A regional air quality forecasting system over Europe: the MACC-II daily ensemble production

3

V. Marécal¹, V.-H. Peuch², C. Andersson³, S. Andersson³, J. Arteta¹, M. 4 Beekmann⁴, A. Benedictow⁵, R. Bergström³, B. Bessagnet⁶, A. Cansado⁷, F. 5 Chéroux¹, A. Colette⁶, A. Coman⁴, R. L. Curier⁸, H. A. C. Denier van der Gon⁸, A. 6 Drouin¹, H. Elbern⁹, E. Emili¹⁰, R. J. Engelen², H. J. Eskes¹¹, G. Foret⁴, E. Friese⁹, 7 M. Gauss⁵, C. Giannaros¹², J. Guth¹, M. Joly¹, E. Jaumouillé¹⁰, B. Josse¹, N. 8 Kadygrov¹, J. W. Kaiser¹³, K. Krajsek¹⁴, J. Kuenen⁸, U. Kumar¹¹, N. Liora¹², E. 9 Lopez⁷, L. Malherbe⁶, I. Martinez⁷, D. Melas¹², F. Meleux⁶, L. Menut¹⁵, P. 10 Moinat¹⁰, T. Morales⁷, J. Parmentier¹, A. Piacentini¹⁰, M. Plu¹, A. Poupkou¹², S. 11 Queguiner¹, L. Robertson³, L. Rouïl⁶, M. Schaap⁸, A. Segers⁸, M. Sofiev¹⁶, L. 12

¹³ Tarasson¹⁷, M. Thomas³, R. Timmermans⁸, Á. Valdebenito⁵, P. van Velthoven¹¹,

- 14 R. van Versendaal¹¹, J. Vira¹⁶ and A. Ung⁶
- 15 [1] {Groupe d'étude de l'Atmosphère Méréorologique/Centre National de Recherches
- 16 Météorologiques, CNRS-Météo-France, UMR 3589, Toulouse, France}
- 17 [2] {European Centre for Medium-range Weather Forecasts, Reading, UK}
- 18 [3] {Swedish Meteorological and Hydrological Institute, Norrköping, Sweden}
- 19 [4] {Laboratoire Inter-universitaire des Systèmes Atmosphériques, UMR CNRS 7583,
- 20 Université Paris Est Créteil et Université Paris Diderot, Créteil, France}
- 21 [5] {Norwegian Meteorological Institute, Oslo, Norway}
- 22 [6] {Institut National de l'Environnement Industriel et des Risques, Parc Technologique
- 23 Alata, 60550 Verneuil en Halatte, France}
- 24 [7] {AEMET Spanish Meteorological State Agency, Leonardo Prieto Castro 8, 28040, Spain}
- 25 [8] {TNO, Climate Air and Sustainabilty Unit, Utrecht, The Netherlands}
- 26 [9] {Rhenish Institute for Environmental Research at the University of Cologne, Cologne,
- 27 Germany}
- 28 [10] {CERFACS, URA 1875, Toulouse, France}
- 29 [11] {Royal Netherlands Meteorological Institute, De Bilt, The Netherlands}

- 30 [12] {Laboratory Of Atmospheric Physics, Physics Dept., Aristotle University of
- 31 Thessaloniki, Thessaloniki, Greece}
- 32 [13] {Max Planck Institute for Chemistry, Mainz, Germany}
- 33 [14] {Institut für Energie- und Klimaforschung (IEK-8), Forschungszentrum Jülich, Jülich,
- 34 Germany}
- 35 [15] {Laboratoire de Météorologie Dynamique, Ecole Polytechnique, 91128 Palaiseau,
- 36 France}
- 37 [16] {Finnish Meteorological Institute, Erik Palmenin Aukio 1, Helsinki 00560, Finland}

38 [17] {Norwegian Institute for Air Research, 2027 Kjeller, Norway}

- 39 Correspondence to: Virginie Marécal (virginie.marecal@meteo.fr)
- 40

41 Abstract

This paper describes the pre-operational analysis and forecasting system developed during 42 43 MACC (Monitoring Atmospheric Composition and Climate) and continued in MACC-II (Monitoring Atmospheric Composition and Climate: Interim Implementation) European 44 45 projects to provide air quality services for the European continent. This system is based on seven state-of-the art models developed and run in Europe (CHIMERE, EMEP, EURAD-IM, 46 47 LOTOS-EUROS, MATCH, MOCAGE and SILAM). These models are used to calculate multi-model ensemble products. The paper gives an overall picture of its status at the end of 48 49 MACC-II (summer 2014) and analyses the performance of the multi-model ensemble. The MACC-II system provides daily 96h forecasts with hourly outputs of 10 chemical 50 species/aerosols (O₃, NO₂, SO₂, CO, PM₁₀, PM_{2.5}, NO, NH₃, total NMVOCs and PAN+PAN 51 precursors) over 8 vertical levels from the surface to 5km height. The hourly analysis at the 52 surface is done *a posteriori* for the past day using a selection of representative air quality data 53 from European monitoring stations. 54

The performance of the system is assessed daily, weekly and 3 monthly (seasonally) through statistical indicators calculated using the available representative air quality data from European monitoring stations. Results for a case study show the ability of the ensemble median to forecast regional ozone pollution events. The seasonal performances of the individual models and of the multi-model ensemble have been monitored since September 2009 for ozone, NO₂ and PM₁₀. The statistical indicators for ozone in summer 2014 show that

the ensemble median gives best performances on average compared to the seven models. 61 There is very little degradation of the scores with the forecast day but there is a marked 62 diurnal cycle, similarly to the individual models, that can be related partly to the prescribed 63 diurnal variations of anthropogenic emissions in the models. During summer 2014, the diurnal 64 ozone maximum is underestimated by the ensemble median by about 4 μ g m⁻³ on average. 65 Locally, during the studied ozone episodes, the maxima from the ensemble median are often 66 lower than observations by 30 to 50 μ g m⁻³. Overall, ozone scores are generally good with 67 average values for the normalized indicators of 0.14 for the Modified Normalized Mean Bias 68 69 and of 0.30 for the Fractional Gross Error. Tests have also shown that the ensemble median is robust to reduction of ensemble size by one, that is if predictions are unavailable from one 70 model. Scores are also discussed for PM₁₀ for winter 2013-1014. There is an underestimation 71 of most models leading for the ensemble median to a Mean Bias of -4.5 μ g m⁻³. The ensemble 72 median Fractional Gross Error is larger for PM_{10} (~0.52) than for ozone and the correlation is 73 lower (~0.35 for PM_{10} and ~0.54 for ozone). This is related to a larger spread of the 7 model 74 scores for PM₁₀ than for ozone linked to different levels of complexity of aerosol 75 representation in the individual models. In parallel, a scientific analysis of the results of the 76 seven models and of the ensemble is also done over the Mediterranean area because of the 77 specificity of its meteorology and emissions. 78

79 The system is robust in terms of the production availability. Major efforts have been done in 80 MACC-II towards the operationalisation of all its components. Foreseen developments and 81 research for improving its performances are discussed in the conclusion.

82

83 **1. Introduction**

The chemical composition of the air close to the Earth's surface, generally referred as 'air 84 quality' (AQ), directly affects human and animal health and also the vegetation. For instance, 85 ozone has a known impact on the respiratory system (e.g. WHO, 2004) and on the vegetation 86 development (e.g. Fuhrer and Booker, 2003). Recently, the World Health Organisation 87 reported that in 2012 around 3.7 million people deaths are attributable to ambient air pollution 88 (http://www.who.int/phe/health topics/outdoorair/databases/en/). This is why air quality has 89 90 become a major concern, starting in the 1970's, in particular in Europe (e.g. WHO, 2013). Since the Helsinki Protocol in 1985, many regions and countries, including the European 91 92 Union countries, have progressively put in place tools to regulate and to control the emissions

93 of the main air pollutants. This has led to an important effort to monitor the air composition 94 near the surface but also to develop air quality forecasting systems in experimental or 95 operational modes (see reviews by Ebel et al. 2005, Menut and Bessagnet 2010). These tools 96 can be used in cases of high pollution episodes to inform people and to take emergency 97 measures to prevent harming effects. They can also be used for policy makers for the 98 regulations on air pollutant emissions and for monitoring the effect of these regulations on air 99 quality (episodes and also background pollution).

100 The main pollutants under focus for air quality are ozone, nitrogen oxides (NOx= $NO_2 + NO$), sulphur dioxide (SO₂), Volatile Organic Compounds (VOCs), ammonia (NH₃), particulate 101 102 matter, heavy metals (Pb, Cd, Hg) and Persistent Organic Pollutants (POPs, e.g. pesticides and dioxine). Ozone is a secondary pollutant meaning that it is not emitted but produced from 103 104 gaseous precursors (mainly VOCs and NOx) originating from both natural and anthropogenic 105 sources. Particulate matter (PM) corresponds to small size aerosols. PMs are categorised as PM_{10} (size < 10 µm), $PM_{2.5}$ (size < 2.5 µm) and PM_1 (size < 1 µm). These categories were 106 107 chosen because of their known effects on health. In PMs, the distinction between primary (dust, sea salts, black carbon and organic carbon) and secondary aerosols formed from 108 109 gaseous precursors such as SO₂, DMS, H₂S, NH₃. NOx and VOCs is ignored when considering mass or number concentration only. 110

111 Besides the development of surface measurement networks for these main pollutants, there has been a sustained research effort on the atmospheric chemistry modelling for air quality 112 forecasting purposes. Regional and local air quality forecasting systems (Kukkonen et al. 113 2012; Zhang et al., 2012) rely on limited area models that can be based either on an off-line or 114 an on-line approach to take into account the effect of meteorological conditions on air 115 116 composition. Off-line chemistry models, known as Chemistry Transport Models (CTMs), use the meteorological parameters from the analyses or the forecasts provided by a separate 117 numerical weather prediction model. On-line models are meteorological models in which 118 chemical variables and processes are included (Baklanov et al., 2014). On-line models have 119 the capability to represent the feedback of the chemical composition on meteorological 120 121 parameters but they are computationally demanding by design. This is why CTMs are generally preferred for operational air quality forecasting systems. 122

123 The chemical composition of air depends on many processes that need to be well represented 124 in models in order to provide reliable air quality forecasts (e.g. Rao et al. 2011). The 125 composition near the surface is very much driven by emissions but also by chemical processes

(gaseous/heterogeneous reactions and photolysis) including the production of secondary 126 pollutants, by the advection by winds, by the diffusion in the planetary boundary layer, by the 127 scavenging by rain and by the dry deposition at the surface. Each of these processes has its 128 own uncertainty. These uncertainties come, on one hand, from the limit of our current 129 knowledge and, on the other hand, from the need to simplify the process representation in 130 models because of computational constraints. In meteorology and climate studies, and more 131 recently in atmospheric dispersion and chemistry modelling, the approach based on a multi-132 model ensemble of forecasts has been developed to provide better information by combining 133 134 information from different models. The methods vary from very simple such as the average or the median to more elaborated such as weighted averages based on past scores or Bayesian 135 136 models or spectral methods (e.g. Delle Monache 2006, Riccio et al. 2007, Potempski et al. 2010, Galmarini et al. 2013). 137

The European Union is very much involved in air quality issues not only through a series of 138 protocols on emissions and consecutive political actions but also by supporting research 139 activities aiming at developing tools for air quality monitoring for Europe. These activities 140 were initiated in the GEMS (Global and regional Earth-system (Atmosphere) Monitoring 141 using Satellite and in-situ data, FP6, 2005-2009, Hollingworth et al. 2008) and PROMOTE 142 (ESA PROtocol MOniToring for the GMES Service Element: Atmosphere, 2006-2009, 143 http://www.gse-promote.org/) projects and pursued in the MACC (Monitoring Atmospheric 144 Composition and Climate, FP7, 2009-2011), MACC-II (Monitoring Atmospheric 145 Composition and Climate: Interim Implementation, FP7, 2011-2014) and MACC-III 146 147 (Monitoring Atmospheric Composition and Climate-III, H2020, 2014-2015) projects. One of the major achievements accomplished in GMES, MACC and MACC-II for European AQ 148 objectives is the development and the exploitation of a pre-operational analysis and 149 forecasting system run on a daily basis. This system is based on the combined use of an 150 ensemble of 7 air quality models. The general objective of this system is not to provide air 151 quality forecasts and analyses for precise local situations but at the pan-European scale. For 152 this purpose, the horizontal resolution chosen for the individual models is between 10 and 20 153 km, thereby representing large scale phenomena and background air pollution. GEMS 154 involved 10 research and operational models. Evolving towards a pre-operational system, the 155 156 MACC/MACC-II/MACC-III ensemble is, since 2009, based on seven following state-of-theart regional CTMs that are all developed and run in Europe and that have been extensively 157 evaluated: CHIMERE (Menut et al., 2013a), EMEP (MSC-W version) (Simpson et al., 2012), 158

EURAD-IM (Haas et al., 1995; Memmesheimer et al., 2004), LOTOS-EUROS (Schaap et 159 al., 2008), MATCH (Robertson et al., 1999; Andersson et al., 2015), MOCAGE (Josse et al., 160 2004; Dufour et al., 2004) and SILAM (Sofiev et al., 2008). They are used to produce a multi-161 model ensemble for major monitored pollutants. Although each of these models can perform 162 very well on particular days in particular areas, the ensemble approach aims at providing on 163 average forecasts and analyses of better quality than any of the individual. It also gives an 164 indication of the uncertainties through the spread between the models. Similarly to 165 meteorological forecasts, the quality of the AQ forecasts needs to be routinely evaluated to 166 167 provide information to users on its reliability. The performance of the individual and ensemble forecast products is evaluated on a daily basis from comparisons with available 168 169 surface observations by the European AQ station network. Additionally the system has been providing birch pollen forecasts at surface during the pollen season since 2013. All the 170 171 forecast and analysis numerical data are publicly available.

The objectives of the paper are, firstly, to provide a description of the pre-operational analysis 172 173 and forecasting system in place within MACC and MACC-II to provide AQ services for the European continent and, secondly, to document and analyse the performance of the multi-174 model ensemble. Since the system continuously evolves with time, we present here the 175 configuration at the end of the MACC-II project (summer 2014) with a brief description of 176 recent upgrades included before the end of 2014. An overview of the analysis and forecasting 177 system, including the seven models and the Ensemble, is provided in Section 2. Section 3 is 178 devoted to the system performance for case studies and on a seasonal basis. Section 4 gives a 179 180 summary and the perspective on short and mid-term developments of the MACC-II system 181 and associated research.

182

2. Description of the analysis and forecasting systems

184 **2.1 General description of the system**

The MACC-II air quality system aims at providing analyses and forecasts of the main pollutants at the regional scale over the European continent: from 25° W to 45° E and from 30° N to 70° N. Each of the 7 models is run at its own horizontal and vertical resolutions, with the horizontal resolutions varying between ~20 km and ~10 km. This range of resolutions is not designed to reproduce local aspects of air pollution but to provide concentrations of pollutants at the regional scale that can then be used in particular as boundary conditions for AQ forecasts at finer resolution.

The range of the forecasts is 96h from 00 UTC on Day0 with hourly outputs on 8 vertical 192 levels (surface, 50m, 250m, 500m, 1000m, 2000m, 3000m and 5000m). Day0 is defined as 193 the day when the forecast is run. The forecast initial time/date is Day0 at 00UTC and final 194 time/date is Day3 at 24UTC. For each timestep (one hour), the individual model fields are 195 interpolated on these vertical levels and on the same regular 0.1° latitude by 0.1° longitude 196 grid over the MACC-II European domain. It is from these re-gridded fields that the ensemble 197 median and verification products are calculated. Before mid-May 2014, only surface, 500m, 198 1000m and 3000m were produced. The forecast species include O₃, NO₂, SO₂, CO, PM₁₀ and 199 PM_{2.5}, which are called core species hereafter. The core species are monitored in Near Real-200 Time (NRT) by European air quality stations and forecasts can therefore be evaluated 201 routinely against these observations. Forecasts of birch pollen concentrations at surface are 202 also produced during the pollen season (1st of March to 30th of June) since 2013. This product 203 is not discussed in this paper since its description and validation is detailed in Sofiev et al. 204 (2015). Additionally, since mid-May 2014, the production has been extended to other species 205 206 or aggregation of species (NO, NH₃, PAN+PAN precursors, total Non-Methane Volatile Organic Compounds). Additional species are provided primarily for the use as initial and/or 207 208 boundary conditions mainly for finer scale models designed for local AQ purposes.

The analysis at the surface for Day0-1 is run daily *a posteriori* on Day0 using the assimilation of the hourly data from the AQ monitoring stations available in Europe between 00UTC and 23UTC on Day0-1. Like for the forecasts, Day0 is defined as the day when the analysis is run. Day0-1 refers to the day before Day0. The analysis initial time/date is Day0-1 at 00UTC and final time/date is Day0-1 at 23UTC. Similarly to the forecasts, the hourly individual model fields are interpolated on the same 0.1° latitude by 0.1° longitude grid. The analyses are only produced at the surface level.

Table 1 gives the portfolio of the regional data products. All the additional species and 216 217 vertical levels are not yet available from all models but this is planned to be completed in 2015. Table 2 gives the current times of delivery of the ensemble numerical data products. 218 These production times have been shifted earlier since summer 2014 in order to fulfil users' 219 needs, in particular Day0 and Day1 forecasts, that are the mostly used products, are now 220 221 available at 07UTC. This has been made possible by an earlier delivery of the forecasts of each of the 7 models and by replacing the bulk 96h processing of the ensemble by a 222 processing 24 h segments. The delivery time of the analysis has also been shifted earlier in 223 June 2014. 224

The NRT hourly observations of O₃, NO₂, SO₂, CO, PM₁₀ and PM_{2.5} from the European AQ 225 monitoring stations are used for model assimilation to produce the daily analyses and also for 226 the forecast and analysis evaluation. From 2009 until recently, they were gathered country by 227 country through bilateral agreements with the project. Since 2014, a new system has been put 228 in place to gather these observations from the centralised Airbase database maintained by the 229 European Environment Agency (EEA). The database collects the NRT data and validated data 230 from the European countries bound under Decision 97/101/EC to engage in a reciprocal 231 exchange of information (EoI) on ambient air quality. The delivery time of the observations to 232 233 EEA takes place earlier and there is on average more data available than when gathering them bilaterally country by country, although there is a large variability from one day to another in 234 235 the number of data available. For the use in the production of the analyses, we chose after a dataflow monitoring of the EEA database a cut-off time at 07 UTC on Day0 for the dataset 236 237 covering Day0-1. At this time of the day, more than 90% (on average) of all data are available. The 07 UTC cut-off time is therefore a compromise between having enough data 238 239 available for the model assimilation and a reasonable production time for the ensemble analysis that was at 14:30UTC at the end of MACC-II. This production time is still too late 240 241 for the forecasts to be initialised from the analysis, meaning that the forecast and the analysis 242 products are currently run in two separate chains for each model. For the product evaluation, the observations covering Day0-1 available in the EEA database at 23UTC on Day0 are used 243 since there is less constraint on the time of delivery of evaluation products. On average there 244 is about 10% more data available at 23UTC than at 07 UTC. As shown in a MACC-II report 245 (D16 3; http://www.gmes-atmosphere.eu/documents/maccii/deliverables/obs/), the additional 246 data collected at 23UTC compared to 07UTC are mainly data from the end of the previous 247 day. This is because there is a significant number of stations that do not send their late 248 afternoon and evening Day0-1 data before 07UTC on Day0. This means that the 23UTC 249 250 dataset used for verification is homogeneous with approximately the same number of observations in the morning, afternoon and evening. 251

Because the NRT AQ observations used are not validated data, sorting procedures are applied to reject unrealistic observations through a blacklist. The blacklist includes stations identified as unrealistic, such as for instance stations giving the same concentration for each hour of the day. Moreover, only the data representative of the horizontal resolution of the regional models (10-20 km) are selected. There is currently no uniform and reliable metadata on site representativeness available for all regions and countries of Europe. This is why we chose to

follow the work that has been done by Joly and Peuch (2012) to build an objective 258 classification of sites, based on past validated measurements available in the Airbase database 259 (EEA). Stations are classified between 1 and 10 depending on the characteristics of their 260 series of measurements (diurnal cycle, "week-end effect" and high frequency variability with 261 periods lower than 3 days). The original classification of Joly and Peuch (2012) was based on 262 a series of data spanning from 2002 to 2009. It has been updated in MACC-II using version 7 263 of the Airbase database spanning from 2002 to 2011. Classes 1 to 10 cover the range from 264 most rural background sites to most locally polluted sites. Once each station is classified we 265 266 exclude those stations that have a concentration variability that is typical of locations mainly 267 influenced by local phenomena. Only the stations with class numbers ranging from 1 to 5 for 268 all pollutants are kept. The threshold of 5 allows us to remove the stations influenced by local phenomena while keeping a reasonable number of stations for calculating statistical 269 270 indicators. This leads to a typical number in summer 2014 of : ~600 sites for ozone, ~500 sites for NO₂, ~150 sites for SO₂, ~40 sites for CO, ~400 sites for PM₁₀, ~150 sites for PM_{2.5}. 271 272 All these data are used for the verification of the forecast products. For the verification of analyses, the developments done during MACC-II were only put into place after the end of 273 274 the project. This verification is done in the following way: a list of stations not used for the 275 assimilation is kept aside for each pollutant for verification. This list is the same every day and it has been determined so that the stations are well spread inside the domain. The ratio of 276 277 observations that are kept aside for the verification of analyses is roughly 20% of the total amount of observations that are downloaded at 23UTC. 278

The plots of forecasts and analyses from the 7 models and the Ensemble and of their scores against observations are available daily on <u>http://macc-raq.copernicus-atmosphere.eu/</u>. Numerical data are publicly available and can be accessed <u>http://www.gmes-</u> <u>atmosphere.eu/request_regional_data/</u>. The full set of numerical data as listed in Table 1 is made available as soon as produced on Météo-France FTP (File Transfer Protocol) server. A subset of these data can also be interactively accessed through the Deutsches Zentrum für Luft- und Raumfahrt (DLR) World Data Center.

