
- 38 Campus de Montegancedo Boadilla del Monte-28660, Madrid, Spain.
- 39 ^s Emissions and Model Evaluation Branch, Atmospheric Modeling and Analysis Division/NERL/ORD, Research
- 40 Triangle Park, North Carolina, USA
- 41 ^t Leibniz Institute for Tropospheric Research, Permoserstr. 15, D-04318 Leipzig, Germany
- 42 "National Institute of Meteorology and Hydrology, Bulgarian Academy of Sciences, Sofia, Bulgaria
- 43 ^v Department of Marine, Earth and Atmospheric Sciences, 2800 Faucette Drive, Jordan Hall, Campus Box 8208,
- 44 North Carolina State University
- 45 "University of Ljubljana, Faculty of Mathematics and Physics, Ljubljana, Slovenia

- 47 *Corresponding author email address: dominik.brunner@empa.ch (Brunner, D.)
- 48
- 49

50 Abstract

51

The Air Quality Model Evaluation International Initiative (AQMEII) has now reached its second 52 phase which is dedicated to the evaluation of online coupled chemistry-meteorology models. 53 Sixteen modelling groups from Europe and five from North America have run regional air quality 54 models to simulate the year 2010 over one European and one North American domain. The MACC 55 re-analysis has been used as chemical initial (IC) and boundary conditions (BC) by all participating 56 57 regional models in AQMEII-2. The aim of the present work is to evaluate the MACC re-analysis along with the participating regional models against a set of ground-based measurements (O₃, CO, 58 NO, NO₂, SO₂, SO₄²⁻) and vertical profiles (O₃ and CO). Results indicate different degrees of 59 agreement between the measurements and the MACC re-analysis, with an overall better 60 61 performance over the North American domain. The influence of BC on regional air quality simulations is analyzed in a qualitative way by contrasting model performance for the MACC re-62 63 analysis with that for the regional models. This approach complements more quantitative approaches documented in the literature that often have involved sensitivity simulations but 64 65 typically were limited to only one or only a few regional scale models. Results suggest an important influence of the BC on ozone for which the underestimation in winter in the MACC re-analysis is 66 mimicked by the regional models. For CO, it is found that background concentrations near the 67 domain boundaries are rather close to observations while those over the interior of the two 68 69 continents are underpredicted by both MACC and the regional models over Europe but only by 70 MACC over North America. This indicates that emission differences between the MACC reanalysis and the regional models can have a profound impact on model performance and points to 71 the need for harmonization of inputs in future linked global/regional modeling studies. 72

73

74 *Keywords:* Online-coupled meteorology-chemistry modelling; model evaluation; AQMEII-2

- 75
- 76

77 **1. Introduction**

78

Modelling plays an important role in the integrated assessment of air quality issues, contributing 79 80 to strengthening the understanding and characterization of air pollution and eventually leading to 81 well-informed air quality management decisions and strategies. Regional air quality modelling has been the focus of considerable development during recent decades, driven by increased concern 82 regarding the impact of air pollution on human health and the ecosystem. Numerous air quality 83 models have been developed by research groups worldwide and are being widely used for 84 85 designing emission control policies and forecasting air quality. However, unlike other geophysical sciences such as climatology, there have only been limited coordinated international efforts to 86 study and evaluate the performance of the air quality models. 87

Since 2008, the Air Quality Model Evaluation International Initiative (AQMEII, Rao et al., 88 2010) coordinated by the European Joint Research Center (JRC), the U.S. Environmental 89 Protection Agency (EPA), and Environment Canada (EC), has promoted research on regional air 90 quality model evaluation across the atmospheric modelling communities of Europe and North 91 America. AQMEII has now reached its second phase which is dedicated to the evaluation of online 92 coupled chemistry-meteorology models, as opposed to Phase 1 where only offline models were 93 considered. At the European level, AQMEII collaborates with the COST Action European 94 95 framework for online integrated air quality and meteorology modelling (EuMetChem, http://eumetchem.info). AQMEII-2 has the goal of validating the many different models used to 96 97 estimate air quality at local levels around the world, how these models might be used in ensemble simulations and whether these models can be used to simulate feedbacks between weather and 98 99 chemistry and to predict ways in which climate change will interact with air quality.

Two spatial domains were used in the exercise - one over Europe and one over North America. 100 All groups participating in AQMEII-2 performed simulations on one or both of these domains 101 using the same input data. Simulation outputs were regridded onto the same horizontal and vertical 102 grid and, for comparison with observations, were interpolated to prescribed sets of measurement 103 station locations (receptors). All outputs were collected, together with evaluation data sets 104 (including measurements from surface in-situ networks from AirBase and EMEP, vertical profiles 105 106 from ozonesondes and aircraft from MOZAIC, and ground-based remote sensing from AERONET), by the JRC. 107

Regional models need to constrain the concentrations at the domain boundaries: the initial and lateral boundary conditions (IC, BC hereafter) were shared between all groups and were provided by the MACC re-analysis of the IFS-MOZART model ("MACC re-analysis" hereafter) (Stein et al., 2011; Inness et al., 2013). Emissions were provided by the TNO/MACC database for anthropogenic emissions for the European domain and by U.S. EPA for the North American domain.

The focus of the present paper is on performing an operational evaluation (Dennis et al., 2010) 114 115 of the MACC re-analysis data in the same way as all the regional models participating in AQMEII-2 have been evaluated (Im et al., 2014a,b; Brunner et al., 2014) and to assess its influence as 116 chemical BC. The current study complements the work by Im et al. (2014 a,b) by expanding the 117 list of variables that are analyzed and by systematically contrasting seasonal and spatial patterns 118 119 of model performance for the MACC re-analysis with the performance of the regional models. It also complements the work of Inness et al. (2013) by evaluating the MACC re-analysis fields for 120 121 additional pollutants (NO, SO₂, sulfate) and at additional monitoring sites. It should be noted that for a full quantification of the influence of BC on the results of regional models, sensitivity 122 123 simulations with varying BC would need to be performed by the regional models. Such sensitivity simulations were beyond the scope of the second Phase of AQMEII. A limited quantitative analysis 124 125 was performed by Hogrefe et al. (2014) for ozone over the North American domain for the months January and July. Replacing the MACC re-analysis by the GEMS re-analysis used during AQMEII 126 127 phase 1 had a significant impact on near-surface ozone concentrations with differences of the order of 7 ppb over large portions of the domain in January and 3 ppb in July. 128

129 Here, we analyze the influence of BC in a more qualitative way by evaluating the performance of the MACC re-analysis and by analyzing the differences between its results and those of the 130 131 participating models. This phenomenological analysis of one global model in comparison to a large number of regional models driven by boundary conditions from this global model complements 132 more quantitative approaches documented in the literature that often have involved sensitivity 133 simulations but typically were limited to only one or only a few regional scale models (e.g. Makar 134 et al, 2010; Katragkou et al, 2010; Hogrefe et al., 2011; Schere et al., 2012; and Akriditis et al., 135 136 2013). In addition, none of these earlier studies investigated the impact of boundary conditions on regional air quality model performance over two continents in a systematic manner. 137

138 In air quality numerical simulations, the influence of the BC on the results in the interior of the

model domain varies strongly from species to species depending on its lifetime. For species with 139 a lifetime comparable or exceeding the average time it takes to transport an air mass across the 140 model domain, the BC will potentially have a large influence on the modeling results. Therefore, 141 we evaluate the MACC re-analysis against observations for a range of species provided as BC for 142 the regional models with a large range of different lifetimes, and compare the results of the MACC 143 re-analysis with the results of the regional models. For longer lived species we expect the regional 144 models to follow more closely the MACC re-analysis, whereas for shorter lived species the 145 differences may be larger and dominated by differences in emissions within the model domain or 146 differences in photochemistry, deposition, and other factors. In cases where concentrations of 147 longer-lived species with large primary sources such as CO differ between the MACC re-analysis 148 and the regional-scale models, these differences may point to inconsistencies in the emission 149 150 inventories used in the global and regional scale simulations. Documenting such instances in this study will provide motivation for future work aimed at better linking global and regional scale 151 152 modeling systems, including a harmonization of emission inventories.

Since this study compares domain-averaged results from the regional models with both the MACC re-analysis and with observations, it also presents a limited evaluation of the regional models for a range of species for which the MACC reanalysis provided the boundary conditions (O_3 , CO, NO_x , SO_2 , SO_4^{2-}). While O_3 and $PM_{2.5}$ concentrations from the regional models have already been analyzed by Im et al. (2014a,b), the current study complements this work by including additional species and explicitly contrasting model performance between the global and regional models.

160

161 **2. Models**

162

163 2.1. Regional Models Participating In AQMEII Phase 2

164

In the context of AQMEII-2, 16 models from Europe (EU) and 5 models from North America (NA) have been used to simulate the year 2010 (see Table 1). Only one model (BG2) has been run offline, while the other models are online coupled. The online coupled models are separated into online access models (NL2, UK5, DE3 and US6) and online integrated models (all the remaining). Online access models are defined as models that use independent meteorology and chemistry modules that might even have different grids, but exchange meteorology and chemistry data on a regular and frequent basis. In contrast, online integrated models simulate meteorology and chemistry over the same grid in one model using one main time step for integration as defined in (Baklanov et al., 2014).

Nine groups used the WRF-Chem model (Grell et al., 2005) and its variant (e.g., Wang et al., 2014), having different gas-phase mechanisms but similar aerosol modules that employ different size distributions approaches (modal/bin and inorganic/organic aerosol treatments). The IT2 simulation is performed with an experimental version of WRF-Chem v. 3.4, where the new secondary organic aerosol scheme VBS is coupled to the aerosol indirect effects modules.

Anthropogenic emissions for AQMEII-2 were provided by U.S. EPA for North America 179 (Pouliot et al., 2014), and TNO (Nederlandse Organisatie voor Toegepast Natuurwetenschappelijk 180 Onderzoek) for Europe (Kuenen et al., 2014, Pouliot et al., 2014). Each participating group had 181 the freedom to choose a grid coordinate system that should however cover the prescribed domains 182 183 (EU or NA). Emissions were therefore re-gridded for each model. Pouliot et al. (2014) provides quantitative explanations of the aerosol loading in the coupled model runs for 2010 and a 184 185 quantitative analysis of changes in emissions between 2006 and 2010. Biomass burning emissions for the European domain were provided by FMI (http://is4fires.fmi.fi/) and by SMART- FIREv2 186 for NA. Other natural emissions such as sea salt, mineral dust or biogenic VOCs were not 187 prescribed but simulated online by the individual models. 188

The simulations were conducted for continental-scale domains of EU and NA covering the continental U.S., southern Canada and northern Mexico. To facilitate the cross-comparison between models, the participating groups interpolated their model output to a common regular grid with 0.25° resolution for both continents. The native grids used by the different groups varied in both resolution and grid projection but typically had a resolution on the order of 0.25° (Im et al. 2014a; Brunner et al., 2014). In addition, model values at observation stations (receptors) were obtained by interpolation from the original model output files for comparison to observations.

