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

V. Marécal[1], V.-H. Peuch[2], C. Andersson[3], S. Andersson[3], J. Arteta[1], M. Beekmann[4], A. Benedictow[5], R. Bergström[3], B. Bessagnet[6], A. Cansado[7], F. Chéroux[1], A. Colette[6], A. Coman[4], R. L. Curier[8], H. A. C. Denier van der Gon[8], A. Drouin[1], H. Elbern[9], E. Emili[10], R. J. Engelen[2], H. J. Eskes[11], G. Foret[4], E. Friese[9], M. Gauss[5], C. Giannaros[12], J. Guth[1], M. Joly[1], E. Jaumouillé[10], B. Josse[1], N. Kadygrov[1], J. W. Kaiser[13], K. Krajsek[14], J. Kuenen[8], U. Kumar[11], N. Liora[12], E. Lopez[7], L. Malherbe[6], I. Martinez[7], D. Melas[12], F. Meleux[6], L. Menut[15], P. Moinat[10], T. Morales[7], J. Parmentier[1], A. Piacentini[10], M. Plu[1], A. Poupkou[12], S. Queguiner[1], L. Robertson[3], L. Rouïl[6], M. Schaap[8], A. Segers[8], M. Sofiev[16], L. Tarasson[17], M. Thomas[3], R. Timmermans[8], Á. Valdebenito[5], P. van Velthoven[11], R. van Versendaal[11], J. Vira[16], and A. Ung[6]

[1]Groupe d'étude de l'Atmosphère Méréorologique/Centre National de Recherches Météorologiques, CNRS-Météo-France, UMR 3589, Toulouse, France
[2]European Centre for Medium-range Weather Forecasts, Reading, UK
[3]Swedish Meteorological and Hydrological Institute, Norrköping, Sweden
[4]Laboratoire Inter-universitaire des Systèmes Atmosphériques, UMR CNRS 7583, Université Paris Est Créteil et Université Paris Diderot, Créteil, France
[5]Norwegian Meteorological Institute, Oslo, Norway
[6]Institut National de l'Environnement Industriel et des Risques, Parc Technologique Alata, 60550 Verneuil en Halatte, France
[7]AEMET Spanish Meteorological State Agency, Leonardo Prieto Castro 8, Madrid 28040, Spain
[8]TNO, Climate Air and Sustainability Unit, Utrecht, the Netherlands
[9]Rhenish Institute for Environmental Research at the University of Cologne, Cologne, Germany
[10]CERFACS, URA 1875, Toulouse, France
[11]Royal Netherlands Meteorological Institute, De Bilt, the Netherlands
[12]Laboratory of Atmospheric Physics, Physics Dept., Aristotle University of Thessaloniki, Thessaloniki, Greece
[13]Max Planck Institute for Chemistry, Mainz, Germany
[14]Institut für Energie- und Klimaforschung (IEK-8), Forschungszentrum Jülich, Jülich, Germany
[15]Laboratoire de Météorologie Dynamique, Ecole Polytechnique, 91128 Palaiseau, France
[16]Finnish Meteorological Institute, Erik Palmenin Aukio 1, Helsinki 00560, Finland
[17]Norwegian Institute for Air Research, 2027 Kjeller, Norway

*Correspondence to:* V. Marécal (virginie.marecal@meteo.fr)

**Abstract.** This paper describes the pre-operational analysis and forecasting system developed during MACC (Monitoring Atmospheric Composition and Climate) and continued in the MACC-II (Monitoring Atmospheric Composition and Climate: Interim Implementation) European 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, 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 MACC-II (summer 2014) and analyses the performance of the multi-model ensemble. The MACC-II system provides daily 96 h forecasts with hourly outputs of 10 chemical species/aerosols ($O_3$, $NO_2$, $SO_2$, CO, $PM_{10}$, $PM_{2.5}$, NO, $NH_3$, total NMVOCs (non-methane volatile organic compounds) and PAN+PAN

precursors) over eight vertical levels from the surface to 5 km height. The hourly analysis at the surface is done a posteriori for the past day using a selection of representative air quality data from European monitoring stations.