Major sources of uncertainties in regional AQ forecast and analyses are the quality of the emissions used, the meteorological forcings, the representation of the atmospheric physical and chemical processes, the initial and boundary conditions for the chemical species and the uncertainties in observations and assimilation methods impacting the analysis. The approach chosen in MACC-II is to use the best available emissions over Europe, high quality

9

meteorological forecasts and chemical boundary conditions in all seven chemistry-transport models. Therefore the variability between the forecasts of the seven models used in the ensemble comes mainly from differences in the models in the treatment of the chemical processes (homogeneous and heterogeneous, photolysis), the advection, the convective transport, the turbulent mixing and the wet and dry depositions. Other differences stem from the use of different vertical and horizontal grids. For the production of the analysis, each model uses its own assimilation system.

298 The inventory used for anthropogenic emissions was built primarily for modelling purposes in the frame of the MACC-II project (Kuenen et al., 2014). This is an updated version of the 299 300 MACC inventory (Kuenen et al. 2011). Its resolution is 1/8° longitude x 1/16° latitude which is approximately 7 km x 7 km and it covers the UNECE-Europe for the years 2003 to 2009. 301 302 The 2009 inventory is currently used in the MACC-II daily production. An important upgrade of the MACC-II inventory compared to earlier MACC inventory is the provision of a 303 particulate matter split between Elemental Carbon, Organic Carbon, SO₄, Na and other 304 aerosols. More details on this inventory can be found in Kuenen et al. (2014). For the biogenic 305 306 sources, each model deals with its own emissions based on dynamical parameterizations 307 and/or inventories that are detailed in the following individual model description sub-sections. Additionally, emissions from fires are taken into account using the GFASv1.1 product (Kaiser 308 et al., 2012) available daily at 0.1°x0.1° resolution. GFASv1.1 is based on fire radiative power 309 retrievals from data of the Moderate Resolution Imaging Spectroradiometer (MODIS) 310 instruments aboard the Terra and Aqua satellites. The GFAS product for Day0-1 is available 311 312 around 06UTC on Day0. This is soon enough to be used in the daily analysis individual 313 production chains. At the time the individual forecasts begin for Day0, only the fire emissions from Day0-2 are available. To have less time gap between the fire emissions and the starting 314 time of the regional forecast runs (usually around 20 UTC), an additional fire emission 315 product available around 20:30UTC on Day0-1 using satellite observations from 15 UTC on 316 Day0-2 to 15 UTC on Day0-1 is currently under test. In the forecasts, a persistence of the fire 317 emissions of 3 days is assumed. This is a rounded average of the fire duration obtained by 318 Turquety et al. (2014) from Euro-Mediterranean region from MODIS MCD64 product (Giglio 319 et al., 2010) in the period 2003-2012. 320

The meteorological fields used to force the 7 CTMs are from the operational IFS (Integrated Forecasting System) daily meteorological forecasts of the European Centre for Medium-Range Weather Forecasts (ECMWF). The IFS forecast starting at 12UTC on Day0-1 is used for the MACC-II air quality 96h forecast starting at 00UTC on Day0. For the analysis on
Day0-1, the IFS forecast starting at 00UTC on Day0-1 is used.

326 The regional domain boundary conditions for the aerosols and gaseous species are provided 327 by the MACC-II global assimilation and forecasting system. This forecasting system is an extension of the ECMWF meteorological IFS running at lower resolution, providing 328 329 concentrations of dust, sea salt, organic matter, black carbon and sulphate aerosols (Morcrette et al. 2009, Benedetti et al. 2009) that are used to force the aerosols in the regional CTMs at 330 331 the boundaries. At the end of MACC-II project (summer 2014), for the chemical species the IFS was two-way coupled to the off-line MOZART global chemical transport model (CTM). 332 333 This allowed assimilation of satellite data for O₃, NO₂, and CO in the IFS itself, while the detailed chemical processes were handled in the MOZART model (Flemming et al. 2009, 334 335 Stein et al. 2012, Inness et al. 2013). Since 18 September 2014, the MACC-II global assimilation and forecasting system has been upgraded to a fully integrated system for 336 aerosols and chemical species. Instead of the coupling with the MOZART model, the 337 chemistry is now treated on-line in the IFS using chemistry modules based on the TM5 model 338 (Huijnen et al., 2010). This new system is named Composition-IFS (C-IFS) and is further 339 described in Flemming et al. (2014). The chemical mechanism in the TM5 operational version 340 of C-IFS is based on a modified version of the Carbon Bond 5 (CB05) scheme (Williams et 341 al., 2013, Yarwood et al., 2005). 342

343 Based on all the inputs described above, each of the centres in charge of the 7 models runs its production locally and transfers its forecast and analysis files to Météo-France (referred to 344 345 central production centre hereafter). The general organisation of the MACC-II air quality forecasts and analysis system is summarized in Fig. 1. Tables 3 and 4 give the general 346 features of the seven individual models and of their analysis system. A short description of the 347 seven individual models and of the ensemble is given in the following sections. More details 348 349 can be found in the MACC-II 6-monthly reports (http://www.gmes-350 atmosphere.eu/documents/maccii/deliverables/ens/).

351

352 2.2 CHIMERE forecast and analysis system

353 CHIMERE is an Eulerian chemistry-transport model able to simulate concentrations fields of 354 gaseous and aerosols species at a regional scale (Menut et al., 2013a). The model is developed 355 under the GPL licence (<u>http://www.lmd.polytechnique.fr/chimere/</u>). CHIMERE is used for analysis of pollution events, process studies, (Bessagnet et al., 2009; Beekmann and Vautard,
2010), experimental and operational forecast (Rouil et al., 2009), regional climate studies and
trends, (Colette et al., 2011), among others.

359 CHIMERE calculates and provides the atmospheric concentrations of tens of gas-phase and aerosol species over local (e.g. urban) to continental domains (from 1 km to 1 degree 360 361 resolution). Vertically, the model is able to simulate the whole troposphere. The gaseous species are calculated using the MELCHIOR2 scheme and the aerosols using the scheme 362 363 developed by Bessagnet et al. (2004). This module takes into account species such as 364 sulphate, nitrate, ammonium, primary organic matter (POM) and elemental carbon (EC), 365 secondary organic aerosols, sea salt, dust and water. These aerosols are represented using 8 bins, from 40 nm to 40 um, in diameter. The life cycle of the aerosols is completely 366 367 represented with nucleation of sulphuric acid, coagulation, adsorption/desorption, wet and dry deposition and scavenging. This scavenging is both represented by coagulation with cloud 368 droplets and precipitation. The formation of SOA is also taken into account (Bessagnet et al., 369 370 2009).

Biogenic emissions are calculated using the MEGAN emissions scheme (Guenther et al.,
2006) which provides fluxes of isoprene and monoterpenes. The mineral dust emissions are
calculated using the (Alfaro and Gomes, 2001) scheme, forced by satellite soil and surface
data (Menut et al., 2013b).

The CHIMERE assimilation system for operational products is based upon hourly optimal interpolation processing of surface observations for O_3 and PM_{10} (Honoré et al, 2008). During MACC-II, an Ensemble Kalman filter has been also developed for ozone analysis (Gaubert et al., 2014).

CHIMERE is fully dedicated to regional air pollution modelling. It includes a comprehensive representation of the aerosol with secondary organic (SOA) and inorganic aerosols (SIA). CHIMERE has a chemical scheme specifically designed to reproduce the photochemical activity in the lower part of the troposphere (for air quality purposes). In terms of point that may need to be improved, the vertical resolution is composed of 8 levels up to 500 hPa, meaning that the models need to be fed with realistic top conditions. The assimilation is so far limited to O_3 and PM_{10} and for the surface layer.

386

387 2.3 EMEP forecast and analysis system

The EMEP/MSC-W model (hereafter referred to as 'EMEP model') has been developed at the EMEP Meteorological Synthesizing Centre-West at the Norwegian Meteorological Institute. The model has been publicly available as open source code since 2008, and a detailed description is given in Simpson et al. (2012).

The numerical solution of advection is based on Bott (1989). The turbulent diffusion 392 393 coefficients are calculated for the whole 3D model domain on the basis of local Richardson 394 number, and the planetary boundary layer (PBL) height is calculated using methods described 395 in Simpson et al. (2003). Dry deposition uses a resistance analogy combined with stomatal and non-stomatal conductance algorithms (Simpson et al., 2003; Tuovinen et al., 2004), 396 397 whereas wet deposition uses scavenging coefficients applied to the 3-D rainfall, including both in-cloud and sub-cloud scavenging of gases and particles. The chemical scheme couples 398 399 the sulphur and nitrogen chemistry to the photochemistry using about 140 reactions between 70 species (Andersson-Sköld and Simpson, 1999; Simpson et al. 2012). 400

The methodology for biogenic emissions builds on maps of 115 forest species generated by Köble and Seufert (2001). Emission factors for each forest species and for other land classes are based on Simpson et al. (1999), updated with recent literature (see Simpson et al. (2012) and references therein), and driven by hourly temperature and light using algorithms from Guenther et al. (1995). Other natural emissions include marine emissions of dimethyl sulphide, and SO₂ from volcanoes.

407 The standard model version distinguishes two size fractions for aerosols, fine aerosol $(PM_{2.5})$ and coarse aerosol (PM_{10} excluding $PM_{2.5}$). The aerosol components presently accounted for 408 are sulphate, nitrate, ammonium, anthropogenic primary particulate matter, sea salt and desert 409 410 dust. Also aerosol water is calculated. The parameterisation of dry deposition for aerosols follows standard resistance formulations, accounting for diffusion, impaction, interception, 411 and sedimentation. Wet scavenging is treated with simple scavenging ratios, taking into 412 account in-cloud and sub-cloud processes. For secondary organic aerosol (SOA) the so-called 413 'EmChem09soa' scheme is used, which is a slightly simplified version of the mechanism 414 described by Bergström et al. (2012). 415

The EMEP data assimilation system (EMEP-DAS) is based on the 3D-Var implementation for the MATCH model (Kahnert, 2008, 2009). The background error covariance matrix is estimated following the NMC method (Parrish and Derber, 1992). Currently, the EMEP-DAS delivers analyses for NO₂, using NO₂ columns of OMI and in situ measurements of NO₂ surface concentrations. The assimilation window is 6 hours, 4 times per day.

The EMEP model performs well especially for particulate matter, as it includes carefully 421 evaluated representations of both primary and secondary organic aerosols, in addition to 422 inorganic aerosols, elemental carbon, sea salt, mineral dust and water. Another strength is that 423 424 its domain extends throughout the whole troposphere, thus taking accurate account of longrange transport of pollutants in the free troposphere. As the EMEP model is designed mainly 425 for background concentrations, urban increments have not been implemented as in some other 426 models with equally coarse resolution, leading to somewhat lower performance in urban and 427 sub-urban areas. However, being one of the main research tools under the UN LRTAP 428 429 convention, the EMEP model is evaluated continuously against measurements of a large range of chemical parameters (including air concentrations, depositions, and trends) ensuring 430 431 modelling capability with very good overall performance (e.g. Jonson et al., 2006; Fagerli and Aas, 2008; Genberg et al., 2013). A weakness of the analysis chain until the end of 2014 was 432 433 that only NO₂ was assimilated. However, since early 2015 ozone has been assimilated.

434

435 2.4 EURAD-IM forecast and analysis system

EURAD-IM is an Eulerian meso-scale chemistry transport model involving advection,
diffusion, chemical transformation, wet and dry deposition and sedimentation of tropospheric
trace gases and aerosols (Hass et al., 1995, Memmesheimer et al., 2004). It includes 3d-var
and 4d-var chemical data assimilation (Elbern et al., 2007) and is able to run in nesting mode.
EURAD-IM has been applied on several recent air pollution studies (Monteiro et al., 2013;
Zyryanov et al., 2012; Monteiro et al, 2012; Elbern et al., 2011; Kanakidou et al., 2011).

The positive definite advection scheme of Bott (1989) is used to solve the advective transport. An Eddy diffusion approach is used to parameterize the vertical sub-grid-scale turbulent transport. The calculation of vertical Eddy diffusion coefficients is based on the specific turbulent structure in the individual regimes of the planetary boundary layer (PBL) according to the PBL height and the Monin-Obukhov length (Holtslag and Nieuwstadt, 1986). A semiimplicit (Crank-Nicholson) scheme is used to solve the diffusion equation.

Gas phase chemistry is represented by the Regional Atmospheric Chemistry Mechanism (RACM) (Stockwell et al., 1997) and an extension based on the Mainz Isoprene Mechanism (MIM) (Geiger et al., 2003). A two-step Rosenbrock method is used to solve the set of stiff ordinary differentials equations (Sandu et al., 2003; Sandu and Sander, 2006). Photolysis frequencies are derived using the FTUV model according to Tie et al. (2003). The radiative

transfer model therein is based on the Tropospheric Ultraviolet-Visible Model (TUV) 453 developed by Madronich and Weller (1990). The modal aerosol dynamics model MADE 454 (Ackermann et al., 1998) is used to provide information on the aerosol size distribution and 455 chemical composition. To solve for the concentrations of the secondary inorganic aerosol 456 components, a FEOM (fully equivalent operational model) version, using the HDMR (high 457 dimensional model representation) technique (Rabitz et al., 1999; Nieradzik, 2005), of an 458 accurate mole fraction based thermodynamic model (Friese and Ebel, 2010) is used. The 459 updated SORGAM module (Li et al., 2013) simulates secondary organic aerosol formation. 460 461 Biogenic emissions are calculated in the EURAD-IM CTM with the Model of Emissions of 462 Gases and Aerosols from Nature (MEGAN) (Guenther et al., 2012).

The gas phase dry deposition modelling follows the method proposed by Zhang et al. (2003). Dry deposition of aerosol species is treated size dependent using the resistance model of Petroff and Zhang (2010). Wet deposition of gases and aerosols is derived from the cloud model in the EPA Models-3 Community Multiscale Air Quality (CMAQ) modelling system (Roselle and Binkowski, 1999).

The EURAD-IM assimilation system includes (i) the EURAD-IM CTM and its adjoint, (ii) 468 the formulation of both background error covariance matrices for the initial states and the 469 470 emission factors, (iii) the observational basis and its related error covariance matrix, and, (iv) the minimisation including the transformation for preconditioning. The quasi-Newton limited 471 memory L-BFGS algorithm described in Nocedal (1980) and Liu and Nocedal (1989) is 472 applied for the minimization. Following Weaver and Courtier (2001) with the promise of a 473 474 high flexibility in designing anisotropic and heterogeneous influence radii, a diffusion approach for providing the background error covariance matrices is implemented. 475

One of the EURAD-IM strength is that it includes a comprehensive treatment of aerosol 476 477 dynamics and chemistry. Parameterisations of the formation of secondary particles are temperature dependent for both the inorganic and organic components. However, the 478 complexity of the aerosol components of EURAD-IM is as yet not supported by sufficiently 479 known emission rates of particle types, as well as for organic gaseous precursor compounds, 480 especially from biogenic sources. Another strength of the EURAD-IM system is its ability to 481 assimilate chemical data from a wide range of instruments ranging from surface or airborne in 482 situ data to retrievals from several satellites, which are then defining the initial values. 483

484

485 2.5 LOTOS-EUROS forecast and analysis system

The 3D chemistry-transport model LOTOS-EUROS (Schaap et al., 2008) is developed by the 486 487 Dutch institutes TNO (www.tno.nl), RIVM (www.rivm.nl) and, more recently, KNMI 488 (www.knmi.nl). It is used for regional-scale air-quality forecasts in Europe and the Netherlands (de Ruyter de Wildt et al., 2011). The LOTOS-EUROS model has participated in 489 490 several international model intercomparison studies addressing ozone (Van Loon et al., 2007; Solazzo et al., 2012a) and particulate matter (Cuvelier et al., 2007; Vautard et al., 2007; Stern 491 492 et al., 2008; Solazzo et al., 2012b). These studies have shown that the model has a performance comparable to other European regional models. In the past three year, three 493 494 major updates of the LOTOS-EUROS model have been implemented, moving from version 1.7 to version 1.10. Detailed update information can be found on the model web 495 496 page, http://www.lotos-euros.nl. Since the end of MACC-II, the latest update to v1.10 implemented operationally consists of changes in the SO₂ to SO₄ conversion rate, use of 497 AQMEII conventions for the fine/coarse dust assignment, update of resistances for e.g. ozone 498 499 (leading to an overall ozone increase), and improvement of the treatment of fire emissions.

The model extends up to 3.5 km above sea level, with three dynamic layers and a fixed 25m 500 thick surface layer. The lowest dynamic layer is the mixing layer, followed by two reservoir 501 502 layers. The height of the mixing layer is obtained from the ECMWF meteorological input data used to drive the model. Transport is based on the monotonic advection scheme developed by 503 504 Walcek (2000). Gas phase chemistry is described using the TNO CBM-IV scheme (Schaap et al., 2008). Hydrolysis of N₂O₅ is described following Schaap et al. (2004). Aerosol chemistry 505 506 is represented using ISORROPIA-2 (Fountoukis and Nenes, 2007). The aerodynamic 507 resistance is calculated for all land use types separately. Below cloud scavenging is described 508 using simple scavenging coefficients for gases (Schaap et al., 2004) and particles (Simpson et al., 2003). Dry deposition is based on the well-known resistance approach, with the DEPAC 509 510 parameterization for gases (Wichink Kruit et al., 2012) and the Zhang et al. (2001) 511 parameterization for particles.