196

197 2.2. Global IFS-MOZART Model Providing Chemical Boundary Conditions

The chemical BC for all models were provided by ECMWF from the MACC re-analysis (Inness 199 et al., 2013). The MACC re-analysis uses an updated data set of anthropogenic emissions 200 201 (MACCity, (Granier et al., 2011)) with assimilation of satellite observations of O₃, CO and NO₂ 202 in the coupled system IFS-MOZART (Flemming et al., 2009). It produced a 10 year long reanalysis of global atmospheric composition for the period 2003-2012. As pointed out in In- ness et al. 203 (2013), the assimilation of satellite observations of O₃, CO, and NO₂ greatly improved total column 204 values, that are generally in very good agreement with independent observations, but profiles can 205 206 show some problems in the boundary layer where concentrations are dominated by emissions. Moreover, most of the assimilated satellite observations had little sensitivity to pollutants near the 207 surface and very coarse (or no) vertical resolution in the troposphere and therefore provided fewer 208 constraints on concentrations in the planetary boundary layer. 209

210 MACC data are available in 3-hour time intervals and were provided in daily files with 8 times per file. The horizontal resolution of the model is $1.125^{\circ} \times 1.125^{\circ}$. Variables were provided as 3D 211 212 fields in pressure hybrid vertical coordinates and included gas phase species (O₃, NO, NO₂, HNO₃, HO₂, NO₂, OH, H₂O₂, CO, CH₄, PAN, SO₂, CH₂O (formaldehyde), C₂H₆ (ethane), CH₃CHO 213 214 (acetaldehyde), BIGENE (C>3 alkenes and alkynes), BIGALK (C>3 alkanes), ISOP (isoprene), TOLUENE) and aerosol species (sea-salt, dust, sulfate, organic matter and black carbon. Note: 215 216 Organic matter and black carbon were described as sum of hydrophobic and hydrophilic). 217 NMVOC species had to be assigned to the most closely matching chemical species depending on 218 the individual model's chemical speciation.

In order to mitigate known biases and issues in the MACC data, a list of recommendations were 219 formulated for the modelers to follow. The organic aerosol concentrations were assigned to 220 primary organic aerosol (POA) since it is unclear how this should be distributed on secondary 221 organic aerosol (SOA) in a given model. Since a preliminary analysis indicated that MACC sea-222 salt fields were significantly biased high, they were not used as input to the regional models, but 223 the simulation grids were large enough such that each model could generate the sea salt fields 224 225 internally using its own sea salt parameterization. Mineral dust aerosols were provided by MACC in three different size ranges $(0.03 - 0.55 \,\mu\text{m}, 0.55 - 0.9 \,\mu\text{m} \text{ and } 0.9 - 20 \,\mu\text{m})$ which had to be 226 227 mapped onto the aerosol size classes used by each of the regional models. The guideline was to use a simple but mass-conserving mapping while taking into account the fact that the IFS-228 229 MOZART model used in MACC placed a too large fraction of total dust mass into the smallest

size bin, a deficiency that was improved in 2012 for the near-real-time analysis product (Jean-Jacques Morcrette, ECMWF, personal communication). It was therefore advised to shift the MACC total dust aerosol mass from the size bins listed above to larger size ranges in the regional models whenever possible, e.g. by summing up the masses of the three size ranges and then attributing 10% to fine and 90% to coarse mineral dust, following Johnson and Osborne (2011).

235

3. Observations and Evaluation Method

237

238 *3.1. Surface Observations*

239

The 2010 hourly measurements of surface concentrations of O₃, CO, SO₂, NO, NO₂ and SO₄²⁻ 240 241 aerosol were provided by EMEP (European Monitoring and Evaluation Programme, http://www.emep.int/) and AirBase (http://acm.eionet.europa.eu/databases/airbase/) in Europe. In 242 243 North America they were provided by AIRS (Aerometric Information Retrieval Systems, http://www.epa.gov/air/data/aqsdb.html/) and NAPS (National Air Pollution Surveillance, 244 http://www.ec.gc.ca/rnspa-naps/). The data were homogenized and ingested into the ENSEMBLE 245 system (Galmarini et al., 2012) by the JRC for the EU case, and by Environment Canada for the 246 247 NA case. Weekly/bi-weekly CO measurements from the NOAA/GMD flask sampling network (Novelli and Masarie, 2009) were obtained from WDCGG (World Data Centre for Greenhouse 248 Gases, http://ds.data.jma.go.jp/gmd/wdcgg/) and were not used in any other AQMEII Phase 2 249 studies such as Im et al. (2014 a,b). 250

251

252 *3.2. Vertical Profiles*

The provided 253 2010 vertical profiles at selected airports were by MOZAIC 254 (http://www.iagos.fr/web/rubrique2.html). The MOZAIC program is designed to collect O₃, CO and water vapor data using automatic equipment installed on-board several long-range passenger 255 airliners flying regularly from Europe to destinations all over the world. Five long-range passenger 256 aircrafts carry MOZAIC instruments and visit 35 different airports around the world. 257 Measurements of O3 are taken every four seconds from takeoff to landing and have an accuracy of 258 ±(2 ppbv+2%) (Thouret et al., 1998). Measurements of CO are taken every 30 seconds from takeoff 259

260 to landing and have an accuracy of ±(5 ppbv±5%) (Nédélec et al., 2003). The original MOZAIC data set is separated into cruising (for which data are temporally averaged) and landing/take off 261 262 phase (for which data are averaged over 100 m vertical intervals). The MOZAIC data considered here were gathered during takeoff and landing phases at the airport of Frankfurt, with the majority 263 of data being in the morning hours between 07 and 12 UTC. Unfortunately, data coverage in 2010 264 was much poorer than in other years and was limited to the winter and fall seasons (Solazzo et al., 265 2013). For this analysis we use measurements at Frankfurt airport location, ingested into the 266 ENSEMBLE system by the JRC to provide measurements at a set of 13 fixed elevations above 267 ground. 268

269

270 *3.3. Evaluation method*

Each modelling group has provided standardized outputs: hourly maps of surface concentrations, re-gridded to the same horizontal resolution of $0.25^{\circ} \times 0.25^{\circ}$, hourly surface concentrations at selected locations (receptor points) and vertical profiles at airport locations. The same standardized output has been extracted from MACC, i.e. MACC fields were interpolated to the common $0.25^{\circ} \times 0.25^{\circ}$ analysis grid for the analysis of gridded fields and to the location of the selected monitoring sites for the analysis at specific receptor locations.

For the present analysis, we retrieved CO, O₃, NO, NO₂, SO₂ and SO₄²⁻ aerosol receptor data 277 from the ENSEMBLE system to compare against station observations. This selection is motivated 278 279 by the requirement that the species should be provided as BC by the MACC re-analysis, that there should be a sufficient number of observations available for validation, and that the compounds 280 should cover a range of lifetimes. Since the regional models use relatively coarse horizontal 281 resolutions (see also Table 1 in Im et al. (2014a)), we compare surface concentrations only at rural 282 283 sites, selecting stations with data availability greater than 75% and altitudes lower than 1000 m 284 ASL. The number of stations with available measurements according to this selection is shown in Table 2. We mainly focus on midday/afternoon values (12-14 UTC over EU, 19-21 UTC over NA) 285 because models are known to have difficulties in representing the nighttime boundary layer 286 287 accurately (Steeneveld et al., 2006; Brunner et al., this issue). The model outputs were first time-288 averaged for each month and then averaged over all selected stations.

For surface concentration maps, we have computed seasonal multi-model means. EU data were

converted to ppb using the provided standard pressure and temperature to be directly comparablewith NA data.

Model performance is quantified through mean, normalized mean bias (NMB), root mean square error (RMSE) and Pearson's correlation coefficient.

294

295 **4. Results**

296

297 4.1. Evaluations of the Models against Surface Observations

298

299 *4.1.1.* O₃

300

 O_3 is one of the most important photooxidants in the atmosphere. High surface O_3 concentrations are of concern as they can cause serious problems to human health and vegetation. It is not emitted directly into the air, but in the troposphere it is formed via photochemical cycles involving volatile organic compounds (VOC) and oxides of nitrogen (NO_x). These precursors of ozone have both natural and anthropogenic sources both of which are accounted for in the models analyzed in this study.

O₃ results were extensively discussed in Im et al. (2014a), where it was found that the MACC 307 re-analysis has a large diurnal amplitude, with too high values during day and too low values 308 during night. We extended their analysis by looking separately at the median monthly midday 309 values (Figure 1) and midnight values (Figure 2). In Table 3 we report performance metrics for 310 all the models, focusing on winter and summer midday periods. The MACC re-analysis midday 311 median values over EU show a strong seasonal cycle, with an underestimation in the winter season 312 (NMB=-26.04%) and an overestimation in summer (NMB=26.4%). On the other hand, MACC re-313 analysis midnight values consistently underestimate O₃ throughout the year. This overestimation 314 315 of the amplitude of the diurnal cycle of O_3 was also reported by Inness et al. (2013) but the reasons could not be identified. This problem is reduced in the new fully integrated model C-IFS 316 (Flemming et al., 2014). Neither MACC nor the regional models capture the elevated observed 317 concentrations during springtime, a shortcoming that has also been pointed out by Inness et al. 318 319 (2013) for the MACC fields. The underestimation of ozone in the regional models during these months is thus likely driven by too little ozone entering the domain through the lateral boundaries.

321 Over NA, the MACC re-analysis median values are closer to the observations, both during daytime

and nighttime, only overestimating midday O_3 in the month of July.