The performance of the system is assessed daily, weekly and every 3 months (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, $NO_2$ and $PM_{10}$. The statistical indicators for ozone in summer 2014 show that the ensemble median gives on average the best performances compared to the seven models. There is very little degradation of the scores with the forecast day but there is a marked diurnal cycle, similarly to the individual models, that can be related partly to the prescribed diurnal variations of anthropogenic emissions in the models. During summer 2014, the diurnal ozone maximum is underestimated by the ensemble median by about $4\,\mu g\,m^{-3}$ on average. Locally, during the studied ozone episodes, the maxima from the ensemble median are often lower than observations by $30$–$50\,\mu g\,m^{-3}$. Overall, ozone scores are generally good with average values for the normalised indicators of 0.14 for the modified normalised mean bias 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 model. Scores are also discussed for $PM_{10}$ for winter 2013–1014. There is an underestimation of most models leading the ensemble median to a mean bias of $-4.5\,\mu g\,m^{-3}$. The ensemble median fractional gross error is larger for $PM_{10}$ ($\sim 0.52$) than for ozone and the correlation is lower ($\sim 0.35$ for $PM_{10}$ and $\sim 0.54$ for ozone). This is related to a larger spread of the seven model scores for $PM_{10}$ than for ozone linked to different levels of complexity of aerosol representation in the individual models. In parallel, a scientific analysis of the results of the seven models and of the ensemble is also done over the Mediterranean area because of the specificity of its meteorology and emissions.

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

# 1 Introduction

The chemical composition of the air close to Earth's surface, generally referred as "air quality" (AQ), directly affects human and animal health and also the vegetation. For instance, ozone has a known impact on the respiratory system (e.g. WHO, 2004) and on the vegetation development (e.g. Fuhrer and Booker, 2003). Recently, the World Health Organization reported that in 2012 around 3.7 million deaths were attributable to ambient air pollution (http://www.who.int/phe/health_topics/outdoorair/databases/en/). This is why air quality has become a major concern, starting in the 1970s, particularly in Europe (e.g. WHO, 2013). Since the Helsinki Protocol in 1985, many regions and countries, including the European Union countries, have progressively put in place tools to regulate and to control the emissions of the main air pollutants. This has led to an important effort to monitor the air composition near the surface but also to develop air quality forecasting systems in experimental or operational modes (see reviews by Ebel et al., 2005; Menut and Bessagnet, 2010). These tools can be used in cases of high pollution episodes to inform people and to take emergency measures to prevent harming effects. They can also be used for policy makers for the regulations on air pollutant emissions and for monitoring the effect of these regulations on air quality (episodes and also background pollution).

The main pollutants under focus for air quality are ozone, nitrogen oxides ($NO_x = NO_2 + NO$), sulfur dioxide ($SO_2$), volatile organic compounds (VOCs), ammonia ($NH_3$), particulate matter, heavy metals (Pb, Cd, Hg) and persistent organic pollutants (POPs, e.g. pesticides and dioxin). Ozone is a secondary pollutant, meaning that it is not emitted but produced from gaseous precursors (mainly VOCs and $NO_x$) originating from both natural and anthropogenic sources. Particulate matter (PM) corresponds to small size aerosols. PM is categorised as $PM_{10}$ (size $< 10\,\mu m$), $PM_{2.5}$ (size $< 2.5\,\mu m$) and $PM_1$ (size $< 1\,\mu m$). These categories were chosen because of their known effects on health. In PM, the distinction between primary (dust, sea salts, black carbon and organic carbon) and secondary aerosols formed from gaseous precursors such as $SO_2$, DMS (dimethyl sulfate), $H_2S$, $NH_3$, $NO_x$ and VOCs is ignored when considering mass or number concentration only.

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 forecasting purposes. Regional and local air quality forecasting systems (Kukkonen et al., 2012; Zhang et al., 2012) rely on limited area models that can be based either on an off-line or an on-line approach to take into account the effect of meteorological conditions on air composition. Offline chemistry models, known as chemistry-transport models (CTMs), use the meteorological parameters from the analyses or the forecasts provided by a separate numerical weather prediction model. On-line models are meteorological models in which chemical variables and processes are included (Baklanov et al., 2014). On-line models have the capability to represent the feedback of the chemical composition on meteorological parameters but they are computationally demanding by design. This is why CTMs are generally preferred for operational air quality forecasting systems.