512 Biogenic isoprene emissions are calculated following the mathematical description of the 513 temperature and light dependence of the isoprene emissions, proposed by Guenther et al. 514 (1993), using the actual meteorological data. For land use the CORINE/Smiatek database has 515 been enhanced using the tree species map for Europe made by Koeble and Seufert (2001). 516 Total PM10 in the LOTOS-EUROS model is composed of chemically unspecified primarily 517 PM in the fine and coarse mode, black carbon, dust, ammonium, sulphate, nitrate and sea salt518 (Na in the fine and coarse mode).

The LOTOS-EUROS model is equipped with a data assimilation package with the ensemble Kalman filter technique (Barbu et al., 2008; Timmermans et al., 2009; Curier et al., 2012). Data assimilation for the MACC-II daily analyses is performed with surface ozone observations (Curier et al., 2012). An extension to other surface and satellite data is foreseen in the near future.

The LOTOS-EUROS model has been designed as a model of intermediate complexity, to 524 favor short computation times. For this, the vertical top of the operational model version is 525 limited and covers only the boundary layer and reservoir layers (up to 3.5 km); effectively, the 526 model therefore employs 4 dynamic layers only. Concentrations from the free-troposphere are 527 taken from the global boundary conditions, and therefore fully incorporate the knowledge, 528 assimilations, and validation efforts present in the global model. A major weakness is that 529 secondary organic aerosols are currently not included; instead, a bias correction for total PM 530 is used to account for the missing aerosols. In spite of the limited complexity, the model 531 performs well in simulation of O₃ (Curier et al., 2012) and has a skill to forecast the observed 532 variability in PM₁₀ (de Ruyter de Wildt et al., 2011). Apart from the relative short run-533 through time, the strength of the model is in the detailed description of anthropogenic 534 emissions, since close cooperation with the developers of the TNO-MACC emission 535 inventory; this is for example shown by excellent simulation of boundary layer NO₂ 536 (Vlemmix et al., 2015). 537

538

539 2.6 MATCH forecast and analysis system

The MATCH model (Multi-scale Atmospheric Transport and Chemistry model) has been
developed at SMHI over the past 20 years and is applied for emergency purposes as well as
for regional scale chemistry modelling (Langner et al., 1998; Robertson et al., 1999).

The transport is described by a Bott-like mass conservative scheme (Bott, 1989; Robertson et al., 1999). For the vertical diffusion an implicit mass conservative scheme is used where the turbulent exchange coefficients for neutral and stable conditions are parameterized following Holtslag and Moeng (1991). In the convective case the turbulent Courant number is directly determined from the turn-over time in the atmospheric boundary layer. The dynamical core of the model contains initialization and adjustment of the horizontal wind components based on a procedure proposed by Heimann and Keeling (1989). This is important to ensure mass conservative transport for interpolated input weather data, specifically for the transport scheme used.

Boundary layer parameterization is determined from surface heat and water vapour fluxes as 552 553 described by van Ulden and Holtslag (1985) for land surfaces, and Burridge and Gadd (1977) 554 for sea surfaces. The boundary layer height is calculated from formulations proposed by 555 Zilitinkevich and Moronov (1996) for the neutral and stable case and from Holtslag et al. (1995) for the convective case. These parameterizations drive the formulations for vertical 556 557 diffusion and dry deposition where for the latter a resistance approach is used (Andersson et al. 2007). In-cloud and sub-cloud wet deposition is implemented following Andersson et al. 558 559 (2007). The photochemistry scheme is to large extent based on the EMEP chemistry scheme (Simpson et al., 1993), with some updates where a modified production scheme for isoprene 560 561 is the most notable based on the so-called Carter-1 mechanism (Carter, 1996; Langner et al., 562 1998).

Aerosols are described for 4 bins and only for secondary inorganic aerosols, dust and primary organic compounds at the moment. Inclusion of SOA is under testing. Sea salt emissions are dynamically described following Foltescu et al. (2005). A module for wind driven dust emissions is under testing that follows Schaap et al. (2005).

567 A 3D variational data assimilation scheme is used with spectral transformation (Kahnert, 2008). The limitation then is that background covariance structures are described as isotropic 568 and homogeneous, however not necessarily the same for different wavenumbers, and derived 569 570 from the so-called NMC-method (Parish and Derber, 1992). The advantage though is that the background error matrix becomes block diagonal and there are no scale separations as the 571 covariance between spectral components are explicitly handled. The block diagonal elements 572 are the covariance between wave components at different model layers and chemical 573 compounds. 574

The strength of the MATCH model is that it spans vertically the troposphere and makes use of the same vertical layers as provided from the IFS model up to 300 hPa. This means about 50 layers in the vertical and the lowest one just 20 m thick and about 15 in the boundary layer. Using the same vertical resolution as the IFS model is an advantage because no vertical interpolation is required. Nevertheless, since the MATCH model has been developed mainly using HIRLAM data with a coarser vertical resolution, the use of the high resolution vertical

- 581 levels from IFS may lead to less accurate chemistry forecasts compared to HIRLAM version.
- 582 A weakness is missing SOA and wind-blown dust in the PM description.
- 583

584 2.7 MOCAGE forecast and analysis system

The MOCAGE (Model Of atmospheric Chemistry At larGE scale) model (Josse et al. 2004, Dufour et al. 2004) has been developed at Météo-France since 2000. Its assimilation system has been developed jointly with CERFACS. This model and its assimilation system have been successfully used for tropospheric and stratospheric research (e.g. Bousserez et al. 2007; Barré et al. 2013, 2014; Lacressonnière et al. 2014) and also for operational purposes (Rouïl et al. 2009).

MOCAGE uses the semi-lagrangian advection scheme from Williamson and Rasch (1989) for 591 592 the grid-scale transport, the parameterization of convective transport from Bechtold et al. (2001) and the turbulent diffusion parameterization from Louis (1979). Dry deposition is 593 594 based on the approach proposed by Wesely (1989). The wet deposition by the convective and stratiform precipitations follows Mari et al. (2000) and Giorgi and Chameides (1986). 595 596 MOCAGE includes the RACM scheme for tropospheric chemistry (Stockwell et al., 1997) and the REPROBUS scheme for stratospheric chemistry (Lefèvre et al., 1994). Biogenic 597 598 emissions in MOCAGE are fixed monthly biogenic emission from Guenther et al. (1995).

The aerosol module of MOCAGE follows a bin approach and includes so far the primary aerosols: dust (Martet et al. 2009), sea salts, black carbon (Nho-Kim et al. 2005) and organic carbon. Recent updates of the primary aerosol module and corresponding evaluation can be found in Sič et al. (2014).

603 MACC-II operations use a variational assimilation system based upon MOCAGE and the PALM coupler, which has been developed during the ASSET European project (Geer et al., 604 605 2006; Lahoz et al., 2007). The system, recently renamed VALENTINA, has been used to 606 compute global and regional re-analyses of atmospheric composition in multiple studies (El 607 Amraoui et al., 2008; Massart et al., 2009; Barré et al. 2013, 2014; Emili et al., 2014). The assimilation algorithm employed for MACC-II analyses is a 3D-Var with assimilation 608 609 windows of one hour length (Jaumouillé et al., 2012), corresponding to the frequency of surface measurements. The assimilation has first been set for surface ozone analyses and in 610 MACC-III it has been extended to surface NO₂. The specification of the background and 611 observation errors is done based on the evaluation of historical time-series of observations and 612

model values. The horizontal error correlation has a Gaussian shape and its typical length is set to 0.4 degrees for ozone and 0.1 degrees for NO_2 , to account for the larger variability of NO_2 at fine spatial scales. The vertical error correlation length is set to 1 model grid point for all species (~100 m). As a consequence, assimilation increments linked to surface observations are confined in the planetary boundary layer.

618 The strength of MOCAGE is that it simulates the air composition of the whole troposphere and lower stratosphere. Thus it provides a full representation of transport processes, in 619 620 particular boundary layer-troposphere and troposphere-stratosphere exchanges, and the time 621 evolution of stratospheric conditions for accurate photolysis rate calculations at the surface. 622 The MOCAGE assimilation system in its MACC configuration produces robust analyses for both O₃ and NO₂ as illustrated in the annual reanalysis reports (http://www.gmes-623 624 atmosphere.eu/documents/maccii/deliverables/eva/). At the end of the MACC-II project, the main weakness of MOCAGE was the lack of secondary aerosols. Inorganic secondary 625 aerosols have been developed recently and will be included in the next MACC operational 626 627 version (Guth et al., 2015). This new feature is also used in the current development of PM_{10} 628 assimilation.

629

630 2.8 SILAM forecast and analysis system

SILAM is a meso-to-global scale dispersion model (Sofiev et al, 2008), see also the review Kukkonen et al., (2012), http://silam.fmi.fi) that is used for atmospheric composition, emergency, composition-climate interactions, and air quality modelling purposes. The model has been applied with resolutions ranging from 1km up to 3 degrees, incorporates 8 chemical and physical transformation modules and covers the troposphere and the stratosphere. The model is publicly available since 2007 and is used as operational and research tool.

The model has two dynamic cores: Lagrangian (Sofiev et al., 2006), primarily used in 637 emergency-type applications, and Eulerian (Galperin, 2000; Sofiev, 2002) used in 638 atmospheric composition, climate, and air quality-related applications, including MACC-II. 639 The MACC-II operational SILAM v.5.2 uses the simple dry deposition scheme of (Sofiev, 640 2000) for gases and a new approach for aerosols (Kouznetsov and Sofiev, 2012), which 641 covers particle sizes from 1 nm up to ~ 50 μ m of effective aerodynamic size. The wet 642 deposition scheme used in MACC-II simulations calculates the 3-D removal coefficient and 643 distinguishes between sub- and in-cloud scavenging, large-scale and convective 644

precipitations, as well as between rain and snow (Sofiev et al., 2006). Boundary layer parameterization follows (Sofiev et al., 2010), whereas in the free troposphere and the stratosphere turbulence is computed following IFS approach and corresponding turbulent length scale.

Two chemical schemes are used: the CBM-4 gas-phase chemistry mechanism and own development for heterogeneous chemical transformations and inorganic aerosol formation after (Sofiev, 2000). Aerosols in SILAM are represented via sectional approach with speciesspecific size spectra. The aerosol species include primary anthropogenic aerosols, divided into $PM_{2.5}$ and PM_{10} , secondary inorganic aerosols (sulphates, nitrates and ammonia), and sea salt aerosols.

The forecasts utilise the BVOC emission term based on NatAir project results (Poupkou et al.,
2010) and own development for the sea-salt emission (Sofiev et al., 2011).

The data assimilation system of SILAM consists of 3D-VAR and 4D-VAR modules (Vira and Sofiev, 2012). The MACC-II near-real time analysis suite uses the 3D-VAR method and assimilates hourly surface observations of NO₂, O₃ and SO₂. PM observations have been assimilated in reanalysis simulations (Vira and Sofiev, 2015). The 4D-VAR methodology is utilised in re-analysis mode for pollen.

The model evolution from the MACC-II v.5.2 towards v.5.4 that will become operational in early-2015, include several important updates. The dry deposition scheme will follow the resistance analogy with extensions after (Simpson et al., 2003). Wind-blown dust will be included via lateral boundary conditions in the next release of operational SILAM v.5.4, together with a secondary organic aerosol module and fire emission.

A strong point of SILAM is the extensive treatment of secondary inorganic aerosol formation, 667 668 which is reproduced quite well, according to several evaluation exercises and model intercomparisons. Together with the detailed deposition scheme, this leads to good scores for 669 PM_{2.5}, especially in winter when inorganic aerosols are dominant. The current limitation of 670 the model is the secondary organic aerosols formation that makes use of the volatility-based 671 model but it is not yet incorporated in the operational simulations, being tested in research 672 projects. A workaround of this limitation is included in the data assimilation modules, which 673 allow assimilation of both in-situ and remote-sensing measurements of gaseous and 674 particulate species. The module now allows for the PM and AOD observations being 675

assimilated into an unspecified particulate matter, which is then treated as inert aerosol, thuscompensating for the lack of secondary organic particles.

678

679 **2.9 ENSEMBLE forecast and analysis system**

To process the ensemble, all seven individual models are first interpolated to a common 680 0.1°x0.1° horizontal grid. For each grid point, the ensemble model (referred as the 681 ENSEMBLE hereafter) value is calculated as the median value of the individual model 682 forecasts or analyses available. The median is defined as the value having 50% of individual 683 684 models with higher values and 50% with lower values. This method is rather insensitive to outliers in the forecasts or analyses and is very efficient computationally. These properties are 685 useful from an operational point of view. The method is also little sensitive if a particular 686 model forecast or analyses is occasionally missing. The performances of the ensemble median 687 are discussed in section 3. For the forecasts, the ensemble is produced for all levels and all 688 species (core and additional). For the analyses, the individual assimilation systems provide 689 only analyses at the surface level and do not produce analyses for all species yet. At the end of 690 MACC-II, ozone was the only species that was produced by 6 of the models. For other 691 692 species, analyses from less than 5 models were available. This is why the ensemble analysis in MACC-II was only calculated for ozone. It has been extended to NO₂ in 2015 since more 693 694 models will produce NO₂ analyses.

695

696 **3. Evaluation of the performances of the system**

697 3.1 General description

The evaluation of the performances of a forecast system is a necessary step rating its quality 698 and thus proving its usefulness. The MACC-II air quality forecasts are evaluated against the 699 700 NRT AQ surface monitoring data detailed in section 2.1. Note that this set of data is fully 701 independent of the forecast since the analyses assimilating the NRT AQ data are produced too 702 late to be used to initialise the forecasts. The tools to assess the performances of the analyses 703 are not yet in place but this is planned to be ready in 2015. Since the focus of the MACC-II regional system over Europe is on air quality, meaning air composition close to the surface, 704 705 no column observations (ground based or from satellite) or upper air in situ measurements (i.e. on board aircraft) are used operationally to evaluate the system performances. 706

The forecast performances are measured using the five statistical indicators detailed in the Appendix: the mean bias (MB), the root mean square error (RMSE), the modified normalized mean bias (MNMB), the fractional gross error (FGE) and the correlation. These statistical measures, when taken together, provide a valuable indication of the model performances. Taylor diagrams are also used to combine root mean square errors and correlations.

The performances of the MACC-II regional AQ forecasts are assessed operationally byseveral means:

- on a daily basis with plots of statistical indicators and charts available on the MACC II regional website (<u>http://macc-raq.gmes-atmosphere.eu/</u>),
- on a 6 monthly basis in reports including plots of statistical indicators over two
 periods of 3 months (winter+spring or summer+autumn) and analysis of these
 indicators (http://www.gmes-atmosphere.eu/documents/maccii/deliverables/ens/).

Additionally, on a 6 monthly basis, reports are especially dedicated to the scientific analysis of the forecasts of the 7 models and of the ensemble in the Mediterranean area (<u>http://www.gmes-atmosphere.eu/documents/maccii/deliverables/ens/</u>). The Mediterranean area is recognized as challenging for models, in particular under summer conditions with very active photochemistry and because of its large variety of emission sources.

The performances of the NRT analysis are not presented in this paper since there is only an ensemble production of one species (ozone) and the daily verification procedure against an independent dataset was not yet in place at the end of MACC-II project.

727

728 3.2 Availability statistics

The MACC-II regional air quality forecasting and analysis system is currently under a pre-729 730 operational status that can be seen as the demonstrator of a future operational system. Correct working of 7 model chains and of the Ensemble chain is monitored on working hours only 731 732 since, at this stage, there was no funding yet for a 7day/7day 24h/24h control. Nevertheless, in 733 its pre-operational configuration the production chains are reliable with availability in time (see Table 2) of the 7 individual forecasts and analysis generally above 85% during MACC-II. 734 During the past year, the production suffered from failures because of the many changes that 735 were applied to the individual and central systems to fit with fully operational standards (data 736 format, file transfer, databases, processing softwares, ...). The operationalisation being nearly 737 fully settled, the reliability has been improved since the end of MACC-II (generally above 738

90%). The Ensemble forecast and analysis productions have been available 100% of the time
since September 2012. This high performance was achieved because the ENSEMBLE can be
produced even if all the 7 models are not available.

742

3.3 Example of the forecast of two ozone episodes between the 10th and the 13th of June 2014

In this section, we illustrate the performances of the MACC-II AQ forecasts for a case study of ozone pollution events that took place between the 10th and the 13th June 2014. A more indepth analysis of the individual model and of the ENSEMBLE performances is done over longer time periods in Section 3.4.

During the case study period, there were two regional areas with high ozone concentrations (> 749 120 mg m⁻³) occurring at the same time, one over Austria and surrounding regions (South of 750 Germany and Hungary), and one over the South East of France and the North of Italy. This is 751 illustrated by Fig. 2 displaying the maps of the 15h forecasts for the 10th of June at 15UTC of 752 ozone at surface from the ENSEMBLE together with the available observations. Note that 753 754 unfortunately no observation in Italy was available during the time period considered. Even if the comparison is limited by the missing observations, Fig. 2 shows that the ensemble median 755 756 captures the two ozone episodes.

757 For illustration of the system performance, the surface station measurements are compared in Fig. 3 to the forecast. We plot the model forecasts using EPSgrams that give a graphical 758 759 representation of the spread of the 7 models and therefore an estimate of the uncertainty over the 4 days of the forecast. Operationally, EPSgrams are built daily for 40 major cities in 760 761 Europe and made available on the MACC regional web site. Here EPSgrams are calculated 762 and plotted for the same locations as the measurements (Fig. 3) from the forecast started on the 10th of June at 00UTC. Note that, in Fig. 3, EPSgrams are 3-hourly while observations are 763 1-hourly. In the observations (left panel), the "Austrian" episode is highest on the 10th of June 764 in Fechenheim (Germany) and on the 11th of June in Hallein (Austria) and Sopron (Hungary) 765 with values reaching 200 μ g m⁻³ at Sopron. The "French" episode peaks at 250 μ g m⁻³ at 766 Sausset (France) with daily maxima over 150 μ g m⁻³ from the 10th to the 13th of June for all 767 three stations. 768

For the ozone peak event around Austria, there is generally a good consistency of the day-to-day trend provided by the seven models compared to the AQ station observations. For Sopron

station there is a main peak in the model on the 11th of June as in the observations. For 771 Fechenheim, the forecast gives highest concentrations on the 10th of June as measured, and 772 also reproduces the anomaly recorded in the observations in the morning of the 11th of June. 773 For this "Austrian" ozone episode, the spread between the 7 models is very reasonable 774 (generally less than 30 μ g m⁻³ for the 25%-75% range), showing the good consistency 775 between the models with slightly more spread between the forecasts for the highest peak 776 times. There is an exception in Fechenheim on the 10th of June where the 7 models exhibit a 777 large spread. This can be explained by the effect of complex topography combined with 778 779 specific meteorological conditions that lead to different behaviours of the models which have different horizontal grids and orography. 780

Even if the day-to-day trend is well reproduced by the models, ozone median values are often lower during daytime peaks than the observations by 30 to 50 μ g m⁻³ but the maxima of the 7 models are nevertheless close to those observed. There are also cases when the ensemble median forecasts higher peaks than measured as in Fechenheim on the 10th of June. This can also be seen in the map in Fig. 2 where some observations are lower than the ensemble median. During nighttime, the ozone median is close to the observed values.