323

324 *4.1.2. CO*

325

326 CO affects the concentrations of O_3 and the OH radical, and therefore plays a fundamental role in global tropospheric chemistry. Anthropogenic sources account for roughly two thirds of all CO 327 emissions (Seinfeld and Pandis, 2006) through combustion of biomass and fossil fuel; in fact, the 328 329 highest concentrations of CO are found in highly industrialized regions. Wildfires are another 330 important source, originating mostly from tropical forest fires but receiving contributions as high as 25% from boreal forest fires in some years (Goode et al., 2000). Photochemical oxidation of 331 332 methane and non-methane hydrocarbons is another important source which is estimated to represent more than half of the total source of CO globally, larger than direct emission (Duncan et 333 al., 2007). 334

335 CO is the longest lived species that has been simulated by all models, including MACC re-336 analysis, and for which we have available observations.

337 The MACC re-analysis for reactive gases has been extensively evaluated in Inness et al. (2013) and Stein et al. (2014). In particular, surface CO was found to agree well with NOAA/GMD 338 observations at a selection of remote sites (Mace Head, Key Biscayne, Tenerife and South Pole); 339 total column values were in good agreement with observations, but vertical profiles showed some 340 differences from observations in the boundary layer (Inness et al., 2013). Since boundary layer 341 concentrations are dominated by emissions, these differences could be an indication of errors in 342 the emission inventory for CO (and VOCs,) underestimation of the chemical source, a lack of 343 efficacy in modelling boundary layer mixing processes, and/or a mismatch in the spatial 344 representativeness of the observations and MACC fields, but no further diagnostic analysis of the 345 relative importance of these potential factors was presented in Inness et al. (2013). 346

In order to conduct a more comprehensive comparison within the AQMEII-2 modelling domains, we have selected a sample of four NOAA/GMD flask sampling sites (Novelli and Masarie, 2013), two for each domain, of which one is representative of remote conditions, as close as possible to the domain border (Mace Head for EU and Key Biscayne for NA), and one station 351 is located in the middle of the domain (Black Sea EU and Park Falls for NA). Figure 3 shows CO concentrations from the MACC model compared with measurements at these selected stations. At 352 353 the remote stations near the domain boundaries (Mace Head and Key Biscayne) the model agrees 354 well with the observations, while there are differences between observations and MACC fields at the Park Falls and Black Sea sites. In particular, the MACC re-analysis exhibits a moderate level 355 of underprediction at the Park Falls site for most of spring and summer, while there is a large 356 negative bias, particularly in winter, at the Black sea site similar to the biases observed at other 357 sites over continental Europe (see below). As explained in Inness et al. (2013) and Stein et al. 358 359 (2014), the emission inventory (MACCity) used by the re-analysis might be the primary source of such large negative biases. 360

In Figure 4 we show the midday median monthly CO surface concentrations for 48 European 361 362 and 17 North American rural stations with continuous measurements. In Table 4 we report performance metrics for all the models, focusing on winter and summer midday periods. In the EU 363 364 case, the MACC re-analysis consistently underestimates CO surface concentrations, particularly in winter, and this behavior is mimicked by all the models. The fact that the MACC re-analysis 365 366 shows only small biases at the background station Mace Head near the western domain boundary suggests that this underestimation within the domain is not caused by underestimated background 367 368 concentrations. Instead, it is likely that CO emissions within the model domain are significantly underestimated both in the MACCity and the TNO/MACC emission inventories especially during 369 370 winter (see Stein et al. (2014) for a more thorough discussion). Potential discrepancies in the spatial scales represented by the observations and model predictions are likely not a major factor because 371 372 the analysis focuses on rural sites. In the NA case the underestimation is less severe for MACC . The regional models do not track CO from MACC as closely as in the EU case, suggesting larger 373 374 differences between the MACCity and the NA emission inventories (see also Section 4.3 below). 375 In fact, the regional models even tend to overestimate CO suggesting an overcompensation of the too low BC by too high CO emissions within the domain. The study of Miller et al. (2008) using 376 tall-tower and aircraft measurements in a model-data assimilation framework indeed indicated that 377 CO emissions over North America as reported by the U.S. EPA might be too high. 378

379

380 *4.1.3. NO and* NO₂

Nitrogen oxides (NO and NO₂, or NO_x) are major precursors for O₃ and nitrate aerosols. NO_x
emissions and their oxidation products strongly influence the concentrations of air pollutants (O₃,
PM) (Seinfeld and Pandis, 2006).

For the year 2010, 211 rural stations reported observations for NO and 285 for NO₂ in Europe, while only 15 rural stations for NO and 59 for NO₂ were available in North America. While measurements at rural sites are expected to be better comparable to the coarse resolution model outputs, it should be noted that standard NO₂ measurements based on molybdenum converters can be significantly positively biased at rural sites due to interferences from other oxidized nitrogen species including PAN and HNO₃ (see for example Steinbacher et al. (2007) and references therein).

Figure 5 and Figure 6 show the midday median monthly surface concentrations of NO and NO₂ respectively. In Table 5 we report performance metrics for all the models, focusing on winter and summer midday periods, for combined NO_x.

In both the EU and NA case, the MACC model consistently underestimates the surface NO and 395 NO₂ concentrations throughout the year 2010. The regional models are closer to the observations 396 than the MACC re-analysis, especially for NO₂. There is significant spread between the models, 397 especially during winter, indicating substantial inter-model differences in aspects such as vertical 398 mixing, deposition, and chemical lifetimes which warrant further investigation in future diagnostic 399 model evaluation studies. Significantly higher NO_x concentrations are simulated by model ES2a/b 400 as compared to other models in winter, which may be due to the missing heterogeneous reaction 401 of N2O5 to HNO3 in this model which is an important sink of NOx at nighttime and in winter (Badia 402 and Jorba, 2014). 403

404 Anthropogenic, biomass burning, and soil decay NO_x emissions are primarily released within the atmospheric planetary boundary layer (PBL) where the lifetime of NO_x is a few hours in 405 406 summer and up to one day in winter (Schaub et al., 2007). Given such a short lifetime, the ability of the models to reproduce NO and NO₂ surface values is not directly affected by the MACC 407 408 boundary conditions, but the consistent underestimation of NO2 by most models probably indicates a lack of emissions in the inventories used by both the regional models and the MACC re-analysis. 409 410 On the other hand, since we are looking at rural, relatively emissions-poor locations, the 411 underestimation can be due to other factors as well including an underestimation of the chemical lifetime of NO_x, too high dry deposition, an underestimation of natural emissions from soils as 412

413 suggested by Jaeglé et al. (2005), or the missing source from lightning in the regional models 414 which, however, would mainly affect the results in summer. Knote et al. (2014) investigated 415 differences between the chemical mechanisms used in the AQMEII-2 models and found 416 differences in simulated radical concentrations of up to 40% for OH and >100% for NO₃ which 417 would indicate a large spread in chemical lifetimes between the models but based on their study it 418 is not possible to conclude on a general positive or negative bias.

Another likely contribution to the model underestimation is a systematic positive bias of the NO₂ measurements. Standard air pollution monitors use molybdenum converters which are known to be cross-sensitive to other reactive nitrogen species including PAN and HNO₃. Steinbacher et al. (2007) found that NO₂ measurements at a rural site in Switzerland from a standard instrument had a rather constant bias of 1.5 to 2 ppb when compared to an accurate, NO₂-specific instrument throughout the year. They concluded that only 70-83% of the actually measured NO₂ signal of the molybdenum-based NO_x monitor was attributable to real NO₂.

426

427 *4.1.4.* SO₂ and SO₄²⁻

428

SO₂ is a major atmospheric pollutant of mainly anthropogenic origin, produced by the combustion of fossil fuel and by industrial facilities. With an average lifetime in the troposphere of few days (Lee et al., 2011), SO₂ is oxidized to form sulfuric acid and sulfate aerosols. SO₂ is the principal precursor of sulfate aerosols and, because of its links to climate, air quality and human health issues, is extensively and continuously monitored on a global scale.

Figure 7 shows the midday median monthly SO₂ surface concentrations averaged over 165 (EU) 434 435 and 44 (NA) rural stations. In table 6 we report performance metrics for all the models. In the EU case, the MACC re-analysis overestimates the surface SO₂ concentrations during the winter season 436 437 (up to 2 ppby with respect to the median of the observations), while it reproduces the measurements 438 well during the rest of the year. Flemming et al. (2014) suggested that the overestimation in winter and the correspondingly too large amplitude of the seasonal cycle could be introduced by the 439 diffusion scheme in IFS-MOZART. The seasonal bias is much reduced in the new model version 440 C-IFS which, however, not only differs from IFS-MOZART with respect to vertical mixing, but 441 442 also with respect to sulfur chemistry and wet and dry deposition. In the NA case, the MACC reanalysis is within the interquartile range of the observations for most of the year, but shows a 443

pronounced seasonal cycle not seen in the observations that manifests itself in a tendency to 444 underestimate surface SO₂ in summer. Due to its rather short lifetime, SO₂ does not get transported 445 efficiently from the borders towards the center of the domain, where the performance of the models 446 447 is not strongly affected by the MACC re-analysis biases. In winter, the regional models show significantly lower SO₂ concentrations than the MACC re-analysis, more consistent with the 448 observations. However, over NA the regional models also exhibit a seasonal cycle that is not 449 visible in the observations. Some models show a significant underprediction, notably UK5 over 450 EU. Over NA, CA2 and CAf2 show much higher SO_2 than all other models and significantly 451 overpredict SO₂ in winter. The reasons for these model-to-model differences as well as the 452 differences between the observed and simulated seasonal cycles over NA should be investigated 453 further in future studies. 454

455 Sulfate is produced by oxidation from SO₂ in both the gas phase and the liquid phase. Oxidation in the gas phase by OH is more important during the summer compared to other seasons. Note 456 457 that in the MACC model there are two representations of SO₂ and sulfate, one from the mainly gas-phase chemistry of MOZART, and one from the MACC aerosol module (Morcrette et al., 458 459 2009) based on a simple SO₂ to sulfate conversion approach. The fields provided as BC for the regional models were SO₂ from the MOZART scheme but sulfate from the MACC aerosol module. 460 461 Different from the regional models, the SO₂ and sulfate fields analyzed for the MACC re-analysis are thus not chemically coupled. The sulfate component of particulate matter is overestimated 462 463 during summer by the MACC re-analysis both in EU and NA, as shown in Figure 8, where we present daily averaged median monthly sulfate aerosol surface concentrations, averaged over 33 464 465 (EU) and 70 (NA) rural stations.