The chemical composition of air depends on many processes that need to be well represented in models in

order to provide reliable air quality forecasts (e.g. Rao et al., 2011). The 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 pollutants, by the advection by winds, by the diffusion in the planetary boundary layer, by the scavenging by rain and by the dry deposition at the surface. Each of these processes has its own uncertainty. These uncertainties come, on the one hand, from the limit of our current knowledge and, on the other hand, from the need to simplify the process representation in models because of computational constraints. In meteorology and climate studies, and more recently in atmospheric dispersion and chemistry modelling, the approach based on a multi-model ensemble of forecasts has been developed to provide better information by combining 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, Bayesian models or spectral methods (e.g. Delle Monache et al., 2006; Riccio et al., 2007; Potempski et al., 2010; Galmarini et al., 2013).

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

2012), EURAD-IM (European Air pollution Dispersion Inverse Model; Haas et al., 1995; Memmesheimer et al., 2004), LOTOS-EUROS (Long Term Ozone Simulation - European Ozone Simulation; Schaap et al., 2008), MATCH (Multi-scale Atmospheric Transport and Chemistry; Robertson et al., 1999; Andersson et al., 2015), MOCAGE (Model Of atmospheric Chemistry At larGE scale; Josse et al., 2004; Dufour et al., 2004) and SILAM (System for Integrated modeLling of Atmospheric coMposition; Sofiev et al., 2008). They are used to produce a multi-model ensemble for major monitored pollutants. Although each of these models can perform very well on particular days in particular areas, the ensemble approach aims at providing, on average, forecasts and analyses of better quality than any of them individually. It also gives an indication of the uncertainties through the spread between the models. Similarly to meteorological forecasts, the quality of the AQ forecasts needs to be routinely evaluated to provide information to users about its reliability. The performance of the individual and ensemble forecast products is evaluated on a daily basis from comparisons with available surface observations by the European AQ station network. Additionally, the system has been providing birch pollen forecasts at the surface during the pollen season since 2013. All the forecast and analysis numerical data are publicly available.

The objectives of the paper are, firstly, to provide a description of the pre-operational analysis 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-model ensemble. Since the system continuously evolves with time, we present here the configuration at the end of the MACC-II project (summer 2014) with a brief description of recent upgrades included before the end of 2014. An overview of the analysis and forecasting system, including the seven models and the ensemble median, is provided in Sect. 2. Section 3 is devoted to the system performance for case studies and on a seasonal basis. Section 4 gives a summary and the perspective on short- and mid-term developments of the MACC-II system and associated research.

## 2 Description of the analysis and forecasting systems

### 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 to 70° N. Each of the seven models is run at its own horizontal and vertical resolutions, with the horizontal resolutions varying between ∼ 20 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 96 h from 00:00 UTC on Day0 with hourly outputs on eight vertical levels (surface, 50, 250, 500, 1000, 2000, 3000 and 5000 m). Day0 is defined as the day when the forecast is run. The forecast initial time/date is Day0 at 00:00 UTC and final time/date is Day3 at 24:00 UTC. For each timestep (1 h), the individual model fields are interpolated on these vertical levels and on the same regular 0.1° latitude by 0.1° longitude grid over the MACC-II European domain. It is from these re-gridded fields that the ensemble median and verification products are calculated. Before mid-May 2014, only the surface, 500, 1000 and 3000 m levels were produced. The forecast species include $O_3$, $NO_2$, $SO_2$, $CO$, $PM_{10}$ and $PM_{2.5}$, which are called core species hereafter. The core species are monitored in near-real time (NRT) by European air quality stations and forecasts can therefore be evaluated routinely against these observations. Forecasts of birch pollen concentrations at surface are also produced during the pollen season (1 March–30 June) since 2013. This product is not discussed in this paper since its description and validation is detailed in Sofiev et al. (2015). Additionally, since mid-May 2014, the production has been extended to other species or aggregation of species (NO, $NH_3$, PAN+PAN precursors, total non-methane volatile organic compounds – NMVOCs). Additional species are provided primarily for the use as initial and/or boundary conditions mainly for finer-scale models designed for local AQ purposes.