787 For the ozone event in the south of France, the comparison shows also a good consistency between the diurnal variations of the models compared to observations. Most of the 7 models 788 are over 120 µg m⁻³ for each of the 4 days forecasted. Nevertheless, at Sausset the very high 789 ozone peak measured on the 10^{th} of June (over 240 µg m⁻³) is underestimated in all 7 790 forecasts. The very small spread of the models indicates a possible error in the meteorological 791 792 forecast for this day and/or in the emissions. Sausset is located on the Mediterranean coast, west to Marseille industrial city. On this particular day IFS forecasts an eastern wind with 793 high NOx from Marseille limiting the daytime production of ozone by the models compared 794 to observations. For Plan d'Aups, there is a very large spread of the models, particularly on 795 the 10th of June. This can be explained by the effect of sea and land breezes on this date 796 797 combined with steep orography and the presence in the vicinity of a pollution plume with NOx titrated ozone that leads to model differences. In this case (Plan d'Aups), models do not 798 799 reproduce observations. While for the other stations, the behaviour of the ensemble is good. For St Rémi, the models perform well with a small spread and diurnal variations close to the 800 801 measurements. In particular, the increase of the night minimum concentration from day to day is well forecasted. Similarly to the Austrian area, the ozone median concentrations are more 802 often lower than observed, but not always, as shown at Sausset on the 13th of June and St 803

Rémi on the 10th of June. In the case of underestimation of the ensemble median, the maximum of the individual models is generally close to those measured at the French stations. Note that for both the "Austrian" and "French" ozone episodes, there is no significant degradation of the forecast skills at Day3 and Day4 indicating that uncertainty in ozone forecasts is more driven by inherent uncertainty in chemistry transport models and part of its input than by uncertainty of meteorological forecast.

For further evaluation, Fig. 4 displays the MB, MNMB, RMSE, FGE and R (defined in 810 811 Appendix) of ozone of the seven individual models and the ENSEMBLE calculated using the representative observations available over the whole European domain. These statistics are 812 based on seven consecutive 96h forecasts run every day from the 9th of June to the 15th of 813 June. This figure shows that there is a spread of the 7 models and that the ENSEMBLE 814 815 generally gives the best scores with MNMB between 0.2 and -0.1, FGE between 0.15 and 0.4 and correlations up to 0.75 during daytime. All the models, including the ENSEMBLE, 816 exhibit a diurnal cycle with higher correlations and lower RMSE and FGE during daytime 817 818 (when ozone is high) than during nighttime. Five of the models have a positive MB on average and the other two a negative MB on average. The statistics are only calculated here 819 on one week but there is a good consistency with scores based on longer time series, as shown 820 and analysed in detail in Section 3.4. 821

To illustrate the behaviour of the MACC-II ensemble system in the case when some of the 7 models are missing for the production of the ENSEMBLE, we selected as an example the period of the 9th to the 15th of June 2014 corresponding to the ozone episodes discussed in Section 3.2 and we compared the following ensembles:

- 'MEDIAN 7', the operational ensemble method which is the median of the 7 models

827 (=ENSEMBLE) as presented in Fig. 4,

- 'MEDIAN 5', built on 5 individual models, after filtering out the best and the worst models,

- 829 according to the criterion described below,
- 'MEDIAN 3', built on 3 individual models, after filtering out the two best and the two worst
 models, according to the criterion described below,
- e '1BEST', the best model.

By removing at the same time the best (or two best) and the worst (or two worst) models we estimate an "average situation". Since the relative performances of individual models vary in time and space, the criterion to order the 7 individual models from worst to best is measured by their RMSE over the 7 days of the verification between 12 UTC and 18 UTC (ozone peak time). This criterion is chosen on the basis that we look for the model best reproducing the high daytime ozone levels. RMSE is seen as the most objective criterion since MB and NMNB can include compensating effects and since there is a low spread between the models in the FGE. From this, the best model is displayed in purple colour in Fig. 4c and the worst in brown colour.

842 Results of the sensitivity experiments are shown in Fig. 5. This figure confirms that the 843 ensemble median (MEDIAN 7) using all 7 models performs generally better than the best model on all statistical indicators. When only five models (excluding the best and the worst) 844 845 are available to calculate the ensemble, all scores show only very slight differences with the ENSEMBLE (MEDIAN 7) based on 7 models. Going to only three models to calculate the 846 847 ensemble (MEDIAN 3) leads to statistical indicators degraded compared to the ensemble from 7 (MEDIAN 7) or 5 (MEDIAN 5) models but performs generally better than the best 848 model (1BEST). This indicates that using an ensemble of models, even if reduced, is more 849 useful than using a single model event of very good quality. This also shows that with 5 850 models available (that may happen in case of problems of production of 2 of the 7 models), 851 the ensemble is still robust compared to observations. 852

In our tests we disregarded the worst (or 2 worst) and best (or two best) models on a RMSE criterion but Kioutsioukis and Galmarini (2014) showed that there is an impact of the quality of the models chosen on AQ ensemble performances. To go a step further, a more comprehensive study is done in Section 3.4 over a longer time period.

857

858 **3.4 Statistical performances of the forecasts on a seasonal basis**

859 Additionally to the production of daily skill scores, statistical indicators are calculated for ozone, NO_2 and PM_{10} at surface on a seasonal basis since September 2009 for each of the 860 861 seven models and for the ENSEMBLE. These skill scores and the analysis of their seasonal and year-to-year evolutions are presented in 6 monthly reports, each including 2 seasons. The 862 model statistical indicators are calculated against measurements from the European AQ 863 surface station network available in NRT and selected as detailed in Section 2.1. So far, the 864 data provision in NRT is not fully operational. Therefore, there is some variability with time 865 of the number of data available and of their location. Also, the spatial coverage of the surface 866 AQ network in Europe is very inhomogeneous with a high density of stations in France, 867

Germany, UK, Belgium and The Netherlands. Thus the statistical indicators are morerepresentative of the system skills for these countries.

Here we only focus on the performances of the system for the last year of MACC-II to document its status at the end of the project. We only analyse the two main pollutants for the season during which exceedances of regulatory levels are more likely to be encountered: ozone in summer (1st of June to 1st of September 2014) and PM₁₀ in winter (1st of December 2013 to 1st of March 2014). We are not showing scores for previous years since the use of a different set of surface observations from one year to another does not allow a fair comparison of the model skills.

MB, MNMB, RMSE, FGE and R for ozone in summer 2014 from the 7 models and the 877 ENSEMBLE are shown in Fig. 6. One main feature, that is common to all models including 878 the ENSEMBLE, is that there is no day-to-day degradation of MB, MNMB, and FGE 879 indicators from Day0 to Day3 and a slight increase of the RMSE around 15UTC (about 1 µg 880 881 m^{-3}). This indicates that the values of the surface ozone concentrations are not affected on average by the day-to-day degradation of the meteorology but are rather driven by other 882 883 processes such as emissions of precursors and chemistry. However, correlations (Fig. 6e) tend to decrease from Day0 to Day3. This tendency was also found in scores calculated for 884 previous years. Correlations give a measure of the ability of each model to fit the time 885 886 variations of ozone regardless of concentration biases. Therefore, correlations are more sensitive to the meteorological forecast skills than MB, MNMB, RMSE and FGE. 887 Nevertheless, the decrease of the correlation with forecast day is slow. 888

In Fig. 6 there is a marked diurnal cycle of all statistical indicators for all models which leads 889 to a similar diurnal cycle in the ENSEMBLE scores. MNMB, FGE and R show best 890 performances peaking at 15UTC and worst peaking at 06UTC for each of the 4 days of the 891 forecast. This means that all models are able to simulate the ozone daytime photochemistry 892 893 with the given setup of MACC-II (IFS forecasts for meteorology, C-IFS for chemical boundary conditions and GFAS and TNO emissions). For all models, the diurnal cycle in the 894 895 statistical indicators can be at least partly explained by uncertainties in the diurnal cycle of the 896 emissions of ozone precursors used in the individual models. This is illustrated by CHIMERE correlation at night which is better than most of the other models. CHIMERE has developed 897 898 diurnal factors for traffic emissions based on an objective analysis of NO₂ measurements in the different countries in Europe which improves ozone titration at night (Menut et al., 2012). 899 900 Other reasons of the diurnal cycle in the model scores could also be errors in the diurnal cycle

of the boundary layer height and associated vertical diffusion. For instance, the boundary 901 layer in the LOTOS-EUROS simulations is described with a single model level, with a diurnal 902 variation in the boundary layer height obtained 3-hourly from the ECMWF forecasts. This 903 904 differs from the description of vertical mixing in the other models and may be responsible for 905 the low correlation feature at around 9 UTC. MATCH shows the largest diurnal variability that can be partly related to a combination of chemistry, deposition and the vertical resolution, 906 where the latter is inherited from the IFS model with a rather shallow lowest model layer 907 (~20m). The ozone depletion processes at the surface appears too strong and not enough 908 909 compensated by the vertical diffusion. The MB is then more pronounced during night time, and a modification of the vertical diffusion has shown to improve MATCH skill. 910

Figure 6 also shows that there is generally a positive bias (both in MB and MNMB) of the 911 912 ENSEMBLE for each of the 4 days of the forecast except around end afternoon when there is a slight ozone underestimation. The ENSEMBLE MB varies from -6 to 15.5 μ g m⁻³. This is 913 consistent with the behaviour of the individual models, most of them having a positive bias on 914 915 average. Only the MATCH model shows a negative bias (MB and MNMB) which has the same explanation as described above. The SILAM model has the highest positive ozone bias 916 917 (MB and MNMB) on average. High ozone concentrations in SILAM are largely explained by too low dry deposition velocity over terrestrial areas, especially on the vegetated surfaces. The 918 919 new scheme explicitly accounting for the leaf area index is being tested in pre-operational regime. 920

The normalized indicators (MNMB and FGE) of all 7 models vary from -0.45 to 0.45 and from 0.16 to 0.59, respectively. This means that all models perform fairly well for ozone in summer 2014. For the ENSEMBLE, MNMB and FGE are on average 0.14 (varying from -0.03 to 0.33) and 0.30 (varying from 0.16 to 0.45). These values, being well below 1, confirm the good skills of the ENSEMBLE. The ENSEMBLE correlation varies from 0.42 at night to 0.65 during daytime, consistently with the scores of the individual models.

The seven models show a fairly similar overall behaviour against observations because of their common framework (meteorology, chemical boundary conditions and emissions). Nevertheless, there are differences between ozone forecasted by each of the individual models because of their specificities (different chemistry schemes, different implementations for use of input data, different physical parameterizations). The ENSEMBLE gives generally better scores for ozone than any of the individual models.

In Section 3.3, we made a first investigation of the robustness of the ensemble median method 933 934 with regard to the number of models available on a case study of one week. To go a step further, we ran a series of tests by removing one or more models in the calculation of the 935 ensemble median over the three months of summer 2014. The removal is done randomly on 936 each of the daily forecasts. Figure 7 shows the statistical results against observations for the 937 ENSEMBLE (7 models) and the other ensemble medians calculated by removing randomly 1, 938 2, 3 or 4 models. For MB, MNMB and FGE, there is hardly any difference between all 939 ensembles. Only RMSE and R (correlation) give significant changes. As expected, decreasing 940 941 the number of models used in the ensemble tends to degrade its performances. Using 6 942 models gives RMSE and R close to the full ensemble based on 7 models. The scores for 943 ensembles with 4 and 5 models are close to each other but are degraded compared to when 7 or 6 models are used. When only 3 models are used, RMSE and R are worse compared to the 944 other configurations by ~0.5 μ g/m³ and ~0.05, respectively. This shows that, the multi-model 945 ENSEMBLE at the end of MACC-II, which is based on the median of 7 models, is robust 946 947 even if 2 to 3 models are unavailable. These results are consistent with the results discussed in 948 Section 3.3 that were calculated on one week and with a different method for the model removal. 949

950 Figure 8 shows the ensemble scores for PM_{10} for the last winter (2013-2014) of MACC-II. Since there were much fewer observations available at 00UTC compared to other times of the 951 952 day, the values given at the forecasts times of 0h, 24h, 48h, 72h and 96h show a specific behaviour that is not analysed since not typical. Among the 7 models, MATCH PM₁₀ scores 953 954 are the poorest for the period. This has been traced down to an error in the sea-salt emissions leading to too strong emissions and a coding error regarding summation of the various aerosol 955 components that builds up PM₁₀ in the model. Secondary organic aerosols are not yet included 956 and there should be an underestimate, rather than an overestimate of PM₁₀ by MATCH. The 957 poor correlations in the period are partly related to a too strong signal from sea salt. The 958 verification for 2015 shows clearly that correction of these errors has improved the simulation 959 of PM₁₀. There is also a positive bias (~3.5 μ g m⁻³ on average for MB and 0.3 for MNMB) for 960 CHIMERE. This could be due to specific set-up in CHIMERE to be more efficient in the 961 detection of PM₁₀ threshold exceedances during wintertime. It is achieved with a correction 962 for lowering the wind over urban areas and with the modulation of the emissions from 963 domestic heating to account for the impact of extremely low temperatures, occurring during 964 cold surge for instance. The other extreme models are MOCAGE and SILAM that exhibits a 965

negative bias of about -6 µg m⁻³ on average for MB and -0.37 for MNMB. For MOCAGE, 966 this is due to the lack of secondary aerosols in the model. Although secondary aerosols are not 967 dominant in PM₁₀ in winter, there is an expected contribution of sulfates mainly in Eastern 968 Europe. Substantial bias of SILAM PM₁₀ is caused by the missing SOA. However, the model 969 showed the highest spatial correlation with the PM₁₀ observations, which largely follows from 970 971 the detailed treatment of sea salt emission and transport. As a result, SILAM also showed among the lowest RMSEs. LOTOS-EUROS, EURAD-IM, EMEP have a smaller negative 972 MB and NMNB. LOTOS-EUROS has an advanced treatment of the SIA fraction, but the 973 974 secondary organic aerosols (SOA) are not included yet, which explains part of the negative bias. The bias in EMEP and EURAD-IM is relatively small as these two models include a 975 comprehensive treatment of SIA and SOA. The ENSEMBLE MB and MNMB both indicate a 976 low bias related to the fact that 5 of the 7 models have a negative bias. RMSE and FGE (Figs. 977 8c and 8d) are consistent with bias scores with largest values for MATCH and MOCAGE. 978

979 In Fig. 8, MB, MNMB, RMSE and FGE are best during daytime (generally around 06-07UTC and 15UTC) with diurnal variations fairly similar for all models. This is related to the fact that 980 PM₁₀ are dominated by primary anthropogenic emissions of black and organic carbon which 981 are prescribed in all model by the same TNO inventories and which have maxima in the 982 983 morning and in the afternoon. Worst MB, MNMB, RMSE and FGE are at night, as for ozone. 984 This may be linked to uncertainties in the boundary layer height at night, in vertical diffusion 985 and/or to an underestimation of emissions. The diurnal cycle is less marked in the correlation but there is a significant day-to-day decrease of skills. As for ozone, this decrease is likely 986 987 linked to a decrease of the meteorological forecast skills with time which affects more the correlation of pollutants than the other statistical indicators. Correlation values are fairly low, 988 a bit higher than 0.4 at maximum. This is due to the lack of certain types of aerosols (SIA 989 and/or SOA) in some models but also likely to uncertainties in the diurnal cycle of the 990 991 anthropogenic emissions prescribed in the models and of the boundary layer height and vertical diffusion. The important contribution of traffic and residential heating to the 992 993 anthropogenic emissions of aerosols in Europe is generally modeled as two fixed peaks that do not fully take into account the differences in habits between the countries in Europe. Also, 994 sea salts contribute to PM₁₀ on the Western side of Europe with emissions and scavenging 995 depending closely to meteorological conditions and therefore directly affected by 996 meteorological uncertainties. 997

The ENSEMBLE scores are best compared to the 7 individual models for RMSE and FGE.For MB and MNMB, EURAD-IM and EMEP models perform better than the ENSEMBLE.

1000 This shows that there are compensating positive/negative biases in EURAD-IM and EMEP 1001 that are removed when RMSE and FGE scores are considered. The ENSEMBLE correlation 1002 is higher than the other models except SILAM.

Figures 6 and 8 show that the seven forecasts on which the ENSEMBLE is calculated are less skillful in modeling the aerosols than ozone. This is a common feature of most chemistry models since there are still large uncertainties on primary aerosol emissions and processes of production and evolution of secondary aerosols, particularly of secondary organic aerosols. Moreover, because of the operational context of MACC-II production, the seven forecasts models are optimized to run in short times. This constrains the level of detail of aerosol processes that can be afforded.

1010

3.5 Example of the specific evaluation for the Mediterranean area

1012 Within the European continent, the Mediterranean area is characterized by special features -1013 high emission densities due to concentration of human activities in surrounding coastal areas, 1014 intense photochemistry, high background pollution, small scale meteorology that make air quality forecasting specially challenging. This is why work has been specifically carried out 1015 to evaluate the seven models and the ENSEMBLE in this region. This is complementary to 1016 1017 the systematic daily and seasonal evaluation performed over the whole European continent. 1018 Its aim is not about scoring the system but on a better scientific understanding of the behaviour of the seven models and the ENSEMBLE in the Mediterranean region. This work 1019 1020 is based, firstly, on two high resolution models run daily over eastern (Greece) Mediterranean 1021 and western (Spain) areas and surface station measurements that are not used in the 1022 operational MACC evaluation and, secondly, on scientific analyses of case studies.

1023 For the Eastern Mediterranean area, the LAP-AUTH forecasting system is run daily. It 1024 consists of the Weather Research and Forecasting mesoscale meteorological model (WRF 1025 version 3.2) (Skamarock et al., 2008) and the chemistry transport model Comprehensive Air 1026 quality Model with extensions (CAMx version 5.30) (ENVIRON, 2010). The anthropogenic 1027 emission data, used as CAMx input data, are from Kuenen et al. (2014) for the reference year 2009. Anthropogenic emissions data are temporally processed using the Model for the Spatial 1028 1029 and tEmporal diStribution of emissionS (MOSESS) (Markakis et al., 2013). The emissions originating from natural sources are calculated with the use of the emission model namely 1030 NEMO (Natural Emission MOdel) (Markakis et al. 2009). Wind erosion dust, sea salt and 1031 biogenic NMVOCs emissions are calculated using the WRF model meteorology. The air 1032

1033 quality forecasting system derives meteorological initial and boundary conditions from the operational 12:00 UTC forecast of ECMWF while chemical boundary conditions derived 1034 from the IFS-MOZART global model forecast and replaced by C-IFS from September 2014. 1035 The domain of the WRF - CAMx implementation is the South-East Europe/Eastern 1036 Mediterranean region from 18°N-30° N and 34.9°E–44.5E. The grid resolution is 10 x 10 km². 1037 The air quality modelling system runs on a daily basis in order to produce 72-hours air quality 1038 forecasts. For the verification, the WRF-CAMx, the ENSEMBLE and the seven models are 1039 compared with available air quality data from the GMEECC (Greek Ministry of Environment 1040 1041 Energy and Climatic Change) air pollution monitoring network as well as from the background station of Finokalia, operated by the University of Crete (Greece). 1042

AEMET runs a version of the MOCAGE (Josse et al. 2004) model at 0.05° horizontal 1043 1044 resolution in the Western Mediterranean coast, daily up to 48h using the ENSEMBLE forecasts as chemical lateral boundary conditions. Meteorological forcings for the high 1045 resolution domain come from operational HIRLAM run every 6 hours at AEMET (Navascues 1046 1047 et al. 2013). Emissions over land in this domain come from the GEMS-TNO inventory (Visschedijk et al, 2007). The domain is 44°N-36°N-5°W-5°E. The ENSEMBLE has been 1048 compared to the AEMET forecasts and to observations from EMEP/GAW Spanish stations 1049 1050 and from different local and regional Air Quality Monitoring networks. From these high resolution daily forecasts, a collection of case studies in which high resolution could have 1051 been an advantage, has been selected and analysed. These comparisons show the high 1052 variability of results between model forecasts depending on the location, time and day, 1053 1054 whereas, sometimes, model forecast agreement is quite noticeable.