Performance metrics for all the models are reported in table 7. For the EU case, the 466 measurements are reported as PM10 (SO4²⁻) but in NA as PM2.5 (SO4²⁻). One possible reason for 467 the positive bias in the MACC re-analysis in the summer months is that the MACC aerosol model 468 469 does not contain a representation of ammonium nitrate aerosol which represents a large component of the European aerosol loading. Therefore the assimilation of satellite AOD will tend to increase 470 471 the other aerosol components to give the correct AOD overall. In the EU case (top panel), model 472 CH1 follows closely the MACC re-analysis predictions and shows a similar strong positive bias pattern. This overestimation contrasts with the findings of a previous evaluation of COSMO-ART 473 474 (Knote et al., 2011) which suggested under- rather than over-prediction of sulfate. This is probably

475 due to the implementation of a comprehensive wet chemistry scheme (Knote and Brunner, 2013) that had not been included in the previous evaluation. The other regional models do not show the 476 477 same positive bias as MACC and COSMO-ART, indicating that the sulfate simulations are not very sensitive to the choice of chemical boundary conditions, but further analysis is needed to 478 investigate the relative role of long-range transport vs. within-domain formation in simulating 479 regional-scale sulfate. The underestimation of sulfate in winter by a number of models is likely 480 related to an underestimation of SO₂ to sulfate oxidation in clouds as also noted by Im et al. (2014) 481 482 or missing heterogeneous oxidation of SO₂ as reported in Wang et al. (2014). Balzarini et al. (2014) investigated the in-cloud oxidation pathway in more detail for two different versions of WRF-483 Chem with (SI2) and without (IT1) SO₂ oxidation in cloud water. The model IT1 indeed is one of 484 the models that strongly underestimate sulfate in winter while SI2 is rather closely following the 485 486 observations. Although the other WRF-Chem models did include aqueous phase chemistry, they follow more closely the seasonal profile of IT1 than the profile of SI2. However, these models 487 488 used different modules for both aqueous chemistry and wet deposition than SI2. A further analysis 489 of the impact of different configurations available in WRF-Chem on SO₂ and sulfate would thus 490 be highly valuable.

Using a diagnostic analysis approach, Wang et al. (2014) investigated several possible factors contributing to the underpredictions of sulfate in the US. They found that model performance for sulfate can be improved by the use of a more detailed convective cloud chemistry, addition of SO₂ heterogeneous chemistry, and the use of a more realistic surface wind drag parameterization, among which the addition of SO₂ heterogeneous chemistry contributes the most to the improvement.

The models UK5 and NL2 also represent sulfate concentrations fairly accurately and at the same time are among the models with rather low SO₂ concentrations in winter (especially UK5), suggesting a more efficient SO₂ to sulfate conversion as compared to other models.

In the NA case (bottom panel of Figure 8), CA2 and CA2f show a similar pattern of strong positive bias, similarly to SO₂, but with the largest overestimation in summer, as opposed to winter for SO₂. The other models are able to well reproduce sulfate surface concentrations, but this pattern is not correlated with the positive bias shown by the MACC re-analysis.

504

505 4.2. Vertical Profiles at Frankfurt Airport

Most air pollutants have a longer lifetime at higher altitudes and show a much smoother distribution than in the boundary layer. This is a consequence of slower photochemistry due to generally lower concentrations and lower temperatures and the absence of emission sources and (dry) deposition sinks. Therefore, the influence of the MACC re-analysis is expected to be larger in the free and upper troposphere than in the boundary layer.

In Figure 9, we show the winter averaged vertical profiles at Frankfurt airport for O₃ (top panel)
and CO (bottom panel), in which all regional models and the MACC re-analysis are compared to
MOZAIC aircraft measurements.

Despite the large spread at the surface, O_3 concentrations in the models are similar in the lower 514 troposphere and start to diverge again above 6 km. While most models used a vertical resolution 515 of 30 and 40 levels in their simulations (Brunner et al., 2014), the spacing of these levels in the 516 517 tropopause region varies between models and may affect the downward mixing of stratospheric ozone into the upper troposphere. However, the two simulations performed with the NMMB-BSC 518 519 model at two largely differing vertical resolutions (ES2a: 24 levels, ES2b: 48 levels), differ only 520 little in the upper troposphere, indicating that factors other than vertical resolution must dominate 521 these results such as differences in the treatment of ozone in the stratosphere or different vertical 522 model tops. Note that due to the lack of observations in spring and summer, the profiles are only representative of the situation in fall and winter. In these seasons, O₃ shows a pronounced vertical 523 profile with low values near the surface possibly influenced by the titration of O₃ by NO. Above 524 the PBL the regional models quite closely track the MACC re-analysis, indicating a larger 525 526 influence of BC at higher altitudes, though as noted above the spread between the regional models 527 begins to increase in the upper troposphere.

Over Frankfurt, CO concentrations are severely underestimated up to ~6 km, with a bias close 528 to the surface exceeding 200 ppb. All the models, except for DE3 and UK5, follow closely the 529 vertical profile of MACC re-analysis, with an increasing spread towards the surface. This result is 530 consistent with the analysis of CO concentrations at ground stations and suggests that the emission 531 532 inventories used by the MACC re-analysis and the regional models might be the cause of large 533 negative biases closer to the surface, in particular at urban and polluted sites. It further suggests 534 that CO concentrations in the free troposphere are quite closely tied to the concentrations of the 535 model providing the boundary conditions. The deviations of the models DE3, UK5 and BG2 from the general behavior may point at problems with the technical implementation of the chemical 536

537 boundary conditions in these models.

538

539 4.3. Spatial analysis of CO and O₃

540

541 Due to simulation strategy and resources available, all the participating regional modeling 542 groups only performed one simulation for the year 2010 using the same BC. The lack of sensitivity 543 simulations with alternate sets of BC precludes us from quantitatively determining the effect of 544 boundary conditions on the simulation results. Nevertheless, it is possible to qualitatively 545 demonstrate the influence of the use of the MACC re-analysis as BC, in particular by analyzing 546 CO concentrations, the longest-lived species simulated and delivered as a gridded field by all 547 models.

In Figures 10 and 11 we show the winter and summer average surface concentrations over EU 548 for O_3 and CO, respectively. The superposed green dots on the maps represent the ground stations 549 550 selected for the analysis in Section 4.1.1 and 4.1.2. These maps show clearly the impact of the 551 difference in emissions between the models and the MACC re-analysis: While the CO 552 concentrations in the regional models are similar to MACC near the domain borders, they are significantly higher over the European continent in winter. Emissions outside of Europe were 553 554 missing in the TNO/MACC inventory resulting in a strong underestimation of CO concentrations over North Africa, especially over Cairo. Although the MACC re-analysis and the regional models 555 used different forest fire emission inventories (MACC: GFAS (Kaiser et al., 2012), regional 556 models: IS4FIRES (Sofiev et al., 2009)), the CO concentrations over the Russian forest fire regions 557 in summer are quite similar though somewhat higher in MACC. O₃ concentrations in the regional 558 models are also similar to MACC near the domain borders, but over continental Europe the 559 560 differences become larger. This is especially true for winter where MACC shows much lower O₃ concentrations. As suggested by the analysis in section 4.1.1., this is mostly due to a strong 561 underprediction of ozone at nighttime in the MACC re-analysis. This underestimation could 562 indicate a too stable nocturnal boundary layer which would lead to strong depletion of ozone near 563 564 the surface by dry deposition. Another reason could be the fact that no diurnal cycle is imposed on 565 NO_x emissions in the IFS-MOZART model which leads to too high emissions at nighttime and correspondingly too strong O₃ titration by NO at night. 566

These findings confirm our speculation that strong biases in O_3 and CO are to a large extent 567 due to differences in emission inventories between the MACC re-analysis and the regional models. 568 In Figures 12 and 13 we show the winter and summer average surface concentrations for O₃ 569 and CO over NA. Again, the superposed green dots on the maps represent the ground stations 570 selected for the analysis in Section 4.1.1 and 4.1.2. Similar to the EU case, the differences between 571 the regional models and the MACC re-analysis can be quite large within the interior of the domain 572 573 confirming the importance of the chosen emission inventories even for relatively long-lived 574 species such as CO and ozone. The CO concentrations are much larger in the regional models than in MACC especially in winter, consistent with the findings presented in Section 4.1.2. 575

The differences in summertime CO concentrations over Canada can be explained by the fact that SMARTFIREv2 covers only the US and no Canadian wildfire emissions were contained in the emission inventories used by the regional models. While ozone concentrations are rather similar in the mean of the regional models and the global model in summer, the MACC re-analysis shows generally much lower O₃ in winter similar to the EU case.

It is important to note that the regional models are run at much higher resolution than the MACC re-analysis. As a consequence, the regional models generally show much more small scale structure in the concentrations of O₃ and CO at the surface than the MACC re-analysis. Since there are other factors which could influence the surface concentrations beyond the differences in emission inventories, including the effect of vertical mixing, dry deposition and chemistry schemes, we stress that these differences deserve further analysis.

587

588 5. Summary and Conclusions

589

Sixteen European models and five North American models have been used to simulate the year 2010 within AQMEII-2. All models have taken their boundary conditions from the MACC reanalysis of the global IFS-MOZART model. All groups have provided their model outputs regridded on the same horizontal and vertical grids. We have evaluated the MACC re-analysis along with all other models against a set of ground station observations for O_3 , CO, NO, NO₂, SO₂ and SO₄²⁻ concentrations.

596 We have further investigated the ability of all models to reproduce surface O₃ concentrations

(see also Im et al., 2014a) and we have looked at midday and midnight values separately. This allowed us to isolate the different performance of the MACC re-analysis in simulating O₃ at different times of the day. Over the EU domain, the MACC re-analysis has a general tendency to underestimate O₃ at midnight. At midday, while there is still an underestimate during the winter season, the MACC re-analysis overestimates O₃. The general tendency to underestimate O₃ during winter is shown also in vertical profiles. Over the NA domain, the MACC re-analysis shows a good agreement with observations.

The longest lived species analyzed in this study was CO, for which the MACC re-analysis has negative biases over the European domain, mainly due to differences in emission inventories, while at remote NOAA stations near the domain boundaries the MACC re-analysis well reproduces CO concentrations.