The analysis at the surface for Day0–1 (the day before Day0) is run daily a posteriori on Day0 using the assimilation of the hourly data from the AQ monitoring stations available in Europe between 00:00 and 23:00 UTC on Day0–1. Like for the forecasts, Day0 is defined as the day when the analysis is run. . The analysis initial time/date is Day0–1 at 00:00 UTC and final time/date is Day0–1 at 23:00 UTC. 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 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. These production times have been shifted earlier since summer 2014 in order to fulfil the users' needs, in particular Day0 and Day1 forecasts, which are the mostly used products, are now available at 07:00 UTC. This has been made possible by an earlier delivery of the forecasts of each of the seven models and by replacing the bulk 96 h processing of the ensemble by processing 24 h segments. The delivery time of the analysis has also been shifted earlier in June 2014.

The NRT hourly observations of $O_3$, $NO_2$, $SO_2$, $CO$, $PM_{10}$ and $PM_{2.5}$ from the European AQ monitoring stations are used for model assimilation to produce the daily analyses and

also for the forecast and analysis evaluation. From 2009 until recently, they were gathered country by country through bilateral agreements with the project. Since 2014, a new system has been put in place to gather these observations from the centralised AirBase database maintained by the European Environment Agency (EEA). The database collects the NRT data and validated data from the European countries bound under Decision 97/101/EC to engage in a reciprocal exchange of information (EoI) on ambient air quality. The delivery time of the observations to 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 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 07:00 UTC on Day0 for the data set covering Day0–1. At this time of the day, more than 90 % (on average) of all data are available. The 07:00 UTC cut-off time is therefore a compromise between having enough data available for the model assimilation and a reasonable production time for the ensemble analysis that was at 14:30 UTC at the end of MACC-II. This production time is still too late for the forecasts to be initialised from the analysis, meaning that the forecast and the analysis 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 23:00 UTC on Day0 are used since there is less constraint on the time of delivery of evaluation products. On average there is about 10 % more data available at 23:00 UTC than at 07:00 UTC. As shown in a MACC-II report (D16_3; http://www.gmes-atmosphere.eu/ documents/maccii/deliverables/obs/), the additional data collected at 23:00 UTC compared to 07:00 UTC are mainly data from the end of the previous day. This is because there is a significant number of stations that do not send their late afternoon and evening Day0–1 data before 07:00 UTC on Day0. This means that the 23:00 UTC data set used for verification is homogeneous with approximately the same number of observations in the morning, afternoon and evening.

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 classification of sites, based on past validated measurements available in the AirBase database (EEA). Stations are classified between 1 and 10 depending on the characteristics of their series of measurements (diurnal cycle, "weekend effect" and high frequency variability with periods lower than 3 days). The orig-

**Table 1.** Portfolio of the MACC-II regional data products. Each product is provided once daily. Core species correspond to $O_3$, $NO_2$, CO, $SO_2$, $PM_{10}$, $PM_{2.5}$. Additional species correspond to NO, $NH_3$, PAN+PAN precursors, and total non-methane volatile organic compounds. Birch pollen concentrations are only available during the season from 1 March to 30 June each year. Old levels refer to surface, 500, 1000, 3000 and 5000 m, corresponding to the production before mid-May 2014. All levels refers to surface, 50, 250, 500, 1000, 2000, 3000 and 5000 m, produced from mid-May 2014. The analysis is run a posteriori on Day0 for Day−1 (00:00–24:00 UTC).