1055 We are presenting here a brief summary of the analysis of the case study that occurred between the 15th and the 18th of July 2013, when high values of ozone were measured in many 1056 Spanish Air Quality Monitoring Stations due to very strong solar radiation and high 1057 temperatures together with persistent anticyclonic conditions and very weak pressure 1058 gradients. Ozone concentrations at surface above 140 μ g/m³ were not rare at the stations used 1059 in this period and values above 120 μ g/m³ were common. Figure 9 shows two maps with the 1060 18th July 2013 ENSEMBLE and AEMET model at H+18 forecasts and the observations over-1061 1062 plotted using the same colour intervals. The ENSEMBLE forecasts generally fit well to the measurements. The main characteristic of the ENSEMBLE forecasts (left) is that it is too 1063 1064 smooth to capture all the small scale features occurring in reality because of its horizontal resolution (~15 km). As an example, we can look at Fig. 9 in which the Madrid area has been 1065

magnified to observe how ozone values between 100 and 160 μ g/m³ were measured by different Air Quality Networks (belonging to Madrid Regional Authorities and Madrid City Council) whereas in the ENSEMBLE forecasts all the concentrations are lying in the 100-120 μ g/m³ interval. In the same period, the AEMET forecasts provides values in this area with a higher spread, between 100 and 160 μ g/m³ which fits better to observations. Something similar can be observed in the Eastern Spain area, also magnified in the same figure.

Illustrations of the results of ENSEMBLE and AEMET models compared to three EMEP 1072 1073 stations are given in Figs. 10 to 12 for ozone during summer. For Cap de Creus (Fig. 10), 1074 there is a good agreement between observations and models, but the two models have a wider 1075 diurnal cycle in concentrations. This behaviour can be related to the local dynamics. The area is located in a strong wind zone on the North-East coast of Spain. Therefore, ozone formed in 1076 1077 this area can be transported rapidly, prohibiting ozone accumulation and leading to a smoother diurnal cycle than in the models that are not able to represent this local effect. For 1078 Mahon (Fig. 11), both models fit fairly well the observations but observations have generally 1079 1080 a wider diurnal cycle in concentrations, contrarily to Cap de Creus. Mahon station is located in the small island of Menorca in the Balearic Archipelago. It is sometimes exposed to the 1081 pollution produced by ships entering the port early in the morning, leading to high NOx 1082 conditions and low ozone. This could explain the observed decrease of ozone at this time. 1083 This local effect is not captured by the two models. San Pablo de Montes measurements (Fig. 1084 12) show a different behaviour with generally higher concentrations than the ENSEMBLE 1085 and AEMET models. The discrepancy is particularly important during July. This can be 1086 1087 explained by an underestimation of the local isoprene emissions in the two models. Isoprene 1088 concentrations measured at San Pablo de los Montes exhibit higher values than other sites in the same area because of the oak vegetation surrounding the station. 1089

Overall, the quality of the ENSEMBLE forecasts is generally good and the verification scores of the forecasts calculated for the whole period of the project show, most of the time, better results for the ENSEMBLE than for the AEMET forecasts. The limitations of the verification carried out (only the 7 EMEP background air quality stations within the domain have been considered) and the different high resolution emission inventories used in AEMET and ENSEMBLE can be part of the reason for these different results.

1096 Another product we have started to generate at the end of the project is the behaviour of 1097 forecasts of the seven models together with the ENSEMBLE and the AEMET forecasts 1098 against observations from the EMEP Air Quality Network. An example is presented in Fig.

13. In this figure, we can see the ozone forecasts at ES10 station which is located at Cabo de 1099 Creus in the Northeastern corner of Spain (42.32N, 3.32 E). We observe that the spread 1100 between the seven model forecasts in the H+24 to H+48 forecast period from the 9th April 1101 2014 is fairly low with most of the members producing similar forecasts. It changes quickly 1102 1103 on the next day at the same place with the seven models providing very different concentrations leading to a high spread. We have also observed differences in the spread of 1104 the members at other locations on the same day and forecast time. More generally this pattern 1105 with very different spreads (ranging from low to high) depends on the case studies: day, time 1106 1107 period and location. The analysis of the spread between different model forecasts in the same 1108 period can help modellers to understand how their models behave in the Mediterranean area.

1109

1110 4. Conclusion and future developments

In this paper, we give an overview of the current state and performances of the forecasting 1111 system for European air quality that was put in place in the framework of MACC project and 1112 continued during MACC-II project and now in the MACC-III project. Its strength comes from 1113 the fact that it is based on an ensemble of seven state-of-the-art chemistry-transport models 1114 (CHIMERE, EMEP, EURAD-IM, LOTOS-EUROS, MOCAGE, MATCH, SILAM) that are 1115 1116 developed and run by recognized institutes in Europe. It also relies on good quality inputs for meteorological forcings, emissions and chemical boundary conditions. It provides daily 4-1117 days forecasts for 6 major pollutants (O₃, NO₂, SO₂, CO, PM₁₀ and PM_{2.5}) and birch pollen 1118 during pollen season, and also additional species for downscaling air quality modeling 1119 1120 purposes. The production also includes hourly analysis for the previous day. Daily statistical performances of the forecasts against available European air quality monitoring station are 1121 processed on daily, weekly and 3 monthly bases, giving an objective assessment of the 1122 products to users. They are also used to monitor the seasonal and yearly evolutions of the 1123 1124 forecast scores.

Because of the resolution of the seven models (10 to 20 km), this system is not designed and do not attempt to forecast very local concentrations but large scale phenomena and background air pollution. The ENSEMBLE has the capability to forecast pollution episodes at the regional scale as illustrated over the period from the 10th to the 13th of June 2014. On a seasonal basis, the 7 models show good statistical performances for ozone in summer 2014 and the ensemble median outperforms any of the individual models. The normalized indicators, that are less sensitive to outliers than MB and RMSE, are low, varying for the

ENSEMBLE from -0.03 to 0.33 for MNMB and from 0.16 to 0.45 for FGE. The diurnal 1132 ozone peak is underestimated by the ENSEMBLE by about 4 μ g m⁻³ on average during 1133 summer 2014. The underestimation is larger during ozone episodes, such as during 9-15 June 1134 2014, when the ENSEMBLE underestimates ozone daily maxima on average by about 10 µg 1135 m⁻³. Comparing locally to surface station measurements within the pollution episode areas, 1136 the ENSEMBLE low bias is often between 30 and 50 μ g m⁻³ but one of the models is often 1137 close to the observed values. For PM_{10} , there is a negative bias of the ENSEMBLE with MB ~ 1138 -4.5 mg m-3 and MNMB \sim -0.1, on average. The ENSEMBLE FGE is larger than for ozone 1139 1140 (~0.52 on average for PM_{10} and 0.30 for ozone) and the correlation is lower (~0.35 on average for PM_{10} and 0.54 for ozone). There is a large variability of the statistical indicators of the 7 1141 1142 models for the last winter of MACC-II period (winter 2013-2014). This is related to different levels of complexity in the representation of aerosols in the models. PMs are regulatory 1143 pollutants that are difficult to forecast mainly because of uncertainties in primary aerosol 1144 1145 emissions, our partial knowledge on secondary organic aerosol processes and the constraint of timely production that prevents using very sophisticated representations of secondary 1146 aerosols. A possibility of improvement of the ensemble median performances at low cost 1147 1148 could be to remove the bias of the individual models before the ensemble median calculation. To complement the statistical evaluation done over the whole European domain, a scientific 1149 evaluation of the seven models and of the ENSEMBLE is also done for the Mediterranean 1150 1151 region because of its specificities (emissions, population, topography, meteorology, photochemistry). Another important point to note is that major efforts have been put during 1152 MACC-II towards the full operationalisation of the system in order to improve its robustness. 1153

1154 The regional air quality production was extended during MACC-II and further developments are underway to improve the quality, the variety and the timeliness of its products based on 1155 users' feedbacks. In the very short term, the ENSEMBLE analysis that is only provided for 1156 ozone will be extended to NO₂ from January 2015 and verification statistics with independent 1157 data will be produced. It is also planned to shift the ENSEMBLE analysis production time 1158 earlier, at 11UTC from early 2015 following the users' recommendation. One planned change 1159 1160 in the mid-term will be to have all individual models run at a ~10km horizontal resolution. This should improve the performances of the system compared to observations. Also the 1161 regional production benefits and will continue to benefit from the evolutions and 1162 improvements of the global production such as the use of the newly operated C-IFS (fully 1163 1164 coupled chemistry to the IFS meteorological model) since September 2014 for regional
boundary conditions for chemical species and aerosols. In parallel, a dedicated fire emission 1165 1166 product for regional forecast purposes, available earlier than the current operational product, is progressively implemented in the seven models and its usefulness will be assessed. Product 1167 evaluation is done on the basis of the available NRT measurements from the European AQ 1168 monitoring network. Stations used are selected to be as representative as possible of the model 1169 horizontal resolution, by retaining only classes 1 to 5 from the Joly and Peuch (2012) 1170 classification. There is ongoing work to improve the station selection, still based on Joly and 1171 1172 Peuch's classification, by determining the best class ranges to be used individually for each of 1173 the 6 pollutants (O₃, NO₂, SO₂, CO, PM₁₀, PM_{2.5}) based on pollutant lifetime.

1174 Continuous research is pursued to improve the seven individual models and their assimilation systems. In particular, there is an important effort on the use of new satellite data or 1175 1176 combinations of satellite data with surface measurements in the assimilation systems. Also, there is on-going work on ensemble methods in order to extract as much value as possible 1177 from the seven model forecasts. Alternative methods to the median are currently tested: 1178 1179 application of weights on the individual models at each grid point related to the performances from the day before or spectral decomposition (Galmarini et al. 2013). The results of these 1180 alternative methods applied to the MACC-II multi-model ensemble will be the subject of a 1181 forthcoming paper. Another goal in MACC-II was the start of research and developments for 1182 the modelling of CO₂ in the regional models in view of potential future high-resolution 1183 surface CO₂ flux inversion products over Europe. This work will be pursued. 1184

In the next few years, the availability of more daily European surface observations, in a wider European area (i.e. from more countries) and at earlier time is foreseen. More data on a wider area would improve the strength of the statistical product evaluation. The continuous improvement of the quality of the surface monitoring data is also important for performance evaluation. Earlier availability of the surface station data would give the opportunity of an earlier production of the analyses with the goal of using the analyses as the initial state for the forecasts.

1192 Other studies will be conducted on the possibility to provide complementary indicators such 1193 as the exceedences in ozone or PM_{10} . In the future, the production could be extended to other 1194 types of pollens than birch. There are currently some developments to test olive, grass and 1195 ambrosia pollens based on work done at the Finnish Meteorological Institute. Also, the 1196 possibility to produce additional species will be considered for users running forecast systems at finer scales than the MACC-II system, such as the concentrations of different types ofaerosols.

1199

1200 Appendix: Statistical indicators

1201 The forecast performances are measured using five statistical indicators: the mean bias, the 1202 root mean square error, the modified normalized mean bias, the fractional gross error and the 1203 correlation.

1204 The mean bias captures the average deviations between two datasets and is defined as:

$$MB = \frac{1}{N} \sum_{i} (f_i - o_i)$$

1205 Where f_i and o_i are the forecast value at the observation location and the observation value, 1206 respectively.

1207 The root mean square error combines the spread of individual error and is defined as:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i}(f_i - o_i)^2}$$

It should be noted that the RMSE is strongly dominated by the largest values, due to the 1208 squaring operation. Especially in cases where prominent outliers occur, the usefulness of the 1209 RMSE is questionable and the interpretation becomes more difficult. MB and RMSE are not 1210 1211 dimensionless variables, but have the same dimension as the modelled/observed quantity and requires knowledge of typical mean values. By scaling the MB and RMSE to the observations 1212 these metrics can be made relative, dimensionless, and hence more appropriate for use as a 1213 1214 score. This is relevant when comparing bias and RMSE of atmospheric species whose concentrations can vary by orders of magnitude. This is why the modified normalized mean 1215 1216 bias (MNMB) and the fractional gross error (FGE) are also used. MNMB is defined as:

$$MNMB = \frac{2}{N} \sum_{i} \left(\frac{f_i - o_i}{f_i + o_i} \right)$$

1217 This gives a measure of forecast bias bounded by the values -2 to +2. It performs 1218 symmetrically with respect to under and over-prediction of the observations, which is a 1219 desirable feature.

1220 FGE is defined as:

$$FGE = \frac{2}{N} \sum_{i} \left| \frac{f_i - o_i}{f_i + o_i} \right|$$

FGE gives a measure of the overall forecast error. This is proposed in addition to the more traditional RMSE, because due to the squaring procedure the RMSE gives the largest weight to the (possibly spurious) largest observations. FGE is bounded between 0 and 2.

In addition, the correlation coefficient is needed to indicate the extent to which patterns in the forecast match those in the observations. The correlation coefficient R between the forecast and observed values is defined as:

1227
$$R = \frac{\frac{1}{N}\sum_{i}(f_{i} - \overline{f})(o_{i} - \overline{o})}{\sigma_{f}\sigma_{o}}$$

where \overline{f} and \overline{o} are the mean values of the forecast and observed values and σ_f and σ_o are the corresponding standard deviations. The correlation coefficient has a maximum value of unity when, for each observation site, $(f_i - \overline{f}) = c(o_i - \overline{o})$, where *c* is a positive constant. In this case the two datasets have the same pattern of variation but are not identical unless *c*=1 for all sites.

1233

1234 Acknowledgements

1235 This study was funded by the European Commission under the EU Seventh Research Framework Programme (grant agreement no. 283576, MACC II). In-situ air quality data were 1236 1237 provided by the European Environment Agency. Additional financial support at national level was given by the French Ministère de l'écologie, du développement durable et de l'énergie 1238 1239 through the ADONISS project. This work was granted access to the HPC resources of CCRT 1240 under the allocation 2013-6695 made by GENCI (Grand Equipement National de Calcul 1241 Intensif). IASI has been developed and built under the responsibility of the Centre National d'Etudes Spatiales (CNES, France). Developments of SILAM system were supported by 1242 1243 projects ASTREX and IS4FIRES of Academy of Finland. NO₂ column retrievals from AURA/OMI and METOP/GOME-2, and CO profiles from TERRA/MOPITT have been used 1244 for assimilation in some of the individual models. 1245

1247 **References**

- 1248 Ackermann, I.J., Hass, H., Memmesheimer, M., Ebel, A., Binkowski, F.S., and Shankar, U.:
- Modal aerosol dynamics model for Europe: Development and first applications, Atmos.Environ., 32, 2981-2999, 1998.
- 1251 Alfaro, S.C., and Gomes, L.: Modeling mineral aerosol production by wind erosion: Emission
- intensities and aerosol size distribution in source areas, J. Geophys. Res., 106, 18,075-18,084,
- 1253 2001.
- Andersson-Sköld, Y., and Simpson, D.: Comparison of the chemical schemes of the EMEP
 MSC-W and IVL photochemical trajectory models, Atmospheric Environment, 33, 1111–
 1129, 1999.
- Andersson, C., Langner, J., and Bergström, R.: Interannual variation and trends in air
 pollution over Europe due to climat variability during 1958-2001 simulated with a regional
 CTM coupled to the ERA40 reanalysis. Tellus 59B, 77-98, 2007.
- Andersson, C., Bergström, R., Bennet, C., Robertson, L., Thomas, M., Korhonen, H.,
 Lehtinen, K.E.J. and Kokkola, H.: MATCH-SALSA Multi-scale Atmospheric Transport
 and Chemistry model coupled to the SALSA aerosol microphysics model Part 1: model
 description and evaluation. Geosci. Model Dev., 8, 171-189, 2015.
- 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, 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. T., Savage, N., Seigneur, C., Sokhi, R. S., Solazzo, E.,
- 1269 Solomos, S., Sørensen, B., Tsegas, G., Vignati, E., Vogel, B., and Zhang, Y.: Online coupled
- 1270 regional meteorology chemistry models in Europe: current status and prospects, Atmos.
- 1271 Chem. Phys., 14, 317-398, doi:10.5194/acp-14-317-2014, 2014.
- Barbu, A.L., Segers, A.J., Schaap, M., Heemink, A.W., and Builtjes, P.J.H.: A multicomponent data assimilation experiment directed to sulphur dioxide and sulphate over
 Europe, Atmos. Env., 43(9), 1622-1631, 10.1016/j.atmosenv.2008.12.005, 2008.
- 1275 Barré, J., El Amraoui, L., Ricaud, P., Lahoz, W. A., Attié, J.-L., Peuch, V.-H., Josse, B.,
- 1276 and Marécal, V.: Diagnosing the transition layer at extratropical latitudes using MLS O₃ and
- 1277 MOPITT CO analyses Atmos. Chem. Phys., 13, 7225-7240, 2013.

- 1278 Barré, J., Peuch, V.-H., Lahoz, W. A., Attié, J.-L., Josse, B., Piacentini, A., Emerenko, M.,
- 1279 Dufour, G., Nédélec, P., von Clarmann, T., and El Amraoui, L.: Combined data assimilation
- 1280 of ozone tropospheric columns and stratospheric profiles in a high-resolution CTM. Q. J. Roy.
- 1281 Meteorol. Soc., 140, 966-981, doi:10.1002/qj.2176, 2014.
- 1282 Bechtold, P., Bazile, E., Guichard, F., Mascart, P., and Richard, E.: A mass-flux convection
- scheme for regional and global models, Q. J. R. Meteorol. Soc., 127, 869–886, 2001.
- Beekmann, M., and Vautard, R.: A modelling study of photochemical regimes over Europe:
 robustness and variability, Atmos. Chem. Phys., 10, 10067-10084, 2010.
- 1286 Benedetti, A., Morcrette, J.-J., Boucher, O., Dethof, A., Engelen, R. J., Fisher, M., Flentje,
- 1287 H., Huneeus, N., Jones, L., Kaiser, J. W., Kinne, S., Mangold, A., Razinger, M., Simmons,
- 1288 A. J., and Suttie, M.: Aerosol analysis and forecast in the European Centre for Medium-Range
- 1289 Weather Forecasts Integrated Forecast System. 2. Data assimilation: J. Geophys. Res., 114,
- 1290 D13205, doi:10.1029/2008JD011115, 2009.
- 1291 Bergström, R., Denier van der Gon, H. A. C., Prévôt, A. S. H., Yttri, K. E., and Simpson, D.:
- 1292 Modelling of organic aerosols over Europe (2002–2007) using a volatility basis set (VBS)
- 1293 framework: application of different assumptions regarding the formation of secondary organic
- aerosol, Atmos. Chem. Phys., 12, 8499–8527, doi:10.5194/acp-12-8499-2012, 2012. TS4
- Bessagnet, B., Hodzic, A., Vautard, R., Beekmann, M., Cheinet, S., Honoré, C., Liousse, C.,
 and Rouil, L.: Aerosol modeling with CHIMERE: preliminary evaluation at the continental
 scale, Atmospheric Environment, 38, 2803--2817, 2004.
- Bessagnet, B., Menut, L., Curci, G., Hodzic, A., Guillaume, B., Liousse, C., Moukhtar, S.,
 Pun, B., Seigneur, C., and Schulz, M.: Regional modeling of carbonaceous aerosols over
 Europe Focus on Secondary Organic Aerosols, Journal of of Atmospheric Chemistry, 61,
 175-202, 2009.
- Bott, A.: A positive definite advection scheme obtained by nonlinear renormalization of theadvective fluxes, Mon. Weather Rev., 117, 1006-1016, 1989.
- 1304 Bousserez, N., Attié, J.-L., Peuch, V.-H., Michou, M., Pfister, G., Edwards, D., Avery, M.,
- 1305 Sachse, G., Browell, E., and Ferrare, E.: Evaluation of MOCAGE chemistry and transport
- 1306 model during the ICARTT/ITOP experiment, J. Geophys. Res., 112 (D120S42), doi: 10.1029/
- 1307 2006JD007595, 2007.