The negative CO bias in the MACC re-analysis is also seen in the regional models at the surface, and in the vertical profile for all models (with the only exception of DE3). The analysis of surface concentrations of CO over the NA domain shows a much better agreement with observations both in the MACC re-analysis and regional model simulations. On the other hand, the regional models are missing wildfire emissions over Canada during 2010 summer season, since SMARTFIREv2 covers only the US and no Canadian wildfire emissions were contained in the emission inventories used by the regional models.

Although sensitivity simulations are crucial for a precise quantification of BC influence on any 615 numerical experiment, from our analysis clearly emerges that biases in the MACC re-analysis are 616 partly traced by the models, depending on the lifetime of the transported species. This influence is 617 most obvious for ozone, where the underestimation in winter in MACC is mimicked by the 618 619 regional models, while for CO the emissions in the interior of the domain appear to play an equally important role as the boundary conditions. The strong differences in CO between the global and 620 regional simulations are pointing at significant differences in the underlying emission inventories, 621 which calls for a better harmonization of regional and global inventories in the future. For the 622 623 shorter-lived species NO_x, SO₂ and sulfate, the influence of boundary conditions appears to be 624 minor. The sometimes large differences between the regional models and the MACC-reanalysis 625 as well as among the regional models themselves must be due to other factors. A particularly large spread between models and large differences from observations was found for sulfate, indicating 626 that the conversion of SO_2 to sulfate is often not well represented probably due to a 627

misrepresentation or lack of SO₂ oxidation in cloud water and through heterogeneous reactions on
the surface of aerosols.

630

631 6. Acknowledgements

632

We acknowledge NOAA/ESRL for providing their data through the WDCGG database and 633 making them available for use in research. We gratefully acknowledge the support of the European 634 groups through COST Action ES1004 EuMetChem. Individual authors of this article were 635 additionally supported by the following projects and grants: Lea Giordano through Swiss State 636 637 Secretariat for Education, Research and Innovation, project C11.0144. O. Jorba is supported by grant SEV-2011-00067 of Severo Ochoa Program awarded by the Spanish Government. The UPM 638 authors thankfully acknowledge the computer resources, technical expertise and assistance 639 provided by the Centro de Supercomputación y Visualización de Madrid (CESVIMA) and the 640 Spanish Supercomputing Network (BSC). Rahela Žabkar and Luka Honzak were supported by the 641 Centre of Excellence for Space Sciences and Technologies SPACE-SI, partly financed by the 642 European Union, European Regional Development Fund and Republic of Slovenia, Ministry of 643 Higher Education, Science, Sport and Culture. Y. Zhang acknowledges funding support from the 644 NSF Earth System Program (AGS-1049200) and computing support from Yellowstone by 645 NCAR's Computational and Information Systems Laboratory, sponsored by the National Science 646 647 Foundation and Stampede, provided as an Extreme Science and Engineering Discovery Environment (XSEDE) digital service by the Texas Advanced Computing Center (TACC). G. 648 649 Curci and P. Tuccella were supported by the Italian Space Agency (ASI) in the frame of PRIMES 650 project (contract n. I/017/11/0). The technical assistance of Bert van Ulft (KNMI) and Arjo Segers 651 (TNO) in producing the results of the NL2 model is gratefully acknowledged. The authors would also like to thank numerous data providers: the ECMW/MACC project & Météo-France/CNRM-652 653 GAME for chemical boundary conditions; the MOZAIC Data Centre and its contributing airlines for aircraft takeoff and landing vertical profiles. The Joint Research Center Ispra/Institute for 654 Environment and Sustainability provided its ENSEMBLE system for model output harmonization 655 and analyses and evaluation. The views expressed here are those of the authors and do not 656 657 necessarily reflect the views and policies of the U.S. EPA or any other organization participating

658 in the AQMEII project. This paper has been subjected to EPA review and approved for 659 publication.

660

661 **References**

- 662
- Badia, A. and Jorba, O., 2014. Gas-phase evaluation of the online NMMB/BSC-CTM model over
 Europe for 2010 in the framework of the AQMEII-Phase2 project, Atmospheric Environment,
 this issue, doi:10.1016/j.atmosenv.2014.05.055.
- Balzarini, A., Pirovano, G., Honzak, L., Žabkar, R., Curci, G., Forkel, R., Hirtl, M., San José, R.,
 Tuccella, P., Grell, G.A., 2014. WRF-Chem model sensitivity to chemical mechanisms choice
 in reconstructing aerosol optical properties. Atmospheric Environment, this issue,
- 669 doi:10.1016/j.atmosenv.2014.12.033.
- 670 Baklanov, A., Schlünzen, K., Suppan, P., Baldasano, J., Brunner, D., Aksoyoglu, S., Carmichael,
- G., Douros, J., Flemming, J., Forkel, R., Galmarini, S., Gauss, M., Grell, G., Hirtl, M., Joffre,
- 672 S., Jorba, O., Kaas, E., Kaasik, M., Kallos, G., Kong, X., Korsholm, U., Kurganskiy, A.,
- Kushta, J., Lohmann, U., Mahura, A., Manders-Groot, A., Maurizi, A., Moussiopoulos, N.,
- Rao, S., Savage, N., Seigneur, C., Sokhi, R., Solazzo, E., Solomos, S., Sörensen, B., Tsegas,
- G., Vignati, E., Vogel, B., Zhang, Y., 2014. Online coupled regional meteorology chemistry
 models in Europe: Current status and prospects. Atmospheric Chemistry and Physics 14 (1),
 317–398.
- 678 Brunner, D., Jorba, O., Savage, N., Eder, B., Makar, P., Giordano, L., Badia, A., Balzarini, A., Baro, R., Bianconi, R., Chemel, C., Forkel, R., Jimenez- Guerrero, P., Hirtl, M., Hodzic, A., 679 680 Honzak, L., Im, U., Knote, C., Kuenen, J., Makar, P., Manders-Groot, A., Neal, L., Perez, J., Pirovano, G., San Jose, R., Savage, N., Schroder, W., Sokhi, R., Syrakov, D., Torian, A., 681 682 Werhahn, J., Wolke, R., van Meijgaard, E., Yahva, K., Žabkar, R., Zhang, Y., Zhang, J., Hogrefe, C., Galmarini, S., 2014. Comparative analysis of meteorological performance of 683 coupled chemistry-meteorology models in the context of AQMEII phase 2. Atmospheric 684 Environment, this issue. 685
- 686 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., 2010. A framework for evaluating regional-scale
 numerical photochemical modeling systems. Environ Fluid Mech 10, 471-489.

- Duncan, B.N., Logan, J.A., Bey, I., Megretskaia, I.A., Yantosca, R.M., Novelli, P.C., Jones, N.B.,
 Rinsland, C.P., 2007. Global budget of CO, 1988–1997: Source estimates and validation with
- a global model, 2007. Journal of Geophysical Research 112, D22301,
 doi:10.1029/2007JD008459.
- Flemming, J., Inness, A., Flentje, H., Huijnen, V., Moinat, P., Schultz, M. G., Stein, O., 2009.
 Coupling global chemistry transport models to ECMWF's integrated forecast system.
 Geoscientific Model Development 2 (2), 253–265, doi:10.5194/gmd-2-253-2009.
- Flemming, J., Huijnen, V., Arteta, J., Bechtold, P., Beljaars, A., Blechschmidt, A.-M., Josse, B.,
 Diamantakis, M., Engelen, R. J., Gaudel, A., Inness, A., Jones, L., Katragkou, E., Marecal, V.,
 Peuch, V.-H., Richter, A., Schultz, M. G., Stein, O., and Tsikerdekis, A., 2014. Tropospheric
 chemistry in the integrated forecasting system of ECMWF, Geoscientific Model Development
 Discuss. 7, 7733-7803, doi:10.5194/gmdd-7-7733-2014.
- Galmarini, S., Bianconi, R., Appel, W., Solazzo, E., Mosca, S., Grossi, P., Moran, M., Schere, K.,
 Rao, S.T., 2012. ENSEMBLE and AMET: Two systems and approaches to a harmonized,
 simplified and efficient facility for air quality models development and evaluation.
 Atmospheric Environment 53, 51–59.
- Goode, J., Yokelson, R., Ward, D., Susott, R., Babbitt, R., Davies, M., Hao, W., 2000.
 Measurements of excess O₃, CO₂, CO, CH₄, C₂H₄, C₂H₂, HCN, NO, NH₃, HCOOH,
 CH₃COOH, HCHO, and CH₃OH in 1997 Alaskan biomass burning plumes by airborne Fourier
 transform infrared spectroscopy (A-FTIR). Journal of Geophysical Research D: Atmospheres
 105 (D17), 22147–22166.
- 710 Granier, C., Bessagnet, B., Bond, T., D'Angiola, A., van der Gon, H., Frost, G., Heil, A., Kaiser,
- J., Kinne, S., Klimont, Z., Kloster, S., Lamarque, J.-F., Liousse, C., Masui, T., Meleux, F.,
- Mieville, A., Ohara, T., Raut, J.-C., Riahi, K., Schultz, M., Smith, S., Thompson, A., van
- Aardenne, J., van der Werf, G., van Vuuren, D., 2011. Evolution of anthropogenic and biomass
- burning emissions of air pollutants at global and regional scales during the 1980-2010 period.
- 715 Climatic Change, 109 (1), 163–190, doi: 10.1007/s10584-011-0154-1.
- Grell, G., Peckham, S., Schmitz, R., McKeen, S., Frost, G., Skamarock, W., Eder, B., 2005. Fully
 coupled "online" chemistry within the WRF model. Atmospheric Environment 39 (37), 6957–
 6975.