| Model name | Forecast or analysis | Species | Time span | Vertical levels | Format |
|---|---|---|---|---|---|
| CHIMERE | Forecast | Core + additional | 0–96 h, hourly | All levels | NetCDF |
| CHIMERE | Forecast | Birch pollen | 0–96 h, hourly | Surface | NetCDF |
| CHIMERE | Analysis | $O_3$, $PM_{10}$ | −24 to −1 h, hourly | Surface | NetCDF |
| EMEP | Forecast | Core + additional | 0–96 h, hourly | All levels | NetCDF |
| EMEP | Forecast | Birch pollen | 0–96 h, hourly | Surface | NetCDF |
| EMEP | Analysis | $NO_2$ | −24 to −1 h, hourly | Surface | NetCDF |
| EURAD-IM | Forecast | Core + additional | 0–96 h, hourly | All levels | NetCDF |
| EURAD-IM | Forecast | Birch pollen | 0–96 h, hourly | Surface | NetCDF |
| EURAD-IM | Analysis | $O_3$, $NO_2$, CO, $SO_2$, $PM_{10}$ | −24 to −1 h, hourly | Surface | NetCDF |
| LOTOS-EUROS | Forecast | Core + NO | 0–96 h, hourly | Old levels | NetCDF |
| LOTOS-EUROS | Forecast | Birch pollen | 0–96 h, hourly | Surface | NetCDF |
| LOTOS-EUROS | Analysis | $O_3$ | −24 to −1 h, hourly | Surface | NetCDF |
| MATCH | Forecast | Core + additional | 0–96 h, hourly | All levels | NetCDF |
| MATCH | Forecast | Birch pollen | 0–96 h, hourly | Surface | NetCDF |
| MATCH | Analysis | $O_3$, $NO_2$, CO, $PM_{10}$, $PM_{2.5}$ | −24 to −1 h, hourly | Surface | NetCDF |
| MOCAGE | Forecast | Core + additional (except $NH_3$) | 0–96 h, hourly | All levels | NetCDF |
| MOCAGE | Forecast | Birch pollen | 0–96 h, hourly | Surface | NetCDF |
| MOCAGE | Analysis | $O_3$ | −24 to −1 h, hourly | Surface | NetCDF |
| SILAM | Forecast | Core | 0–96 h, hourly | All levels | NetCDF |
| SILAM | Forecast | Birch pollen | 0–96 h, hourly | Surface | NetCDF |
| SILAM | Analysis | $O_3$, $NO_2$, $SO_2$ | −24 to −1 h, hourly | Surface | NetCDF |
| ENSEMBLE | Forecast | Core + additional | 0–96 h, hourly | All levels | NetCDF + Grib2 |
| ENSEMBLE | Forecast | Birch pollen | 0–96 h, hourly | Surface | NetCDF + Grib2 |
| ENSEMBLE | Analysis | $O_3$ | −24 to −1 h, hourly | Surface | NetCDF + Grib2 |

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

| | Forecast Day0 (0–24 h) | Forecast Day1 (25–48 h) | Forecast Day2 (49–72 h) | Forecast Day3 (73–96 h) | Analysis (−24 to −1 h) |
|---|---|---|---|---|---|
| Core species | 07:00 UTC | 07:00 UTC | 08:00 UTC | 09:00 UTC | 14:30 UTC |
| Additional species | 07:00 UTC | 07:00 UTC | 08:00 UTC | 09:00 UTC | NA |

inal classification of Joly and Peuch (2012) was based on a series of data spanning from 2002 to 2009. It has been updated in MACC-II using version 7 of the AirBase database spanning from 2002 to 2011. Classes 1–10 cover the range from most rural background sites to most locally polluted sites. Once each station is classified we exclude those stations that have a concentration variability that is typical of locations mainly influenced by local phenomena. Only the stations with class numbers ranging from 1 to 5 for 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 indicators. This leads to a typical number in summer 2014 of ∼ 600 sites for ozone, ∼ 500 sites for $NO_2$, ∼ 150 sites for $SO_2$, ∼ 40 sites for CO, ∼ 400 sites for $PM_{10}$, ∼ 150 sites for $PM_{2.5}$. 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 the project. This verification is done in the following way: a list of stations not used for the 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 observations that are kept aside for the verification of analyses is roughly 20 % of the total amount of observations that are downloaded at 23:00 UTC.