- Burridge, D.M., and Gadd, A. J.: The Meteorological Office Operational 10-level numerical 1308 weather prediction model, Sci. Pap., 34, UK Meteorological Office, 1977. 1309
- 1310 Carter, W.P.L: Condensed atmospheric photo oxidation mechanism for isoprene. Atmos. 1311 Environ. 30, 4275-4290, 1996.
- 1312 Colette, A., Granier, C., Hodnebrog, O., Jakobs, H., Maurizi, A., Nyiri, A., Bessagnet, B.,
- D'Angiola, A., D'Isidoro, M., Gauss, M., Meleux, F., Memmesheimer, M., Mieville, A.,
- 1314 Rouil, L., Russo, F., Solberg, S., Stordal, F., and Tampieri, F.: Air quality trends in Europe over the past decade: a first multi-model assessment, Atmos. Chem. Phys., 11, 11,657-11,678,
- 10.5194/acp-11-11657-2011, 2011. 1316

- Curier, R.L., Timmermans, R., Calabretta-Jongen, S., Eskes, H., Segers, A., Swart, D., and 1317 Schaap, M.: Improving ozone forecasts over Europe by synergistic use of the LOTOS-1318 EUROS chemical transport model and in-situ measurements, Atmos. Env., 60, 217-226, 1319 doi:10.1016/j.atmosenv.2012.06.017, 2012. 1320
- Cuvelier, C., Thunis, P., Vautard, R., Amann, M., Bessagnet B., Bedogni, M., Berkowicz, R., 1321
- Brandt, J., Brocheton, F., Builtjes, P., Coppalle, A., Denby, B., Douros, G., Graf, A., 1322
- 1323 Hellmuth, O., Honoré, C., Hodzic, A., Jonson, J., Kerschbaumer, A., de Leeuw, F., Minguzzi,
- E., Moussiopoulos, N., Pertot, C., Pirovano, G., Rouil, L., Schaap, M., Stern, R., Tarrason, L., 1324
- Vignati, E., Volta, M., White, L., Wind, P., and Zuber A.: CityDelta: A model 1325 intercomparison study to explore the impact of emission reductions in European cities in 1326 2010, Atmos. Env., 41, 189-207, 10.1016/j.atmosenv.2006.07.036, 2007. 1327
- Delle Monache, L., Deng, X., Zhou, Y., and Stull, R.: Ozone ensemble forecasts: 1. A new 1328 ensemble design, J. Geophys. Res., 111, D05307, doi:10.1029/2005JD006310, 2006. 1329
- De Ruyter de Wildt, M., Eskes, H., Manders, A., Sauter, F., Schaap, M., Swart, D., and van 1330 Velthoven, P.: Six-Day PM10 Air Quality Forecasts For The Netherlands With The 1331 Transport Model Lotos-Euros,, 45. p5586-5594,, 1332 Chemistry Atmos. Env., doi:10.1016/j.atmosenv.2011.04.049, 2011. 1333
- Dufour, A., Amodei, M., Ancellet, G., and Peuch, V.-H.: Observed and modelled "chemical 1334 weather" during ESCOMPTE, Atmos. Res., 74 (1-4), 161-189, 2004. 1335
- 1336 Ebel, A., Jakobs, H., Memmesheimer, M., Elben, H., and Feldmann, H.: Numerical forecast
- of air pollution: advances and problems, vol. Advances in Air Pollution Modeling for 1337 Environmental Security, Springer, doi:10.1007/1-4020-3351-6_14, 2005. 1338

- El Amraoui, L., Peuch, V.-H., Ricaud, P., Massart, S., Semane, N., Teyssèdre, H., and Karcher, F.: Ozone loss in the 2002–2003 Arctic vortex deduced from the assimilation of Odin/SMR O3 and N2O measurements: N2O as a dynamical tracer. Q. J. R. Meteorol. Soc., 1342 134(630), 217–228. doi:10.1002/qj.191, 2008.
- 1343 Elbern, H., Strunk, A., Schmidt, P., and Talagrand, O.: Emission rate and chemical state 1344 estimation by 4-dimensional variational inversion, Atmos. Chem. Phys., 7, 3749-3769, 2007.
- 1345 Elbern, H., Strunk, A., Friese, E., and Nieradzik, L.: Assessment of Source/Receptor
- 1346 Relations by Inverse Modelling and Chemical Data Assimilation, in Persistent Pollution Past,
- 1347 Present and Future School of Environmental Research Helmholtz-Zentrum Geesthacht,
- 1348 Quante, M.; Ebinghaus, R.; Flöser, G. (Eds.) 1st Edition, ISBN 978-3-642-17420-9, 2011.
- 1349 Emili, E., Barret, B., Massart, S., Le Flochmoen, E., Piacentini, A., El Amraoui, L.,
- 1350 Pannekoucke, O., and Cariolle, D. : Combined assimilation of IASI and MLS observations to
- constrain tropospheric and stratospheric ozone in a global chemical transport model, Atmos.Chem. Phys., 14, 177-198, 2014.
- ENVIRON: User's guide CAMx Comprehensive Air Quality Model with extensions,
 Version 5.30, ENVIRON International Corporation, 415.899.0700, December 2010, 2010.
- Fagerli, H. and Aas, W.: Trends of nitrogen in air and precipitation: Model results and
 observations at EMEP sites in Europe, 1980-2003, Environ. Poll., 154, 448–461, 2008.
- 1357 Flemming, J., Inness, A., Flentje, H., Huijnen, V., Moinat, P., Schultz, M. G., and Stein, O.:
- Coupling global chemistry transport models to ECMWF's integrated forecast system, Geosci.Model Dev., 2, 253-265, 2009.
- 1360 Flemming, J., Huijnen, V., Arteta, J., Bechtold, P., Beljaars, A., Blechschmidt, A.-M., Josse,
- 1361 B., Diamantakis, M., Engelen, R. J., Gaudel, A., Inness, A., Jones, L., Katragkou, E., Marécal,
- 1362 V., Peuch, V.-H., Richter, A., Schultz, M. G., Stein, O., and Tsikerdekis, A.: Tropospheric
- chemistry in the integrated forecasting system of ECMWF, Geosci. Model Dev., 7, 7733-7803, 2014.
- Friese, E., and Ebel, A.: Temperature dependent thermodynamic model of the system H^+ -NH₄⁺-Na⁺-SO₄²⁻-NO₄⁻-Cl⁻-H₂O, J. Phys. Chem. A., 114, 11595-11631, 2010.
- Foltescu, V.L., Pryor, C.S., and Bennet, C.: Sea salt generation, dispersion, and removal on
 the regional scale. Atmos Environ 39, 2113-2133, 2005.

- 1369 Fountoukis, C., and Nenes, A.: ISORROPIA II: a computationally efficient thermodynamic
- 1370 equilibrium model for K+-Ca2+-Mg2+-NH4+-Na+-SO42--NO3--Cl--H2O aerosols, Atmos.
- 1371 Chem. Phys., 7, 4639–4659, doi:10.5194/acp-7-4639-2007, 2007.
- 1372 Fuhrer, J., and Booker, F.: Ecological issues related to ozone: agricultural issues, Environ.
- 1373 Int., 29, 141–154, 2003.
- Galmarini, S., Kioutsioukis, I, and Solazzo, E.: E pluribus unum: ensemble air qualitypredictions, Atmos. Chem. Phys., 13, 7153-7182, 2013.
- 1376 Galperin, M. V.: The Approaches to Correct Computation of Airborne Pollution Advection,
- 1377 in: Problems of Ecological Monitoring and Ecosystem Modelling. XVII (in Russian).
- 1378 Gidrometeoizdat, St.Petersburg, pp. 54–68, 2000.
- 1379 Gaubert, B., Coman, A., Foret, G., Meleux, F., Ung, A., Rouil, L., Ionescu, A., Candau, Y.,
- 1380 and Beekmann, M.: Regional scale ozone data assimilation using an ensemble Kalman filter
- and the CHIMERE chemical transport model, Geosci. Model Dev., 7, 283–302, 2014.
- Geer, A.J., Lahoz, W.A., Bekki, S., Bormann, N., Errera, Q., Eskes, H.J., Fonteyn, D.,
 Jackson, D.R., Juckes, M.N., Massart, S., Peuch, V.-H., Rharmili, S., and Segers, A.: The
 ASSET intercomparison of ozone analyses : method and first results, Atmos. Chem. Phys., 6,
 5445-5474, 2006.
- Geiger, H., Barnes, I., Bejan, I., Benter, T., and Spttler, M.: The tropospheric degradation of
 isoprene: an updated module for the regional atmospheric chemistry mechanism, Atmos.
 Env., 37, 1503-1519, 2003.
- Genberg, J., Denier van der Gon, H. A. C., Simpson, D., Swietlicki, E., Areskoug, H.,
 Beddows, D., Ceburnis, D., Fiebig, M., Hansson, H. C., Harrison, R. M., Jennings, S. G.,
 Saarikoski, S., Spindler, G., Visschedijk, A. J. H., Wiedensohler, A., Yttri, K. E., and
 Bergström, R.: Light-absorbing carbon in Europe measurement and modelling, with a focus
 on residential wood combustion emissions, Atmos. Chem. Phys., 13, 8719–8738,
 doi:10.5194/acp-13-8719-2013, 2013.
- Giglio, L., Randerson, J. T., van der Werf, G. R., Kasibhatla, P. S., Collatz, G. J., Morton, D.
 C., and DeFries, R. S.: Assessing variability and long-term trends in burned area by
 mergingmultiple satellite fire products, Biogeosciences, 7, 1171–1186, doi: 10.5194/bg-71171-2010, 2010.

- Giorgi, F., and Chameides, W. L.: Rainout lifetimes of highly soluble aerosols and gases as
 inferred from simulations with a general circulation model, J. Geophys. Res.: Atmospheres,
 91, 14 367–14 376, 1986.
- 1402 Guenther, A. B., Hewitt, C. N., Erickson, D., Fall, R., Geron, C., Graedel, T., Harley, P.,

Klinger, L., Lerdau, M., Mckay, W. A., Pierce, T., Scholes, B., Steinbrecher, R., Tallamraju,
R., Taylor, J., and Zimmerman, P.: A global model of natural volatile compound emissions, J.

- 1405 Geophys. Res., 100, 8873–8892, doi:10.1029/94JD02950, 1995.
- Guenther, A., Karl, T., Harley, P., Wiedinmyer, C., Palmer, P., and Geron, C.: Estimates of
 global terrestrial isoprene emissions using MEGAN (Model of Emissions of Gases and
 Aerosols from Nature), Atmos. Chem. Phys., 6, 3181--3210, 2006.
- 1409 Guenther, A. B., Jiang, X., Heald, C.L., Sakulyanontvittaya, T., Duhl, T., Emmons, L.K., and
- 1410 Wang, X.: The Model of Emissions of Gases and Aerosols from Nature version 2.1
- 1411 (MEGAN2.1). An extended and updated framework for modeling biogenic emissions. Geosci.
- 1412 Model Dev., 5, 1471-1492, 2012.
- Guenther, A. B., Zimmerman, P. R., Harley, P. C., Monson, R. K., and Fall, R.: Isoprene and
 Monoterpene Emission Rate Variability: Model Evaluations and Sensitivity Analyses, J.
 Geophys. Res., 98, dx.doi.org/10.1029/93JD00527, 1993.
- Guth, J., Josse, B., Marécal, V., and Joly, M.: Simulating Secondary Inorganic Aerosols
 using the Chemistry Transport Model MOCAGE, Geosci. Mod. Dev. Discuss., 8, 3593-3651,
 2015.
- Hass, H., Jakobs, H.J., and Memmesheimer, M.: Analysis of a regional model (EURAD) near
 surface gas concentration predictions using observations from networks, Met. Atmos. Phys.,
 57, 173-200, 1995.
- Heimann, M., and Keeling, C.D.: A three-dimensional model of CO₂ transport based on
 observed winds. Model description and simulated trace experiment. In aspects of Climate
 Variability in the Pacific and Western Americas (ed. D.H Peterson). American Geophysical
 Union, Washington, DC, pp. 237-275, 1989.
- Hollingsworth, A., Engelen, R.J., Textor, C., Benedetti, A., Boucher, O., Chevallier, F.,
 Dethof, A., Elbern, H., Eskes, H., Flemming, J., Granier, C., Kaiser, J.W., Morcrette, J.J.,
 Rayner, P., Peuch, V.H., Rouil, L., Schultz, M.G., and Simmons, A.J.: and The GEMS
- 1429 Consortium: Toward a Monitoring and Forecasting System For Atmospheric Composition:

- 1430 The GEMS Project. Bull. Amer. Meteor. Soc., 89, 11471164, doi:10.1175/2008BAMS2355.1,1431 2008.
- Holtslag, A. A. M., and Moeng, C.-H.: Eddy diffusivity and countergradient transport in the
 convective atmospheric boundary layer. J. Atmos. Sci., 48, 1690-1700, 1991.
- Holtslag, A.A.M., and Nieuwstadt, F.T.M.: Scaling the atmospheric boundary layer,
 Boundary-Layer Met., 36, 201-209, 1986.
- Holtslag. A.A.M., Meigaard, E. van, and De Rooy, W.C: A comparison of boundary layerdiffusion schemes in unstable conditions over land. Boundary Layer Met., 76, 69-95, 1995.
- 1438 Honoré, C., Rouil, L., Vautard, R., Beekmann, M., Bessagnet, B., Dufour, A., Elichegaray,
- 1439 C., Flaud, J.-M., Malherbe, L., Meleux, F., Menut, L., Martin, D., Peuch, A., Peuch, V.-H.,
- 1440 and Poisson, N.: Predictability of European air quality : Assessment of 3 years of operational
- forecasts and analyses by the PREV'AIR system, J. Geophys. Res., D113, D04301,
 doi:10.1029/2007JD008761, 2008.
- 1443 Huijnen, V., Williams, J. E., van Weele, M., van Noije, T. P. C., Krol, M. C., Dentener, F.,
- 1444 Segers, A., Houweling, S., Peters, W., de Laat, A. T. J., Boersma, K. F., Bergamaschi, P., van
- 1445 Velthoven, P. F. J., Le Sager, P., Eskes, H. J., Alkemade, F., Scheele, M. P., Nedelec, P., and

Patz, H.-W.: The global chemistry transport model TM5: description and evaluation of the

tropospheric chemistry version 3.0, Geosci. Model Dev., 3, 445-473, 2010.

- 1448 Inness, A., Baier, F., Benedetti, A., Bouarar, I., Chabrillat, S., Clark, H., Clerbaux, C.,
- 1449 Coheur, P., Engelen, R. J., Errera, Q., Flemming, J., George, M., Granier, C., Hadji-Lazaro,
- 1450 J., Huijnen, V., Hurtmans, D., Jones, L., Kaiser, J. W., Kapsomenakis, J., Lefever, K., Leitão,
- 1451 J., Razinger, M., Richter, A., Schultz, M. G., Simmons, A. J., Suttie, M., Stein, O., Thépaut,
- 1452 J.-N., Thouret, V., Vrekoussis, M., Zerefos, C., and the MACC team: The MACC reanalysis:
- an 8 yr data set of atmospheric composition, Atmos. Chem. Phys., 13, 4073-4109,
 doi:10.5194/acp-13-4073-2013, 2013.
- Jaumouillé, E., Massart, S., Piacentini, A., Cariolle, D., and Peuch, V.-H.: Impact of a timedependent background error covariance matrix on air quality analysis, Geosci. Model
 Dev., 5, 1075-1090, 2012.
- Joly, M., and Peuch, V.-H.: Objective Classification of air quality monitoring sites overEurope, Atmos. Env., 47, 111-123, 2012.

- Jonson, J. E., Simpson, D., Fagerli, H., and Solberg, S.: Can we explain the trends in
 European ozone levels?, Atmos. Chem. Phys., 6, 51–66, doi:10.5194/acp-6-51-2006, 2006.
- Josse, B., Simon, P., and Peuch, V.-H.: Radon global simulations with the multiscale
 chemistry and transport model MOCAGE, Tellus B, 56, 339-356, 2004.
- Kahnert, M.: Variational data analysis of aerosol species in a regional CTM: background error
 covariance constraint and aerosol optical observation operators, Tellus, 60, 753–770,
 doi:10.1111/j.1600-0889.2008.00377.x, 2008.
- Kahnert, M.: On the observability of chemical and physical aerosol properties by optical
 observations: inverse modelling with variational data assimilation, Tellus B, 61, 747–755,
 doi:10.1111/j.1600-0889.2009.00436.x, 2009.
- 1470 Kaiser, J.W., Heil, A., Andreae, M. O, Benedetti, A., Chubarova, N., Jones, L., Morcrette, J.-
- 1471 J., Razinger, M., Schultz, M. G., Suttie, M. and van der Werf, G. R.: Biomass burning
- 1472 emissions estimated with a global fire assimilation system based on observed fire radiative
- 1473 power, Biogeosciences, 9, 527–554, doi:10.5194/bg-9-527-2012, 2012
- Kanakidou, M., Dameris, M., Elbern, H., Beekmann, M., Konovalov, I., Nieradzik, L.,
 Strunk, A., and Krol, M.: Synergistic use of retrieved trace constituents distributions and
 numerical modelling, in "The remote sensing of tropospheric composition from space", J.
 Burrows, U. Platt, P. Borrell (Eds.), Springer, doi 10.1007/978-3-642-14791-3, 2011.
- 1478 Kioutsioukis, I., and Galmarini, S.: *De praeceptis ferendis:* good practice in multi-model 1479 ensembles, Atmos. Chem. Phys., 14, 11791-11815, 2014.
- Köble, R., and Seufert, G.: Novel maps for forest tree species in Europe, in: Proceedings of
 the 8th European symposium on the physico-chemical behaviour of air pollutants: "a changing
 atmosphere, pp. 17–20, 2001.
- Kouznetsov, R. and Sofiev, M.: A methodology for evaluation of vertical dispersion and dry
 deposition of atmospheric aerosols, J. Geophys. Res., 117, D01202,
 doi:10.1029/2011JD016366, 2012.
- 1486 Kuenen, J. J. P., Denier van der Gon, H. A. C., Visschedijk, A., Van der Brugh, H., and Van
- Gijlswijk, R.: MACC European emission inventory for the years 2003-2007, TNO report
 TNO-060-UT-2011-00588, Utrecht, 2011.