- Hogrefe, C., Hao, W., Zalewsky, E. E., Ku, J.-Y., Lynn, B., Rosenzweig, C., Schultz, M. G., Rast,
 S., Newchurch, M. J., Wang, L., Kinney, P. L., and Sistla, G., 2011. An analysis of long-term
 regional-scale ozone simulations over the Northeastern United States: variability and trends,
 Atmos. Chem. Phys., 11, 567-582, doi:10.5194/acp-11-567-2011
- Im, U., Bianconi, R., Solazzo, E., Kioutsioukis, I., Badia, A., Balzarini, A., Baro, R., Bellasio, R.,
- 724 Brunner, D., Chemel, C., Curci, G., Flemming, J., Forkel, R., Giordano, L., Jimenez-Guerrero,
- P., Hirtl, M., Hodzic, A., Honzak, L., Jorba, O., Knote, C., Kuenen, J., Makar, P., Manders-
- Groot, A., Neal, L., Perez, J., Pirovano, G., Pouliot, G., San Jose, R., Savage, N., Schroder, W.,
- 727 Sokhi, R., Syrakov, D., Torian, A., Tuccella, P., Werhahn, J., Wolke, R., Yahya, K., Žabkar,
- R., Zhang, Y., Zhang, J., Hogrefe, C., Galmarini, S., 2014a. Evaluation of operational online-
- coupled regional air quality models over Europe and North America in the context of AQMEII
 phase 2. Part I: Ozone. Atmospheric Environment, this issue.
- Im, U., Bianconi, R., Solazzo, E., Kioutsioukis, I., Badia, A., Balzarini, A., Baro, R., Bellasio, 731 R., Brunner, D., Chemel, C., Curci, G., van der Gon, H. D., Flemming, J., Forkel, R., Giordano, 732 733 L., Jimenez-Guerrero, P., Hirtl, M., Hodzic, A., Honzak, L., Jorba, O., Knote, C., Makar, P., Manders-Groot, A., Neal, L., Perez, J., Pirovano, G., Pouliot, G., San Jose, R., Savage, N., 734 Schroder, W., Sokhi, R., Syrakov, D., Torian, A., Tuccella, P., Wang, K., Werhahn, J., Wolke, 735 R., Yahya, K., Žabkar, R., Zhang, Y., Zhang, J., Hogrefe, C., Galmarini, S., 2014b. Evaluation 736 of operational online-coupled regional air quality models over Europe and North America in 737 the context of AQMEII phase 2. Part II: Particulate matter. Atmospheric Environment, this 738 739 issue.
- Inness, A., Baier, F., Benedetti, A., Bouarar, I., Chabrillat, S., Clark, H., Clerbaux, C., Coheur, P.,
 Engelen, R. J., Errera, Q., Flemming, J., George, M., Granier, C., Hadji-Lazaro, J., Huijnen,
 V., Hurtmans, D., Jones, L., Kaiser, J. W., Kapsomenakis, J., Lefever, K., Leitão, J.,
 Razinger, M., Richter, A., Schultz, M. G., Simmons, A. J., Suttie, M., Stein, O., Thépaut, J.N., Thouret, V., Vrekoussis, M., Zerefos, C., and the MACC team, 2013. The MACC
 reanalysis: an 8 yr data set of atmospheric composition. Atmospheric Chemistry and Physics
 13 (8), 4073–4109, doi:10.5194/acp-13-4073-2013.
- Jaeglé, L., Steinberger, L., Martin, R. L., and Chance, K., 2005. Global partitioning of NO_x sources
 using satellite observations: relative roles of fossil fuel combustion, biomass burning and soil

- emissions. Faraday Discuss. 130, 407-423.
- Johnson, B. T., Osborne, S. R., 2011. Physical and optical properties of mineral dust aerosol
 measured by aircraft during the GERBILS campaign. Quarterly Journal of the Royal
 Meteorological Society 137 (658), 1117–1130, doi:10.1002/qj.777.
- 753 Kaiser, J. W., Heil, A., Andreae, M. O., Benedetti, A., Chubarova, N., Jones, L., Morcrette, J.-J.,
- Razinger, M., Schultz, M. G., Suttie, M., and van der Werf, G. R., 2012. Biomass burning
 emissions estimated with a global fire assimilation system based on observed fire radiative
 power, Biogeosciences, 9, 527-554, doi:10.5194/bg-9-527-2012.
- Katragkou, E., Zanis, P., Tegoulias, I., Melas, D., Kioutsioukis, I., Krüger, B. C., Huszar, P.,
 Halenka, T., and Rauscher, S., 2010. Decadal regional air quality simulations over Europe in
 present climate: near surface ozone sensitivity to external meteorological forcing, Atmos.
 Chem. Phys. 10, 11805-11821, doi:10.5194/acp-10-11805-2010.
- Knote, C., Brunner, D., Vogel, H., Allan, J., Asmi, A., Äijälä, M., Carbone, S., van der Gon, H.
 D., Jimenez, J. L., Kiendler-Scharr, A., Mohr, C., Poulain, L., Prévôt, A. S. H., Swietlicki, E.,
 Vogel, B., 2011. Towards an online-coupled chemistry-climate model: evaluation of trace
 gases and aerosols in COSMO-ART. Geoscientific Model Development 4 (4), 1077–1102,
 doi:10.5194/gmd-4-1077-2011.
- Knote, C. and Brunner, D., 2013. An advanced scheme for wet scavenging and liquid-phase
 chemistry in a regional online-coupled chemistry transport model, Atmospheric Chemistry and
 Physics 13, 1177-1192, doi:10.5194/acp-13-1177-2013.
- Knote, C., Tuccella, P., Curci, G., Emmons, L., Orlando, J.J., Madronich, S., Baró, R., JiménezGuerrero, P., Luecken, D., Hogrefe, C., Forkel, R., Werhahn, J., Hirtl, M., Pérez, J.L., San José,
 R., Giordano, L., Brunner, D., Yahya, K., Zhang, Y., Influence of the choice of gas-phase
 mechanism on predictions of key gaseous pollutants during the AQMEII phase-2
 intercomparison. Atmospheric Environment, this issue.
- Kuenen, J. J. P., Visschedijk, A. J. H., Jozwicka, M., Denier van der Gon, H. A. C., 2014. TNO
- 775 MACC II emission inventory: a multi year (2003-2009) consistent high-resolution European
- emission inventory for air quality modelling. Atmospheric Chemistry and Physics Discussions
 14 (5), 5837–5869., doi:10.5194/acpd-14-5837-2014.
- Lee, C., Martin, R. V., van Donkelaar, A., Lee, H., Dickerson, R. R., Hains, J. C., Krotkov, N.,

- Richter, A., Vinnikov, K., Schwab, J. J., 2011. SO₂ emissions and lifetimes: Estimates from
 inverse modeling using in situ and global, space-based (SCIAMACHY and OMI) observations.
 Journal of Geophysical Research: Atmospheres 116 (D6), doi:10.1029/2010JD014758.
- 782 Makar, P. A., Zhang, J., Gong, W., Stroud, C., Sills, D., Hayden, K. L., Brook, J., Levy, I., Mihele,
- C., Moran, M. D., Tarasick, D. W., He, H., and Plummer, D., 2010. Mass tracking for chemical
 analysis: the causes of ozone formation in southern Ontario during BAQS-Met 2007, Atmos.
- 785 Chem. Phys. 10, 11151-11173, doi:10.5194/acp-10-11151-2010.
- Miller, S. M., Matross, D. M., Andrews, A. E., Millet, D. B., Longo, M., Gottlieb, E. W., Hirsch,
 A. I., Gerbig, C., Lin, J. C., Daube, B. C., Hudman, R. C., Dias, P. L. S., Chow, V. Y., and
 Wofsy, S. C., 2008. Sources of carbon monoxide and formaldehyde in North America
 determined from high-resolution atmospheric data, Atmospheric Chemistry and Physics 8,
 7673-7696, doi:10.5194/acp-8-7673-2008.
- Morcrette, J.J., Boucher, O., Jones, L., Salmond, D., Bechtold, P., Beljaars, A., Benedetti, A.,
 Bonet, A., Kaiser, J.W., Razinger, M., Schulz, M., Serrar, S., Simmons, A.J., Sofiev, M., Suttie,
 M., Tompkins, A.M., Untch, A., 2009. Aerosol analysis and forecast in the European Centre
 for Medium-Range Weather Forecasts Integrated Forecast System: Forward modeling. Journal
 of Geophysical Research: Atmosphere 114, D06206, doi: 10.1029/2008JD011235.
- Nédélec, P., Cammas, J.-P., Thouret, V., Athier, G., Cousin, J.-M., Legrand, C., Abonnel, C.,
 Lecoeur, F., Cayez, G., Marizy, C., 2003. An improved infrared carbon monoxide analyser for
 routine measurements aboard commercial airbus aircraft: technical validation and first
 scientific results of the MOZAIC IIIi programme. Atmospheric Chemistry and Physics 3 (5),
 1551–1564, doi:10.5194/acp-3-1551-2003.
- Novelli, P., Masarie, K., 2009. Atmospheric carbon monoxide dry air mole fractions from the
 NOAA ESRL carbon cycle cooperative global air sampling network, 1988-2008. Tech. rep.,
 NOAA, ftp://ftp.cmdl.noaa.gov/ccg/co/flask/event/
- Novelli, P., Masarie, K., 2013. Atmospheric carbon monoxide dry air mole fractions from the
- 805 NOAA ESRL carbon cycle cooperative global air sampling network, 1988-2012, version:
- 806 2013-06-18. Tech. rep., NOAA, ftp://aftp.cmdl.noaa.gov/data/trace_gases/co/flask/surface/
- 807 Pouliot, G., Denier van der Gon, H., Kuenen, J., Makar, P., Zhang, J., Moran, M., 2014. Analysis
- of the emission inventories and model-ready emission datasets of Europe and North America

- Rao, S., Galmarini, S., Keith, P., 2010. Air Quality Model Evaluation International Initiative
 (AQMEII): Advancing the state of the science in regional photochemical modeling and its
 applications. Bulletin of the American Meteorological Society 92, 23-30.
- 813 Schaub, D., Brunner, D., Boersma, K., Keller, J., Folini, D., Buchmann, B., Berresheim, H.,
- 814 Staehelin, J., 2007. SCIAMACHY tropospheric NO₂ over Switzerland: Estimates of NO_x
- 815 lifetimes and impact of the complex alpine topography on the retrieval. Atmospheric Chemistry

and Physics 7 (23), 5971–5987.