The plots of forecasts and analyses from the seven models and the ensemble median, as well as of their scores against observations are available daily at http://macc-raq.copernicus-atmosphere.eu/. Numerical data are publicly available and can be accessed at http://www.
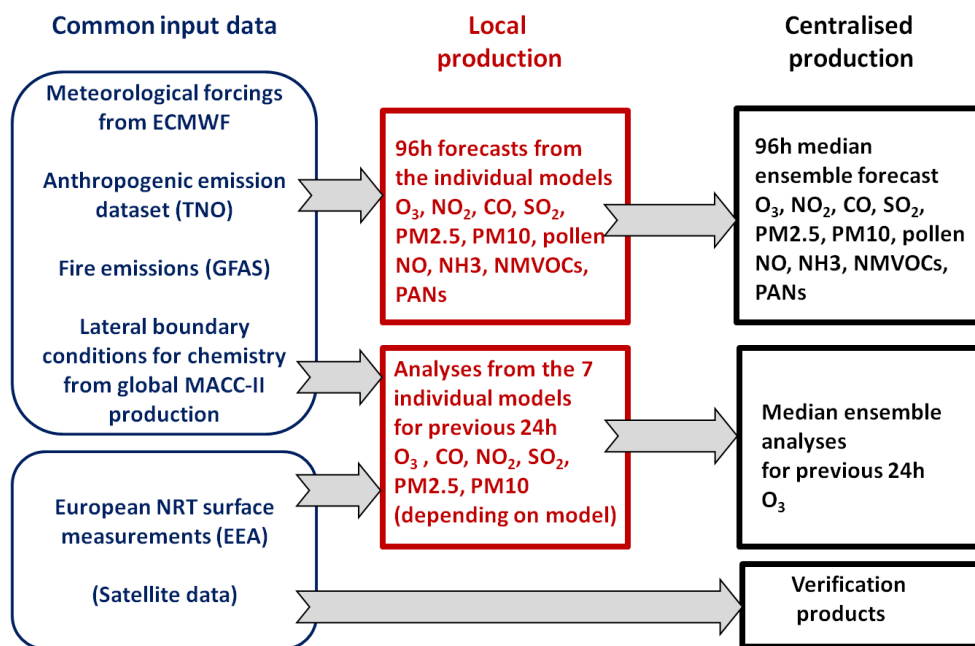
**Figure 1.** Schematic of the general organisation of the MACC-II air quality forecast and analysis system.

gmes-atmosphere.eu/request_regional_data/. The full set of numerical data as listed in Table 1 is made available as soon as it is produced on the Météo-France FTP (file transfer protocol) server. A subset of these data can also be interactively accessed through the Deutsche Zentrum für Luft- und Raumfahrt (DLR) World Data Center.

Major sources of uncertainties in the regional AQ forecasts 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 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.

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 MACC inventory (Kuenen et al., 2011). Its resolution is $1/8°$ longitude $\times 1/16°$ latitude, which is approximately $7\,km \times 7\,km$, and covers the UNECE (Economic Commission for Europe) countries for the years 2003–2009. The

2009 inventory is currently used in the MACC-II daily production. An important upgrade of the MACC-II inventory compared to the earlier MACC inventory is the provision of a particulate matter split between elemental carbon, organic carbon, $SO_4$, Na and other aerosols. More details on this inventory can be found in Kuenen et al. (2014). For the biogenic sources, each model deals with its own emissions based on dynamical parameterisations and/or inventories that are detailed in the following individual model description subsections. Additionally, emissions from fires are taken into account using the GFASv1.1 product (Kaiser et al., 2012) available daily at $0.1° \times 0.1°$ resolution. GFASv1.1 is based on fire radiative power retrievals from data of the Moderate Resolution Imaging Spectroradiometer (MODIS) instruments aboard the Terra and Aqua satellites. The GFAS product for Day0–1 is available around 06:00 UTC on Day0. This is soon enough to be used in the daily analysis of individual production chains. At the time the individual forecasts begin for Day0, only the fire emissions from Day0–2 are available. To have a smaller time gap between the fire emissions and the starting time of the regional forecast runs (usually around 20:00 UTC), an additional fire emission product available around 20:30 UTC on Day0–1 using satellite observations from 15:00 UTC on Day0–2 to 15:00 UTC on Day0–1 is currently under testing. In the forecasts, a persistence of the fire emissions of 3 days is assumed. This is a rounded average of the fire duration obtained by Turquety et al. (2014) from the Euro-Mediterranean region from the MODIS MCD64 product (Giglio et al., 2010) in the period 2003–2012.