- 1489 Kuenen, J. J. P., Visschedijk, A. J. H., Jozwicka, M., and Denier van der Gon, H. A. C.: TNO-
- MACC_II emission inventory: a multi-year (2003–2009) consistent high-resolution European
 emission inventory for air quality modelling, Atmos. Chem. Phys., 14, 10963-10976, 2014.
- 1492 Kukkonen, J., Olsson, T., Schultz, D. M., Baklanov, A., Klein, T., Miranda, A. I., Monteiro,
- 1493 A., Hirtl, M., Tarvainen, V., Boy, M., Peuch, V.-H., Poupkou, A., Kioutsioukis, I., Finardi, S.,
- 1494 Sofiev, M., Sokhi, R., Lehtinen, K. E. J., Karatzas, K., San José, R., Astitha, M., Kallos, G.,
- 1495 Schaap, M., Reimer, E., Jakobs, H., and K. Eben: A review of operational, regional-scale,
- 1496 chemical weather forecasting models in Europe, Atmos. Chem. Phys., 11, 1-87, 2012.
- Lacressonnière, G., Peuch, V.-H., Vautard, R., Arteta, J., Déqué, M., Josse, B., Marécal, V.,
 and Saint-Martin, D.: European air quality in the 2030s and 2050s: Impacts of global and
 regional emission trends and of climate change, Atmos. Env., 92, 348-358, 2014.
- 1500 Lahoz, W.A., Geer, A.J., Bekki, S., Bormann, N., Ceccherini, S., Elbern, H., Errera, Q.,
- 1501 Eskes, H.J., Fonteyn, D., Jackson, D.R., Khattatov, B., Massart, S., Peuch, V.-H., Rharmili,
- 1502 S., Ridolfi, M., Segers, A., Talagrand, O., Thornton, H.E., Vik, A.F., and Von Clarman T.:
- 1503 The Assimilation of Envisat data (ASSET) project, Atmos. Chem. Phys., 7, 1773-1796, 2007.
- Langner, J., Bergström, R., and Pleijel, K.: European scale modeling of sulphur, oxidized nitrogen and photochemical oxidants. Model dependent development av evaluation for the growing season. SMHI report, RMK No. 82, Swedish Met. And Hydrol. Inst., Norrköping, Sweden, 1998.
- Lefèvre, F., Brasseur, G. P., Folkins, I., Smith, A. K., and Simon, P.: Chemistry of the 1991–
 1992 stratospheric winter: Three-dimensional model simulations, J. Geophys. Res. :
 Atmospheres, 99, 8183–8195, 1994.
- 1511 Li, Y.P. Elbern, H., Liu, K.D., Friese, E., Kiendler-Scharr, A., Mentel, Th.F., and Wang, X.S.,
- 1512 Wahner, A., and Zhang, Y.H.: Updated aerosol module and its application to simulate
- secondary organic aerosols during IMPACT campaign May 2008, Atmos. Chem. Phys., 13,
- 1514 6289 6304, 2013, doi:10.5194/acp-13-6289-2013.
- Liu, D.C., and Nocedal, J.: On the limited memory BFGS method for large scale optimization,Math. Programming, 45, 503-528, 1989.
- 1517 Louis, J.-F.: A parametric model of vertical eddy fluxes in the atmosphere, Boundary-Layer
- 1518 Meteorology, 17, 187–202, 1979.

- Madronich, S., and Weller, G.: Numerical integration errors in calculated tropospheric
 photodissociation rate coefficients, J. Atmos. Chem., 10, 289-300, 1990.
- Mari, C., Jacob, D. J., and Bechtold, P.: Transport and scavenging of soluble gases in a deep
 convective cloud, J. Geophys. Res. : Atmospheres, 105, 22 255–22 267, 2000.
- 1523 Markakis, K., Giannaros, T., Poupkou, A., Liora, N., Melas, D., Sofiev, M. and Soares, J.:
- 1524 Evaluating the impact of particle emissions from natural sources in the Balkan region,
- 1525 European Aerosol Conference 2009, 6-9 September 2009, Karlsruhe, Germany, 2009.
- 1526 Markakis, K., Katragkou, E., Poupkou, A. and Melas, D.: "MOSESS: A new emission model
- 1527 for the compilation of model-ready emission inventories. Application in a coal mining area in
- 1528 Northern Greece", Environmental Modeling & Assessment, 18, pp. 509–521, 2013.
- Martet, M., Peuch, V.-H., Laurent, B., Marticorena B., and Bergametti, G.: evaluation of long-range transport and deposition of desert dust with the CTM Mocage, Tellus, 61B, 449-463, 2009.
- Massart, S., Clerbaux, C., Cariolle, D., Piacentini, A., Turquety, S., and Hadji-Lazaro, J.: First
 steps towards the assimilation of IASI ozone data into the MOCAGE-PALM system. Atmos.
 Chem. Phys., 9(14), 5073–5091. doi:10.5194/acp-9-5073-2009, 2009.
- Memmesheimer, M., Friese, E., Ebel, A., Jakobs, H. J., Feldmann, H., Kessler, C., and Piekorz, G.: Long-term simulations of particulate matter in Europe on different scales using sequential nesting of a regional model, Int. J. Environm. and Pollution, 22, (1-2), 108-132, 2004.
- Menut, L., and Bessagnet, B.: Atmospheric composition forecasting in Europe, Ann.Geophys. 28, 61-74, 2010.
- Menut, L., Goussebaile, A., Bessagnet, B., Khvorostiyanov, D., Ung, A.: Impact of realistic
 hourly emissions profiles on modelled air pollutants concentrations, Atmos. Env., pp. 233244, doi:10.1016/j.atmosenv.2011.11.057, 2012.
- 1544 Menut L, Bessagnet, B., Khvorostyanov, D., Beekmann, M., Blond, N., Colette, A., Coll, I.,
- 1545 Curci, G., Foret, G., Hodzic, A., Mailler, S., Meleux, F., Monge, J. L., Pison, I., Siour, G.,
- 1546 Turquety, S., Valari, M., Vautard, R., and Vivanco, M.G.: CHIMERE 2013 a model for
- regional atmospheric composition modelling, Geosci. Model Dev., 6, 981-1028,
 doi:10.5194/gmd-6-981-2013, 2013.

- Menut L., Perez Garcia-Pando, C., Haustein, K., Bessagnet, B., Prigent, C., and Alfaro, S.:
 Relative impact of roughness and soil texture on mineral dust emission fluxes modeling, J.
- 1551 Geophys. Res.: Atmospheres, 118, 6505-6520, doi:10.1002/jgrd.50313, 2013.
- 1552 Monteiro, A., Ribeiro, I., Tchepel, O., Sá, E., Ferreira, J., Carvalho, A., Martins, V., Strunk,
- 1553 A., Galmarini, S., Elbern, H., Schaap, M., Builtjes, P., Miranda, A. I., and Borrego, C.: Bias
- 1554 Correction Techniques to Improve Air Quality Ensemble Predictions: Focus on O3 and PM
- 1555 Over Portugal, Environ. Model. Assess., 18, 533-546, doi:10.1007/s10666-013-9358-2, 2013.
- 1556 Monteiro, A., Strunk, A., Carvalho, A., Tchepel, O., Miranda, A. I., Borrego, C., Saavedra, S.,
- Rodriguez, A., Souto, J., Casares, J., Friese, E., and Elbern, H.: Investigating a very high
 ozone episode in a rural mountain site, Env. Pol., 162, 176-189, 2012.
- 1559 Morcrette, J.-J., Boucher, O., Jones, L., Salmond, D., Bechtold, P., Beljaars, A., Benedetti,
- 1560 A., Bonet, A., Kaiser, J. W., Razinger, M., Schulz, M., Serrar, S., Simmons, A. J., Sofiev, M.,
- 1561 Suttie, M., Tompkins, A. M., and Untch, A.: Aerosol analysis and forecast in the European
- 1562 Centre for Medium-Range Weather Forecasts Integrated Forecast System. 1. Forward
- 1563 modelling: J. Geophys. Res., 114, D06206, doi:10.1029/2008JD011235, 2009.
- 1564 Navascues, B., Calvo, J., Morales, G., Santos, C., Callado, C., Cansado, A., Cuxart, J., Diez,
- 1565 M., del Rio, P., Escriba, P., Garcia-Colombo, O., García-Moya, J.A., Geijo, C., Gutierrez, E.,
- 1566 Hortal, M., Martinez, I., Orfila, B., Parodi, J.A., Rodriguez, E., Sánchez-Arriola, J., Santos-
- 1567 Atienza, I., Simarro, J.: Long term verification of HIRLAM and ECMWF forecasts over
- 1568 Southern Europe. History and perspectives of Numerical Weather Prediction at AEMET,
- 1569 Atmos. Res. 125-126, pp 20-33, doi:10.1016/j.atmosres.2013.01.010, 2013.
- Nho-Kim, E.-Y., Peuch, V.-H., and Oh, S. N.: Estimation of the global distribution of Black
 Carbon aerosols with MOCAGE, the CTM of Météo-France, J. Korean Meteor. Soc., 41(4),
 587-598, 2005.
- 1573 Nieradzik, L.P.: Application of a high dimensional model representation on the atmospheric
 1574 aerosol module MADE of the EURAD-CTM, Master Thesis, Institut für Geophysik und
 1575 Meteorologie der Universität zu Köln, 2005.
- 1576 Nocedal, J.: Updating quasi-Newton matrices with limited storage, Math. Comput., 35 (151),1577 773-782, 1980.
- Parrish, D. F., and Derber, J. C.: The national meteorological center's spectral statisticalinterpolation analysis system, Mon. Weather Rev., 120(8), 1747-1763, 1992.

- Petroff, A., and Zhang, L.: Development and application of a size-resolved particle dry
 deposition scheme for application in aerosol transport models, Geosci. Model Dev., 3, 753769, doi: 10.5197/gmd-3-753-2010.
- 1583 Poupkou, A., Giannaros, T., Markakis, K., Kioutsioukis, I., Curci, G., Melas, D., and Zerefos, C.: A model for European Biogenic Volatile Organic Compound emissions: Software 1584 1585 development and first validation. Environ. Model. Softw. 25, 1845–1856. doi:10.1016/j.envsoft.2010.05.004, 2010. 1586
- Potempski, S., Galmarini, S., Riccio, A., and Giunta, G.: Bayesian model averaging for
 emergency response atmospheric dispersion multimodel ensembles: Is it really better? How
 many data are needed? Are the weights portable?, J. Geophys. Res., 115, D21309,
 doi:10.1029/2010JD014210, 2010.
- Rabitz, H., Ö.F., Alis: General foundations of high-dimensional model representations, J.
 Math. Chem., 25, 197-233, 1999.
- Rao, S.T., Galmarini, S., and Puckett, K.: Air quality model evaluation international initiative
 (AQMEII). Bull. Am. Meteor. Soc., 92, 23-30. DOI:10.1175/2010BAMS3069.1, 2011.
- Riccio, A., Giunta, G., and Galmarini, S.: Seeking for the rational basis of the Median Model:
 the optimal combination of multi-model ensemble results, Atmos. Chem. Phys., 7, 6085–
 6098, doi:10.5194/acp-7-6085-2007, 2007.
- Robertson, L., Langner, J. and Engardt, M.: An Eulerian limited-area atmospheric transportmodel. J. Appl. Met. 38, 190-210, 1999.
- Roselle, S.J., and Binkowski, F.S.: Cloud Dynamics and Chemistry, in: Science algorithms of
 the EPA Models-3 Community multiscale air quality (CMAQ) modeling system, EPA 600/R99-030, EPA, 1999.
- 1603 Rouil, L., Honoré, C., Vautard, R., Beekmann, M., Bessagnet, B., Malherbe, L., Meleux, F.,
- Dufour, A., Elichegaray, C., Flaud, J.-M., Menut, L., Martin, D., Peuch, A., Peuch, V.-H., and
 Poisson, N.: PREV'AIR: an operational forecasting and mapping system for air quality in
- 1606 Europe, Bull. Am. Meteor. Soc., 90(1), 73-83, doi:10.1175/2008BAMS2390.1, 2009.
- Schaap, M., van Loon, M., ten Brink, H. M., Dentener, F. J., and Builtjes, P. J. H.: Secondary
 inorganic aerosol simulations for Europe with special attention to nitrate, Atmos. Chem.
 Phys., 4, 857–874, doi:10.5194/acp-4-857-2004, 2004.

- Schaap, M., Timmermans, R. M. A., Sauter, F. J., Roemer, M., Velders, G. J. M., Boersen, G.
 A. C., Beck, J. P., and Builtjes, P. J. H.: The LOTOS-EUROS model: description, validation
 and latest developments, Int. J. Environ. Pollut., 32, 270–289, 2008.
- Schaap, M., Manders, A.A.M, Hendriks, E.C.J., Cnossen, J.M., Segers, A.J.S., Denier van der
 Gon, H.A.C., Jozwicka, M., Sauter, F.J., Velders, G.J.M., Matthijsen, J., Builtjes P.J.H.:
 Regional modelling of particulate matter for Netherlands' published by the Neterlands
 Research Programme on particulate matter. Report 500099008, ISSN: 1875-2322 (print)
- 1617 ISSN: 1875-2314. 2005.
- Sandu, A., Daescu, D. N., and Carmichael, G.R.: Direct and adjoint sensitivity analysis of
 chemical kinetic systems with KPP: part I theory and software Tools, Atmos. Env., 37,
 5083-5096, 2003.
- Sandu, A., and Sander, R.: Technical node: Simulating chemical systems in Fortran90 andMatlab with the Jinetic PreProcessor KPP-2.1, Atmos. Chem. Phys., 6, 187-195, 2006.
- Sič, B., El Amraoui, L., Marécal, V., Josse, B., Arteta, J., Guth, J., Joly, M., and Hamer, P.
 D.: Modelling of primary aerosols in the chemical transport model MOCAGE: development
 and evaluation of aerosol physical parameterizations, Geosci. Model Dev., 8, 381–408,
 doi:10.5194/gmd-8-381-2015, 2015.
- Simpson, D., Andersson-Sköld, Y., and Jenkin, M.E.: Updating the chemical scheme for the
 EMEP MSC-W oxidant model: current status. EMEP MSC-W Nore 2/93, 1993.
- 1629 Simpson, D., Winiwarter, W., Börjesson, G., Cinderby, S., Ferreiro, A., Guenther, A., Hewitt,
- 1630 C. N., Janson, R., Khalil, M. A. K., Owen, S., Pierce, T. T., Puxbaum, H., Shearer, M., Skiba,
- 1631 U., Steinbrecher, R., Tarrason, L., and Oquist, M. G.: Inventorying emissions from nature in
- 1632 Europe, J. Geophys. Res.: Atmospheres (1984–2012), 104, 8113–8152, 1999.
- Simpson, D., Fagerli, H., Jonson, J. E., Tsyro, S., Wind, P., and Tuovinen, J.-P.:
 Transboundary Acidification, Eutrophication and Ground Level Ozone in Europe, Part 1:
 Unified EMEP Model Description, EMEP Report 1/2003, 2003.
- 1636 Simpson, D., Benedictow, A., Berge, H., Bergström, R., Emberson, L. D., Fagerli, H.,
- 1637 Flechard, C. R., Hayman, G. D., Gauss, M., Jonson, J. E., Jenkin, M. E., Nyiri, A., Richter,
- 1638 C., Semeena, V. S., Tsyro, S., Tuovinen, J.-P., Valdebenito, A., and Wind, P.: The EMEP
- 1639 MSC-W chemical transport model technical description, Atmos. Chem. Phys., 12, 7825–
- 1640 7865, 2012.

- 1641 Skamarock, W.C., Klemp, J.B., Dudhia, J., Gill, D.O., Barker, D.M., Duda, M.G., Huang,
 1642 X.Y., Wang, W., Powers, J.G.: A description of the advanced researcher WRF version 3.
 1643 NCAR Technical Note, NCAR/TN-475+STR, June 2008, Boulder, Colorado, USA, 125 pp,
 1644 2008.
- 1645 Sofiev, M.: A model for the evaluation of long-term airborne pollution transport at regional 1646 and continental scales. Atmos. Environ. 34, 2481–2493, 2000.
- Sofiev, M.: Extended resistance analogy for construction of the vertical diffusion scheme fordispersion models. J. Geophys. Res. 107. doi:10.1029/2001JD001233, 2002.
- Sofiev, M., Galperin, M. V, and Genikhovich, E.: Construction and evaluation of Eulerian
 dynamic core for the air quality and emergency modeling system SILAM, in: Borrego, C.,
 Miranda, A.I. (Eds.), NATO Science for Piece and Security Serties C: Environmental
 Security. Air Pollution Modelling and Its Application, XIX. SPRINGER-VERLAG BERLIN,
 pp. 699–701, 2008.
- Sofiev, M., Genikhovich, E., Keronen, P., and Vesala, T.: Diagnosing the Surface Layer
 Parameters for Dispersion Models within the Meteorological-to-Dispersion Modeling
 Interface. J. Appl. Meteorol. Climatol. 49, 221–233. doi:10.1175/2009JAMC2210.1, 2010.
- Sofiev, M., Siljamo, P., Valkama, I., Ilvonen, M., and Kukkonen, J.: A dispersion modelling
 system SILAM and its evaluation against ETEX data. Atmos. Environ. 40, 674–685.
 doi:10.1016/j.atmosenv.2005.09.069, 2006.
- Sofiev, M., Soares, J., Prank, M., de Leeuw, G., Kukkonen, J.: A regional-to-global model of
 emission and transport of sea salt particles in the atmosphere. J. Geophys. Res. 116.
 doi:10.1029/2010JD014713, 2011.
- Sofiev. M., Berger, U., Prank, M., Vira, J., Arteta, J., Belmonte, J., Bergmann, K.-C. ,
 Cheroux, F., Elbern, H., Friese, E., Galan, C., Gehrig, R., Kranenburg, R., Marécal, V.,
- 1665 Meleux, F., Pessi, A.-M., Robertson, L., Ritenberga, O., Rodinkova, V., Saarto, A., Segers,
- 1666 A., Severova, E., Sauliene, I., Steensen, B. M., Teinemaa, E., Thibaudon, M., and Peuch, V.-
- 1667 H.: Multi-model simulations of birch pollen in Europe by MACC regional ensemble, Atmos.
- 1668 Chem. Phys., 15, 8115-8130, 2015.
- 1669 Solazzo, E., Bianconi, R., Vautard, R., Appel, K. W., Moran, M. D., Hogrefe, C., Bessagnet,
- 1670 B., Brandt, J., Christensen, J. H., Chemel, C., Coll, I., Denier van der Gon, H., Ferreira, J.,
- 1671 Forkel, R., Francis, X. V., Grell, 5 G., Grossi, P., Hansen, A. B., Jericevic, A., Kraljevic, L.,

- 1672 Miranda, A. I., Nopmongcol, U., Pirovano, G., Prank, M., Riccio, A., Sartelet, K. N., Schaap,
- 1673 M., Silver, J. D., Sokhi, R. S., Vira, J., Werhahn, J., Wolke, R., Yarwood, G., Zhang, J., Rao,
- 1674 S. T., and Galmarini, S.: Model eva- lution and ensemble modelling of surface-level ozone in
- 1675 Europe and North America in the context of AQMEII, Atmos. Env., 53, 60–74, 2012a.
- 1676 Solazzo, E., Bianconi, R., Pirovano, G., Matthias, V., Vautard, R., Moran, M. D., Wyat
- 1677 Appel, K., Bessagnet, B., Brandt, J., Chris- tensen, J. H., Chemel, C., Coll, I., Ferreira, J.,
- 1678 Forkel, R., Francis, X. V., Grell, G., Grossi, P., Hansen, A. B., Miranda, A. I., Nop-mongcol,
- 1679 U., Prank, M., Sartelet, K. N., Schaap, M., Silver, J. D., Sokhi, R. S., Vira, J., Werhahn, J.,
- 1680 Wolke, R., Yarwood, G., Zhang, J., Rao, S. T., and Galmarini, S.: Operation model evalu-
- ation for particulate matter in Europe and North America in the context of AQMEII, Atmos.
 Env., 53, 75–92, 2012b.
- Stein, O., Flemming, J., Inness, A., Kaiser, J. W., and Schultz, M. G.: Global reactive gases
 and reanalysis in the MACC project, Journal of Integrative Environmental Sciences, 9, Iss.
 sup1, 57–70, doi:10.1080/1943815X.2012.696545, 2012.
- Stern, R., Builtjes, P., Schaap, M., Timmermans, R. M. A., Vautard, R., Hodzic, A.,
 Memmesheimer, M., Feldmann, H., Renner, E., Wolke, R., and Kerschbaumer, A.: A model
 inter-comparison study focussing on episodes with elevated PM10 concentrations, Atmos.
 Env., 42, 4567-4588, 10.1016/j.atmosenv.2008.01.068, 2008.
- 1690 Stockwell, W. R., Kirchner, F., Kuhn, M., and Seefeld, S.: A new mechanism for regional 1691 atmospheric chemistry modeling, J. Geophys. Res. : Atmospheres, 102, 25 847–25 879, 1997.
- Tie, X., Madronich, S., Walters, S., Zhang, R., Rasch, P., and Collins, W.: Effect of clouds on
 photolysis and oxidants in the troposphere, J. Geophys. Res., 108 (D20), 4642, doi:
 10.1029/2003JD003659, 2003.
- Timmermans, R.M.A., Schaap, M., Elbern, H., Siddans, R., Tjemkes, S.A.T., Vautard, R., and
 Builtjes, P.J.H.: An Observing System Simulation Experiment (OSSE) for Aerosol Optical
 Depth from satellites, J. Atmos. Ocean. Technol., 26, 2673-2682, 2009.
- Tuovinen, J.-P., Ashmore, M., Emberson, L., and Simpson, D.: Testing and improving the
 EMEP ozone deposition module, Atmos. Env., 38, 2373–2385, 2004.
- 1700 Turquety, S., Menut, L., Bessagnet, B., Anav, A., Viovy, N., Maignan, F., and Wooste, M.:
- APIFLAME v1.0: high-resolution fire emission model and application to the Euro
 Mediterranean region, Geosci. Model Dev., 7, 587–612, 2014.