- 817 Schere K., J. Flemming, R. Vautard, C. Chemel, A. Colette, C. Hogrefe, B. Bessagnet, F.
- 818 Meleux, R. Mathur, S. Roselle, R.-Ming Hu, R. S. Sokhi, S. T.Rao, S. Galmarini, 2012. Trace
- gas/aerosol boundary concentrations and their impacts on continental-scale AQMEII
- modeling domains, Atmospheric Environment 53, 38-50,
- 821 <u>doi:10.1016/j.atmosenv.2011.09.043</u>.
- 822 Seinfeld, J., Pandis, S., 2006. Atmospheric chemistry and physics: from air pollution to climate
- change. A Wiley-Interscience publications. Wiley, University of Michigan, USA, ISBN
 0471720186, 9780471720188, 1232 pp.
- Sofiev, M., Vankevich, R., Lotjonen, M., Prank, M., Petukhov, V., Ermakova, T., Koskinen, J.,
 Kukkonen, J., 2009. An operational system for the assimilation of the satellite information on
 wild-land fires for the needs of air quality modelling and forecasting. Atmospheric Chemistry
 and Physics 9 (18), 6833–6847.
- Solazzo, E., Bianconi, R., Pirovano, G., Moran, M. D., Vautard, R., Hogrefe, C., Appel, K. W.,
 Matthias, V., Grossi, P., Bessagnet, B., Brandt, J., Chemel, C., Christensen, J. H., Forkel, R.,
- Francis, X. V., Hansen, A. B., McKeen, S., Nopmongcol, U., Prank, M., Sartelet, K. N.,
- Segers, A., Silver, J. D., Yarwood, G., Werhahn, J., Zhang, J., Rao, S. T., Galmarini, S., 2013.
- Evaluating the capability of regional-scale air quality models to capture the vertical distribution
- of pollutants. Geoscientific Model Development 6 (3), 791–818, doi:10.5194/gmd-6-791-2013.
- 836 Steeneveld, G., Van de Wiel, B., Holtslag, A., 2006. Modeling the evolution of the atmospheric
- boundary layer coupled to the land surface for three contrasting nights in CASES-99. Journal
- of the Atmospheric Sciences 63 (3), 920-935.

for phase 2 of the AQMEII project, Atmospheric Environment, this issue.

- 839 Stein, O., Schultz, M., Flemming, J., Inness, A., Kaiser, J., Jones, L., Benedetti, A., Morcrette, J.-
- J., 2011. MACC global air quality services-technical documentation. Tech. rep.,
- 841 Forschungszentrum Jülich and ECMWF.
- 842 Stein, O., Schultz, M. G., Bouarar, I., Clark, H., Huijnen, V., Gaudel, A., George, M., and
- 843 Clerbaux, C., 2014. On the wintertime low bias of Northern Hemisphere carbon monoxide
- found in global model simulations, Atmospheric Chemistry and Physics 14, 9295-9316,
- doi:10.5194/acp-14-9295-2014.
- Steinbacher, M., Zellweger, C., Schwarzenbach, B., Bugmann, S., Buchmann, B., Ordez, C.,
 Prevot, A., Hueglin, C., 2007. Nitrogen oxide measurements at rural sites in Switzerland: Bias
 of conventional measurement techniques, Journal of Geophysical Research: Atmospheres 112
- 849 (11), doi:10.1029/2006JD007971.
- Thouret, V., Marenco, A., Logan, J. A., Nédélec, P., Grouhel, C., 1998. Comparisons of ozone
 measurements from the MOZAIC airborne program and the ozone sounding network at eight
 locations. Journal of Geophysical Research: Atmospheres, 103 (D19), 25695–25720,
 doi:10.1029/98JD02243.
- Wang, K., Yahya, Y., Zhang, S.-Y. W., Grell, G., 2014. Implementation and initial application of
 a new chemistry-aerosol option in WRF-Chem for simulation of secondary organic aerosols
 and aerosol indirect effects. Atmospheric Environment, this issue.
- 857

Tables

Table	1:	Overview	of	Participating	Models
-------	----	----------	----	---------------	--------

Domain	Group	Specifications	Model
EU	BG2	offline	WRF-CMAQ
EU	NL2	online access	RACMO LOTUS-EUROS
EU	DE3	online access	COSMO-MUSCAT
EU	UK5	online access	WRF-CMAQ
NA	US6	online access	WRF-CMAQ
EU	AT1	online integrated	WRF-CHEM
EU	DE4	online integrated	WRF-CHEM
EU	ES1	online integrated	WRF-CHEM
EU	ES3	online integrated	WRF-CHEM
EU	IT1	online integrated	WRF-CHEM
EU	IT2	online integrated	WRF-CHEM
EU	SI1	online integrated	WRF-CHEM
EU	US6	online integrated	WRF-CHEM
EU	US8	online integrated	WRF-CHEM
EU	CH1	online integrated	COSMO-ART
EU	ES2a	online integrated	NMMB-BSC-CTM
EU	ES2b	online integrated	NMMB-BSC-CTM
EU	UK4	online integrated	METUM-UKCA RAQ
NA	CA2	online integrated	GEM-MACH
NA	CA2f	online integrated	GEM-MACH

	Number of stat	tions
Chemical		
species		
	EU	NA
O ₃	401	271
со	48	17
NO	211	15
NO ₂	285	59
SO ₂	165	44
SO ₄	33	70

Table 2: Stations selection

Model	O ₃									
	mean (p	opb)	NMB (9	%)	RMSE		r			
	S	W	S	W	(ppb) S	W	S	W		
MACC (EU)	54.01	18.37	26.4	-26.04	16.43	11.94	0.58	0.51		
BG2	50.47	27.23	17.98	9.68	14.65	15.97	0.56	0.31		
NL2	46.78	22.55	9.4	-9.23	12.14	10.1	0.64	0.5		
DE3	44.33	17.52	3.65	-29.41	12.31	13.07	0.54	0.41		
UK5	56.76	29.44	32.8	18.64	18.29	11.07	0.62	0.53		
AT1	41.12	21.43	-3.81	-13.67	11.22	10.92	0.63	0.42		
DE4	46.05	19.79	7.7	-20.32	11.8	11.43	0.62	0.47		
ES1	39.35	20.11	-7.96	-19.01	11.64	11.31	0.63	0.41		
ES3	38.64	20.5	-9.61	-17.41	12.14	10.8	0.61	0.49		
IT1	37.92	20.57	-11.3	-17.18	12.62	10.83	0.58	0.49		
IT2	35.27	19.69	-17.53	-20.59	13.47	11.34	0.63	0.48		
SI1	41.2	22.48	-3.62	-9.47	11.29	10.39	0.63	0.43		
SI2	41.07	22.54	-3.94	-9.21	11.26	10.32	0.63	0.44		
CH1	41.52	15.78	-2.9	-36.46	13.84	13.23	0.48	0.53		
ES2a	50.36	16.61	17.8	-33.12	14.75	13.22	0.6	0.52		
ES2b	48.72	15.44	13.96	-37.85	14.01	13.85	0.61	0.53		
UK4	55.36	16.66	29.47	-32.92	18.68	13.05	0.61	0.54		
MACC	52.7	25.95	23.13	-23.69	15.46	12.14	0.64	0.45		
(NA)										
CA2f	45.88	32.47	7.21	-4.55	13.59	8.64	0.59	0.48		
CA2	46.52	32.5	8.69	-4.46	13.88	8.64	0.59	0.48		
US6	48.94	30.79	14.04	-9.84	11.6	10.45	0.75	0.49		
US7	51.64	28.86	20.35	-15.56	17.28	10.46	0.55	0.46		
US8	40.64	25.06	-5.21	-26.4	11.64	12.38	0.64	0.53		

Table 3: Seasonal (summer and winter) midday model perfomance metrics for O_3 at ground stations

Model	со									
	mean (p	opb)	NMB (9	%)	RMSE		r			
					(ppb)					
	S	W	S	W	S	W	S	W		
MACC (EU)	117.19	150.91	-32.00	-53.12	141.85	306.80	0.13	0.19		
BG2	93.09	166.89	-45.89	-48.14	152.82	296.80	0.14	0.21		
NL2	106.62	150.42	-38.32	-53.25	149.64	305.42	0.07	0.23		
DE3	132.67	203.43	-23.19	-36.78	139.26	278.64	0.09	0.22		
UK5	105.51	175.38	-38.69	-45.47	147.43	289.45	0.09	0.26		
AT1	111.01	153.74	-35.58	-52.27	146.93	304.07	0.09	0.22		
DE4	107.51	156.72	-37.60	-51.33	148.46	302.55	0.09	0.21		
ES1	107.69	155.41	-37.50	-51.72	148.17	303.30	0.09	0.21		
ES3	98.51	157.62	-42.83	-51.05	152.40	301.77	0.08	0.22		
IT1	96.37	154.36	-44.04	-51.11	153.76	300.87	0.07	0.21		
IT2	94.91	154.41	-44.88	-50.55	154.58	301.03	0.08	0.21		
SI1	111.08	156.60	-35.53	-51.37	147.11	302.43	0.09	0.22		
SI2	110.61	156.41	-35.80	-51.43	147.03	302.56	0.09	0.22		
CH1	110.26	167.84	-35.98	-47.80	151.80	296.95	0.03	0.19		
ES2a	113.86	177.61	-33.86	-44.91	142.34	287.87	0.17	0.27		
ES2b	117.65	178.97	-31.64	-44.46	141.45	287.41	0.15	0.27		
UK4	126.68	198.61	-27.56	-38.37	179.44	274.32	0.05	0.32		
MACC	123.06	127.59	-2.68	-8.16	109.74	125.63	0.3	0.32		
(NA)										
CA2f	167.7	189.57	36.5	37.12	184.57	172.56	0.16	0.05		
CA2	169.84	192.94	38.24	39.58	185.18	176.63	0.16	0.05		
US6	145.21	170.43	16.66	23.22	119.97	159.91	0.2	0.05		
US7	159.48	192.56	27.49	38.96	115.22	158.97	0.34	0.06		
US8	164.81	198.89	32.35	43.44	139.74	155.7	0.27	0.13		

Table 4: Seasonal (summer and winter) midday model perfomance metrics for CO at ground stations