The meteorological fields used to force the seven CTMs are from the operational IFS (integrated forecasting system)

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

| Model | Operated by | Horizontal resolution | Vertical levels top height |
|---|---|---|---|
| CHIMERE | INERIS (Institut National de l'Environnement Industriel et des Risques) France | 0.1° × 0.1° | 8 levels Top at 500 hPa |
| EMEP | MET Norway (Meteorologisk institutt) Norway | 0.25° × 0.125° | 20 levels top at 100 hPa |
| EURAD-IM | RIU UK (Rheinisches Institut Für Umweltforschung an der Universität zu Köln E. V.) Germany | 15 km on a Lambert conformal projection | 23 levels Top at 100 hPa |
| LOTOS-EUROS | KNMI (Koninklijk Nederlands Meteorologisch Instituut) the Netherlands | 0.25° × 0.125° | 4 levels Top at 3.5 km |
| MATCH | SMHI (Sveriges Meteorologiska och Hydrologiska Institut) Sweden | 0.2° × 0.2° | 52 levels Top at 300 hPa |
| MOCAGE | Météo-France France | 0.2° × 0.2° | 47 levels Top at 5 hPa |
| SILAM | FMI (Ilmatieteen Laitos) Finland | 0.15° × 0.15° | 8 levels Top at 6.7 km |

daily meteorological forecasts of the European Centre for Medium-Range Weather Forecasts (ECMWF). The IFS forecast starting at 12:00 UTC on Day0–1 is used for the MACC-II air quality 96 h forecast starting at 00:00 UTC on Day0. For the analysis on Day0–1, the IFS forecast starting at 00:00 UTC on Day0–1 is used.

The regional domain boundary conditions for the aerosols and gaseous species are provided 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 concentrations of dust, sea salt, organic matter, black carbon and sulfate aerosols (Morcrette et al., 2009; Benedetti et al., 2009) that are used to force the aerosols in the regional CTMs at 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 (Model for OZone And Related chemical Tracers) global CTM. This allowed for assimilation of satellite data for $O_3$, $NO_2$, and CO in the IFS itself, while the detailed chemical processes were handled in the MOZART model (Flemming et al., 2009; 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 aerosols and chemical species. Instead of the coupling with the MOZART model, the chemistry is now treated on-line in the IFS using chemistry modules based on the TM5 model (Huijnen et al., 2010). This new system is named Composition-IFS (C-IFS) and is further described in

Flemming et al. (2015). The chemical mechanism in the TM5 operational version of C-IFS is based on a modified version of the Carbon Bond 5 (CB05) scheme (Williams et al., 2013; Yarwood et al., 2005).

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

## 2.2 CHIMERE forecast and analysis system

CHIMERE is an Eulerian chemistry-transport model able to simulate concentration fields of gaseous and aerosols species at a regional scale (Menut et al., 2013a). The model is developed under the General Public License licence (http://www.lmd.polytechnique.fr/chimere/). CHIMERE is used for analysis of pollution events, process studies, (Bessagnet et al., 2009; Beekmann and Vautard, 2010), experimental and operational forecasts (Rouïl et al., 2009), regional climate studies and trends (Colette et al., 2011), among others.

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

| Model | Assimilation method | Observation assimilated | Species analysed |
|---|---|---|---|
| CHIMERE | Optimal interpolation | $O_3$ and $PM_{10}$ from surface stations, | $O_3$, $PM_{10}$ |
| EMEP | 3DVar | $NO_2$ columns from OMI and $NO_2$ from surface stations | $NO_2$ |
| EURAD-IM | 3DVar | $O_3$, NO, $NO_2$, $SO_2$, CO, $PM_{10}$, $PM_{2.5}$ from surface stations, OMI and GOME-2 $NO_2$ column retrievals, MOPITT CO profiles | $O_3$, $NO_2$, $SO_2$, CO, $PM_{10}$ |
| LOTOS-EUROS | Ensemble Kalman filter | $O_3$ from surface stations | $O_3$ |
| MATCH | 3DVar | $O_3$, $NO_2$, CO, $PM_{10}$, $PM_{2.5}$ from surface stations | $O_3$, $NO_2$, CO, $PM_{10}$, $PM_{2.5}$ |
| MOCAGE | 3DVar | $O_3$ from surface stations | $O_3$ |
| SILAM | 4DVar | $O_3$, $NO_2$ and $SO_2$ from surface stations | $O_3$, $NO_2$, $SO_2$ |