- 1703 Van Loon, M., Vautard, R., Schaap, M., Bergstrom, R., Bessagnet, B., Brandtn J., Builtjes,
- 1704 P.J.H, Christensen, J.H., Cuvelier, C., Graff, A., Jonson, J.E., Krol, M., Langner, J., Roberts,
- 1705 P., Rouil, L., Stern, R., Tarrason, L., Thunis, P., Vignati, E., and White, L.: Evaluation of
- 1706 long-term ozone simulations from seven regional air quality models and their ensemble,
- 1707 Atmos. Env., 41, 2083-2097, 10.1016/j.atmosenv.2006.10.073, 2007.
- 1708 Van Ulden, A.P, and Holtslag, A.A.M.: Estimation of atmospheric boundary layer parameters
 1709 fort diffusion applications, J. Climate. Appl. Met., 24, 1196-1207, 1975.
- Vautard, R., Builtjes, P.H.J., Thunis, P., C. Cuvelier, Bedogni, M., Bessagnet, B., Honoré, C.,
 Moussiopoulos, N., Pirovano, G., Schaap, M., Stern, R., Tarrason, L., and Wind,
 P.:Evaluation and intercomparison of Ozone and PM10 simulations by several chemistry
 transport models over four European cities within the CityDelta project, Atmos. Env., 41,
 1714 173–188, 2007.
- 1715 Vira, J., and Sofiev, M.:On variational data assimilation for estimating the model initial
 1716 conditions and emission fluxes for short-term forecasting of SOx concentrations. Atmos.
 1717 Env., 46, 318–328. doi:10.1016/j.atmosenv.2011.09.066, 2012.
- 1718 Vira, J. and Sofiev, M.: Assimilation of surface NO2 and O3 observations into the SILAM
 1719 chemistry transport model, Geosci. Model Dev., 8, 191–203, doi:10.5194/gmd-8-191-2015,
 1720 2015.
- Visschedijk, A. J. H., Zandveld, P. Y. J., and Denier van der Gon, H. A. C. A.: High
 resolution gridded European database for the EU Integrate Project GEMS, TNO-report 2007A-R0233/B, 2007.
- Walcek, C.J.: Minor flux adjustment near mixing ratio extremes for simplified yet highly
 accurate monotonic calculation of tracer advection, J. Geophy. Res., 105(D7), 9335-9348,
 2000.
- Wichink Kruit, R. J., Schaap, M., Sauter, F. J., van Zanten, M. C., and van Pul, W. A. J.:
 Modeling the distribution of ammonia across Europe including bi-directional surface–
 atmosphere ex- change, Biogeosciences, 9, 5261–5277, doi:10.5194/bg-9-5261-2012, 2012.
- Vlemmix, T., Eskes, H. J., Piters, A. J. M., Schaap, M., Sauter, F. J., Kelder, H., and Levelt,
 P. F., MAX-DOAS tropospheric nitrogen dioxide column measurements compared with the
 Lotos-Euros air quality model, Atmos. Chem. Phys., 15, 1313-1330, doi:10.5194/acp-151313-2015, 2015.

- Weaver, A., and Courtier, P.: Correlation modeling on the sphere using a generalizeddiffusion equation, Q. J. R. Meteorol. Soc., 127, 1815-1846, 2001.
- Wesely, M.: Parameterization of surface resistances to gaseous dry deposition in regionalscale numerical models, Atmos. Env. (1967), 23, 1293 1304, 1989.
- 1738 Williams, J. E., van Velthoven, P. F. J., and Brenninkmeijer, C. A. M.: Quantifying the
- 1739 uncertainty in simulating global tropospheric composition due to the variability in global
- 1740 emission estimates of Biogenic Volatile Organic Compounds, Atmos. Chem. Phys., 13, 2857-
- 1741 2891, doi:10.5194/acp-13-2857-2013, 2013.
- Williamson, D.L., and Rasch, R. P.: Two-Dimensional Semi-Lagrangian Transport with
 Shape-Preserving Interpolation, American Meteorological Society, 117, 102–129, 1989.
- WHO (World Health Organization): Health aspects of air pollution results from the WHO
 project "Systematic review of health aspects of air pollution in Europe", Technical Report,
 2004.
- WHO (World Health Organization): Review of evidence on health aspects of air pollution –
 REVIHAAP Project, Technical Report, 2013.
- Yarwood, G., Rao, S., Yocke, M., and Whitten, G.Z.: Updates to the carbon bond chemical
 mechanism: CB05, Report to the U.S. Environmental Protection Agency, RT-04-00675,
 Yocke and Company, Novato, California, United States, 2005.
- Zhang, L., Brook, J.R., and Vet, R.: A revised parameterization for gaseous dry deposition in
 air-quality models, Atmos. Chem. Phys., 3, 2067-2082, 2003.
- Zhang, L, Gong, S., Padro, J., and Barrie, L.: A size-segregated particle dry deposition
 scheme for an atmospheric aerosol module. Atmos. Env., 35(3), 2001, 549–560, 2001.
- Zhang, Y., Bocquet, M., Mallet, V., Seigneur, C., and Baklanov, A.: Real-time air quality
 forecasting, part I: History, techniques, and current status, Atmos. Environ., 60, 632–655,
 2012.
- Zilitinkevich, S. and Mornom, D.V.: A multi-limit formulation for the equilibrium depth of a
 stable stratified boundary layer. Max-Planck-Institute for Meteorology. Report No. 185, ISSN
 0397-1060, 30 pp., 1996.
- 1762 Zyryanov, D., Foret, G., Eremenko, M., Beekmann, M., Cammas, J.-P., D'Isidoro, M., Elbern,
- 1763 H., Flemming, J., Friese, E., Kioutsioutkis, I., Maurizi, A., Melas, D., Meleux, F., Menut, L.,

- 1764 Moinat, P., Peuch, V.-H., Poupkou, A., Razinger, M., Schultz, M., Stein, O., Suttie, A. M.,
- 1765 Valdebenito, A., Zerefos, C., Dufour, G., Bergametti, G., and Flaud, J.-M.: 3-D evaluation of
- 1766 tropospheric ozone simulations by an ensemble of regional Chemistry Transport Models,
- 1767 Atmos. Chem. Phys., 12, 3219-3240, 2012, doi:10.5194/acp-12-3219-2012.

Table 1. Portfolio of the MACC-II regional data products. Each product is provided once 1770 daily. Core species correspond to O₃, NO₂, CO, SO₂, PM10, PM2.5. Additional species 1771 correspond to NO, NH₃, PAN+PAN precursors, total Non-Methane Volatile Organic 1772 Compounds. Birch pollen concentrations are only available during season from 1st of March 1773 to 30th of June each year. Old levels refer to surface, 500m, 1000m, 3000m and 5000m, 1774 corresponding to the production before mid-May 2014. All levels refers to surface, 50m, 1775 1776 250m, 500m, 1000m, 2000m, 3000m and 5000m, produced from mid-May 2014. The analysis is run a posteriori on Day0 for Day-1 (00 to 24UTC). 1777

Model name	Forecast or	Species	Time span	Vertical	Format
	Analysis			levels	
CHIMERE	Forecast	core +	0h to 96h,	All levels	Netcdf
		additional	hourly		
CHIMERE	Forecast	Birch pollen	0h to 96h,	Surface	Netcdf
			hourly		
CHIMERE	Analysis	O ₃ , PM10	-24h to -1h,	Surface	Netcdf
			hourly		
EMEP	Forecast	core +	0h to 96h,	All levels	Netcdf
		additional	hourly		
EMEP	Forecast	Birch pollen	0h to 96h,	Surface	Netcdf
			hourly		
EMEP	Analysis	NO ₂	-24h to -1h,	Surface	Netcdf
			hourly		
EURAD-IM	Forecast	core +	0h to 96h,	All levels	Netcdf
		additional	hourly		
EURAD-IM	Forecast	Birch pollen	0h to 96h,	Surface	Netcdf
			hourly		
EURAD-IM	Analysis	O ₃ , NO ₂ , CO,	-24h to -1h,	Surface	Netcdf
		SO ₂ , PM10	hourly		
LOTOS-	Forecast	core + NO	0h to 96h,	Old levels	Netcdf
EUROS			hourly		

LOTOS- EUROS	Forecast	Birch pollen	0h to 96h, hourly	Surface	Netcdf
LOTOS- EUROS	Analysis	O ₃	-24h to -1h, hourly	Surface	Netcdf
MATCH	Forecast	core + additional	0h to 96h, hourly	All levels	Netcdf
MATCH	Forecast	Birch pollen	0h to 96h, hourly	Surface	Netcdf
MATCH	Analysis	O ₃ , NO ₂ , CO,PM10, PM2.5	-24h to -1h, hourly	Surface	Netcdf
MOCAGE	Forecast	core + additional (except NH ₃)	0h to 96h, hourly	All levels	Netcdf
MOCAGE	Forecast	Birch pollen	0h to 96h, hourly	Surface	Netcdf
MOCAGE	Analysis	O ₃	-24h to -1h, hourly	Surface	Netcdf
SILAM	Forecast	core	0h to 96h, hourly	All levels	Netcdf
SILAM	Forecast	Birch pollen	0h to 96h, hourly	Surface	Netcdf
SILAM	Analysis	O ₃ , NO ₂ , SO ₂	-24h to -1h, hourly	Surface	Netcdf
ENSEMBLE	Forecast	core + additional	0h to 96h, hourly	All levels	Netcdf + Grib2
ENSEMBLE	Forecast	Birch pollen	0h to 96h, hourly	Surface	Netcdf + Grib2
ENSEMBLE	Analysis	O ₃	-24h to -1h,	Surface	Netcdf +

	hourly	Grib2
1778		
1779		
1780		

	Forecast Day0 (0h-24h)	Forecast Day1 (25h-48h)	Forecast Day2 (49h-72h)	Forecast Day3 (73h-96h)	Analysis (-24h-0h)
Core species	07 UTC	07 UTC	08 UTC	09 UTC	14:30 UTC
Additional species	07 UTC	07 UTC	08 UTC	09 UTC	N/A

Table 2. Time of delivery of the ENSEMBLE numerical products. Core species for theanalysis is restricted to ozone only.

Model	Operated by	Horizontal	Vertical levels
		resolution	Top height
CHIMERE	INERIS (Institut National de	0.1° x 0.1°	8 levels
	l'Environnement Industriel et des		Top at 500 hPa
	Risques)		
	France		
EMEP	MET Norway (Meteorologisk institutt)	0.25° x 0.125°	20 levels
	Norway		top at 100 hPa
EURAD-IM	RIUUK (Rheinisches Institut Fuer	15 km on a	23 levels
	Umwelt-Forschung an der Universitaet	Lambert	Top at 100 hPa
	zu Koeln E. V.)	conformal	
	Germany	projection	
LOTOS-	KNMI (Koninklijk Nederlands	0.25° x 0.125°	4 levels
EUROS	Meteorologisch Instituut)		Top at 3.5km
	Netherlands		
MATCH	SMHI (Sveriges Meteorologiska och	$0.2^{\circ} \ge 0.2^{\circ}$	52 levels
	Hydrologiska Institut)		Top at 300 hPa
	Sweden		
MOCAGE	Météo-France	0.2° x 0.2°	47 levels
	France		Top at 5 hPa
SILAM	FMI (Ilmatieteen Laitos)	0.15° x 0.15°	8 levels
	Finland		Top at 6.7 km

1785 Table 3. General characteristics of the regional models at the end of MACC-II project.

Table 4. Characteristics of the daily assimilation chains of the regional models at the end ofMACC-II project.

Model	Assimilation method Observation assimil		Species analysed
CHIMERE	Optimal interpolation	O_3 and PM_{10} from surface stations,	O ₃ , PM ₁₀
EMEP	3D- Variational	NO2columnsfromOMIandNO2fromsurface stationssurface stationssurface stations	NO ₂
EURAD-IM	3D- Variational	$\begin{array}{ccc} O_3, NO, NO_2, SO_2, CO, \\ PM_{10}, & PM_{2.5} & from \\ surface stations, & OMI \\ and & GOME-2 & NO_2 \\ column & retrivals, \\ MOPITT CO profiles \end{array}$	O ₃ , NO ₂ , SO ₂ , CO, PM ₁₀
LOTOS-EUROS	Ensemble Kalman filter	O ₃ from surface stations	O ₃
МАТСН	3D- Variational	O ₃ , NO ₂ , CO, PM ₁₀ , PM _{2.5} from surface stations	O ₃ , NO ₂ , CO, PM ₁₀ , PM _{2.5}
MOCAGE	3D- Variational	O ₃ from surface stations	O ₃
SILAM 4D-Variational		O ₃ , NO ₂ and SO ₂ from surface stations	O ₃ , NO ₂ , SO ₂



Figure 1. Schematic of the general organisation of the MACC-II air quality forecast and analysis system.



Figure 2. Zoomed map of ozone concentrations at the surface in $\mu g m^{-3}$ of the 15h forecast for the 10th of June 2014 at 15 UTC of the ensemble median constructed with the 7 model forecasts. NRT AQ observations available (circles) for the same date/time are overpoltted on the maps using the same colour scale.



1805

Figure 3. Left panel: Ozone measurements from surface stations in µg m⁻³ from the 10th of 1806 June 2014 at 00UTC to the 14th of June 2014 at 00 UTC located (a) at 47.67°N/13.11°E 1807 (Hallein, Austria), (b) at 47.69°N/16.58°E (Sopron, Hungaria), (c) at 50.13°N/8.75°E 1808 1809 (Fechenheim, Germany), (d) at 43.33°N/5.12°E (Sausset, France), (e) at 43.34°N/5.73°E (Plan d'Aups, France) and (f) at 43.79°N/4.83°E (St Rémi, France). Right panel: EPSgrams 1810 giving median, 90% percentile, 75% percentile, 25% percentile, 10% percentile, minimum 1811 and maximum from 3-hourly outputs of the 96h forecasts of the 7 models from the 10th of 1812 June 2014 at 00UTC to the 14th of June 2014 at 00 UTC. Model outputs are interpolated at the 1813 location of the stations shown in the left panel. The red dashed line corresponds to 120 µg m⁻ 1814 3 1815





1818

Figure 4. Statistical indicators (see Appendix) for ozone as a function of the forecast time in hour for the ensemble median (in turquoise) and the seven models (other colours) compared to the hourly surface station measurements available for the period from the 9th to the 15th of June 2014 over the MACC-II European domain. (a) MB in μ g m⁻³, (b) MNMB, (c) RMSE in μ g m⁻³, (d) FGE and (e) correlation.



Figure 5. Statistical indicators (see Appendix) for ozone as a function of the forecast time in hour MEDIAN 7, MEDIAN 5, MEDIAN 3 and 1BEST (see text for their definition) compared to the hourly surface station measurements available for the period from the 9th to the 15th of June 2014 over the MACC-II European domain. (a) MB in μ g m⁻³, (b) MNMB, (c) RMSE in μ g m⁻³, (d) FGE and (e) correlation.



1833

Figure 6. Statistical indicators (see Appendix) for ozone as a function of the forecast time in hour for the ensemble median compared to the hourly surface station measurements available for the period from the 1st of June at 00UTC to the 1st of September at 00UTC over the MACC-II European domain for 2014: (a) MB in μ g m⁻³, (b) MNMB, (c) RMSE in μ g m⁻³, (d) FGE and (e) correlation.

- 1839
- 1840



1841

Figure 7. Statistical indicators (see Appendix) for ozone as a function of the forecast time in hour for an ensemble of 7, 6, 5, 4 and 3 models compared to the hourly surface station measurements available for the period from the 1st of June 2014 at 00UTC to the 1st of September 2014 at 00UTC over the MACC-II European domain: (a) MB in μ g m⁻³, (b) MNMB, (c) RMSE in μ g m⁻ 1846 ³, (d) FGE and (e) correlation.



1849

Figure 8. Statistical indicators (see Appendix) for PM_{10} as a function of the forecast time in hour for the ensemble median compared to the hourly surface station measurements available for the period from the 1st of December 2013 at 00UTC to the 1st of March 2014 at 00UTC over the MACC-II European domain: (a) MB in μ g m⁻³, (b) MNMB, (c) RMSE in μ g m⁻³, (d) FGE and (e) correlation.

Surface Ozone Concentrations



1856

1857

Figure 9. ENSEMBLE (left) and AEMET (right) surface ozone concentrations in μ g m⁻³ from a forecast (H+18) started on 18th July 2013 at 00UTC for the Western Mediterranean area. Observations from different air quality networks have been plotted on the map. The Madrid and the Eastern Spain areas appear magnified.




Figure 10. Ozone concentrations in µg m-3 as a function of days at Cap de Creus (42.32°N,
3.32°W). (Top panel) for June 2013, (middle panel) for July 2013 and (bottom panel) for
August 2013. The blue colour line is for EMEP observations, the red colour line is for the
ENSEMBLE and the green colour line for the AEMET model.



1872 Figure 11. As figure 10 but for Mahon (39.867°N, 4.32°W).



1875 Figure 12. As figure 12 but for San Pablo de los Montes (39.55°N, 4.35°E).



Figure 13. Ozone concentrations (in µg m⁻³) from a 48h forecast of ENSEMBLE, AEMET and the 7 individual models at ES10 EMEP air quality station which is located at Cabo de Creus in the Northeastern corner of Spain (42.32°N, 3.32°E). The forecast is started on the 9th of April 2014. CHI, EMP, KNM, FMI, MFM, RIU, SMH, ENS, MACCH3 and OBS correspond to CHIMERE, EMEP, LOTOS-EUROS, SILAM, MOCAGE, EURAD-IM, MATCH, ENSEMBLE, AEMET models and observations, respectively.