Model	NOx									
	mean (ppb)		NMB (9	NMB (%)			r			
					(ppb)					
	S	W	S	W	S	W	S	W		
MACC	0.95	7.32	-76.73	-45.64	4.47	12.69	0.15	0.31		
BG2	2.19	13.57	-46.22	0.58	3.94	11.09	0.24	0.42		
NL2	1.43	6.6	-65.04	-51.11	4.21	12.82	0.17	0.37		
DE3	2.49	5.58	-38.85	-58.69	3.87	13.46	0.18	0.35		
UK5	1.34	8.11	-67.24	-39.83	4.28	11.95	0.24	0.41		
AT1	1.28	6.06	-68.64	-55.04	4.35	13.23	0.19	0.33		
DE4	1.3	6.53	-68.1	-51.56	4.36	12.91	0.18	0.35		
ES1	1.19	6.13	-70.77	-54.46	4.49	13.18	0.15	0.33		
ES3	1.27	5.87	-68.8	-56.42	4.35	13.27	0.17	0.36		
IT1	1.38	6.25	-66.02	-52.33	4.28	13.04	0.18	0.35		
IT2	1.87	6.57	-54.06	-48.38	4.06	12.8	0.19	0.36		
SI1	1.35	5.82	-67.02	-56.79	4.34	13.35	0.18	0.34		
SI2	1.31	5.73	-67.75	-57.42	4.32	13.37	0.19	0.35		
CH1	2.44	10.57	-38.3	-21.6	9.7	11.43	0.05	0.39		
ES2a	2.15	17.14	-47.06	26.81	4.71	14.65	0.14	0.35		
ES2b	2.63	17.69	-35.2	30.94	5	14.68	0.13	0.36		
UK4	2.44	13.31	-40.04	-1.47	4.42	12.23	0.21	0.41		
MACC	0.9	4.89	-74.73	-52.53	4.4	11.66	0.29	0.42		
CA2f	2.35	7.24	-33.51	-29.4	4.6	11.14	0.26	0.46		
CA2	2.31	7.66	-34.52	-25.24	4.58	11.21	0.27	0.47		
US6	2.37	10.69	-29.06	9.29	3.76	11.29	0.37	0.51		
US7	2.57	12.3	-23.55	23.27	3.47	12.45	0.42	0.44		
US8	2.29	10.52	-30.86	6.75	4.38	12.53	0.24	0.39		

Table 5: Seasonal (summer and winter) midday model performance metrics for NO_x at ground stations

Model	SO ₂									
	mean (p	opb)	NMB (9	NMB (%)			r			
					(ppb)					
	S	W	S	W	S	W	S	W		
MACC (EU)	0.32	2.56	-63.48	39.89	2.34	3.03	0.1	0.49		
BG2	0.55	1.66	-36.74	-9.06	2.17	2.77	0.34	0.5		
NL2	0.38	0.82	-56.45	-54.9	2.23	3.08	0.3	0.4		
DE3	0.64	0.96	-27.13	-47.51	2.21	3.06	0.27	0.39		
UK5	0.2	0.42	-76.76	-77	2.29	3.34	0.34	0.28		
AT1	0.48	0.86	-44.74	-53.11	2.19	2.99	0.32	0.45		
DE4	0.5	1.24	-43	-31.92	2.19	2.92	0.32	0.42		
ES1	0.33	0.91	-61.78	-50.35	2.23	2.97	0.33	0.45		
ES3	0.47	0.94	-46.32	-48.49	2.19	2.97	0.32	0.44		
IT1	0.52	0.89	-40.48	-50.96	2.18	2.99	0.33	0.46		
IT2	0.53	1.25	-39.13	-31.3	2.18	2.93	0.32	0.43		
SI1	0.37	0.79	-57.59	-56.69	2.22	3.02	0.32	0.45		
SI2	0.36	0.79	-58.61	-56.77	2.22	3.02	0.32	0.45		
CH1	0.67	1.59	-24.11	-12.83	2.27	2.91	0.24	0.41		
ES2a	0.73	1.55	-16.59	-14.93	2.17	2.88	0.32	0.45		
ES2b	0.77	1.42	-12.4	-22.02	2.18	2.92	0.31	0.43		
UK4	0.74	1.65	-15.38	-9.78	2.26	3.07	0.3	0.46		
MACC (NA)	0.42	2.38	-81.25	-13.96	5.67	7.15	0.16	0.19		
CA2f	2.52	3.64	12.34	33.71	5.32	7.38	0.34	0.27		
CA2	2.54	3.86	13.04	41.63	5.34	7.49	0.34	0.27		
US6	1.4	1.91	-29.24	-31.66	4.6	6.94	0.41	0.41		
US7	1.47	2.12	-26.06	-25.71	4.6	6.91	0.4	0.42		
US8	1.16	1.7	-37.99	-38.71	4.74	7.11	0.36	0.38		

Table 6: Seasonal (summer and winter) midday model performance metrics for SO_2 at ground stations

Table 7: Seasonal (summer and winter) daily model performance metrics for SO^{2-} at₄ground stations

Model	SO ₄ ²⁻								
	mean (ppb)		NMB (%)		RMSE (ppb)		r		
	S	W	S	W	S	W	S	W	

MACC	3.19	1.7	89.13	-29.37	2.56	2.54	0.43	0.41
(EU)								
BG2	1.94	2.28	15.59	-5.46	1.26	2.11	0.59	0.57
NL2	1.77	1.35	5.64	-44.13	1.56	2.61	0.39	0.36
DE3	1.15	2.28	-28.72	-5.72	1.29	2.18	0.55	0.54
UK5	1.54	1.62	-8.33	-32.88	1.37	2.36	0.44	0.49
AT1	1.11	0.78	-34.22	-67.82	1.31	2.74	0.52	0.55
DE4	1.07	0.36	-36.08	-84.97	1.33	3.22	0.51	0.2
ES1	0.95	0.42	-43.14	-82.62	1.4	3.1	0.47	0.42
ES3	1.5	0.8	-11.01	-66.97	1.42	2.73	0.5	0.53
IT1	1.46	0.65	-13.4	-72.92	1.42	2.95	0.47	0.36
IT2	0.91	0.38	-45.68	-83.1	1.39	3.2	0.52	0.2
SI1	2.22	2.16	31.67	-10.24	1.74	2.21	0.45	0.57
SI2	2.22	2.16	31.53	-10.29	1.75	2.21	0.45	0.57
CH1	2.87	2.21	67.06	-8.37	2.13	2.46	0.4	0.46
MACC	2.84	3.02	48.19	130.85	1.75	2.85	0.58	0.44
(NA)								
CA2f	4.48	2.07	132.31	52.64	4.06	1.53	0.65	0.64
CA2	4	1.49	107.91	12.36	3.45	1.14	0.68	0.6
US6	1.62	1.01	-15.58	-23.16	0.99	0.89	0.82	0.68
US7	2.34	0.91	20.91	-29.98	1.23	1.18	0.73	0.33
US8	1.64	1.53	-14.8	14.98	1.12	1.4	0.77	0.68

6 Figures





Figure 1: Time series for 2010 of median monthly O_3 concentrations, midday values, for EU (top, 12–14 UTC) and NA (bottom, 19–21 UTC) domains. The gray shaded area is the interquartile range of the distribution of observations across all selected stations. The white line represents the median of the observations.





Figure 2: Time series for 2010 of median monthly O₃ concentrations, midnight values, for EU (top, 00–02 UTC) and NA (bottom, 07–09 UTC) domains. The gray shaded area is the interquartile range of the distribution of observations across all selected stations. The white line represents the median of the observations.



Figure 3: 2010 time series of mean CO concentrations from the MACC model (black line) and from NOAA/GMD ground-based measurements (red points) at Mace Head, Black Sea, Key Biscayne and Park Falls stations.





Figure 4: Time series for 2010 of median monthly CO concentrations, midday values, for EU (top, 12–14 UTC) and NA (bottom, 19–21 UTC) domains. The gray shaded area is the interquartile range of the distribution of observations across all selected stations. The white line represents the median of the observations.





Figure 5: Time series for 2010 of median monthly NO concentrations, midday values, for EU (top, 12–14 UTC) and NA (bottom, 19–21 UTC) domains. The gray shaded area is the interquartile range of the distribution of observations across all selected stations. The white line represents the median of the observations.



AUG

SEP

OCT

NOV

JUL

DEC



50

51

52

47 48 49

JAN

FEB

MAR

APR

Figure 6: Time series for 2010 of median monthly NO₂ concentrations, midday values, for EU (top, 12–14 UTC) and NA (bottom, 19–21 UTC) domains. The gray shaded area is the interquartile range of the distribution of observations across all selected stations. The white line represents the median of the observations.

MAY

JUN







Figure 7: Time series for 2010 of median monthly SO₂ concentrations, midday values, for EU (top, 12–14 UTC) and NA (bottom, 19–21 UTC) domains. The gray shaded area is the interquartile range of the distribution of observations across all selected stations. The white line represents the median of the observations.







Figure 8: Top panel: time series for 2010 of median monthly PM_{10} -SO₄ concentrations for EU. Bottom panel: time series for 2010 of median monthly $PM_{2.5}$ -SO₄ concentrations for NA. The gray shaded area is the interquartile range of the distribution of observations across all selected stations. The white line represents the median of the observations.



Figure 9: 2010 winter vertical profiles of median O_3 (top) and CO (bottom) concentrations at Frankfurt airport. The gray shaded area is the interquartile range of the distribution of observations across all selected stations. The white line represents the median of the MOZAIC observations.



O3 MACC Winter [ppb]





O3 EU Summer Mean [ppb]







Figure 10: Winter (top panels) and summer (bottom panels) surface 2010 means for O_3 in the EU case. Right: MACC re-analysis. Left: Average over the 16 regional models that simulated the EU case. The superposed green dots on the maps represent the ground stations selected for the analysis at receptor points. Stripes near the domain borders are due to the different model domains of the regional models leading to a varying number of models contributing to the mean.



Figure 11: Winter (top panels) and summer (bottom panels) surface 2010 means for CO in the EU case. Right: MACC re-analysis. Left: Average over the 16 regional models that simulated the EU case. The superposed green dots on the maps represent the ground stations selected for the analysis at receptor points. Stripes near the domain borders are due to the different model domains of the regional models leading to a varying number of models contributing to the mean.

O3 NA Winter Mean [ppb]

O3 MACC Winter [ppb]







55 55 20 20 45 45 latitude latitude 5 5 35 35 30 30 25 52 -130 -120 -110 -100 -90 -80 -70 -60 -130 -120 -110 -100 -90 -80 -70 longitude longitude CO NA Summer Mean [ppb] CO MACC Summer [ppb] 55



CO NA Winter Mean [ppb]



CO MACC Winter [ppb]





115

116

117

Figure 13: Winter (top panels) and Summer (bottom panels) surface 2010 means for CO in the NA case. Right: MACC re-analysis. Left: Average over the 5 regional models that simulated the NA case. The superposed green dots on the maps represent the ground stations selected for the analysis at receptor points. Stripes near the domain borders are due to the different model domains of the regional models leading to a varying number of models contributing to the mean.

118 119

111

-60