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° 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 developed by Bessagnet et al. (2004). This module takes into account species such as sulfate, nitrate, ammonium, primary organic matter (POM) and elemental carbon (EC), secondary organic aerosols, sea salt, dust and water. These aerosols are represented using eight bins, from 40 nm to 40 µm, in diameter. The life cycle of the aerosols is completely represented with nucleation of sulfuric acid, coagulation, adsorption/desorption, wet and dry deposition and scavenging. This scavenging is both represented by coagulation with cloud droplets and precipitation. The formation of SOA (secondary organic aerosol) is also taken into account (Bessagnet et al., 2009).

Biogenic emissions are calculated using the MEGAN (Model of Emissions of Gases and Aerosols from Nature) 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 was 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 SOA and secondary 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 points that

may need to be improved, the vertical resolution is composed of eight levels up to 500 hPa, meaning that the models need to be fed with realistic top conditions. The assimilation is thus far limited to $O_3$ and $PM_{10}$ and for the surface layer.

### 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 coefficients are calculated for the whole 3-D model domain on the basis of local Richardson number, and the planetary boundary layer (PBL) height is calculated using methods described 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), 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 the sulfur and nitrogen chemistry to the photochemistry using about 140 reactions between 70 species (Andersson-Sköld and Simpson, 1999; Simpson et al., 2012).

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 sulfide and $SO_2$ from volcanoes.

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 are sulfate, nitrate, ammonium, anthropogenic primary particulate matter, sea salt and desert dust. Aerosol water is also calculated. The parameterisation of dry deposition for aerosols follows standard resistance formulations, accounting for diffusion, impaction, interception, and sedimentation. Wet scavenging is treated with simple scavenging ratios, taking into account in-cloud and sub-cloud processes. For SOA the so-called "EmChem09soa" scheme is used, which is a slightly simplified version of the mechanism described by Bergström et al. (2012).

The EMEP data assimilation system (EMEP-DAS) is based on the 3DVar (3-dimensional variational) implementation for the MATCH model (Kahnert, 2008, 2009). The background error covariance matrix is estimated following the so-called NMC (National Meteorological Center) method (Parrish and Derber, 1992). Currently, the EMEP-DAS delivers analyses for $NO_2$, using $NO_2$ columns of OMI (Ozone Monitoring Instrument) and in situ measurements of $NO_2$ surface concentrations. The assimilation window is 6 h, 4 times per day.

The EMEP model performs well especially for particulate matter, as it includes carefully evaluated representations of both primary and secondary organic aerosols, in addition to inorganic aerosols, elemental carbon, sea salt, mineral dust and water. Another strength is that its domain extends throughout the whole troposphere, thus taking accurate account of long-range transport of pollutants in the free troposphere. As the EMEP model is designed mainly for background concentrations, urban increments have not been implemented as in some other models with equally coarse resolution, leading to somewhat lower performance in urban and sub-urban areas. However, being one of the main research tools under the UN LRTAP (Long-range Transboundary Air Pollution) convention, the EMEP model is evaluated continuously against measurements of a large range of chemical parameters (including air concentrations, depositions, and trends) ensuring 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 that only $NO_2$ was assimilated. However, since early 2015 ozone has been assimilated.

### 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 3DVar and 4DVar 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 parameterise 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 PBL according to the PBL height and the Monin–Obukhov length (Holtslag and Nieuwstadt, 1986). A semi-implicit (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 differential equations (Sandu et al., 2003; Sandu and Sander, 2006). Photolysis frequencies are derived using the FTUV (fast tropospheric ultraviolet–visible) model according to Tie et al. (2003). The radiative transfer model therein is based on the TUV model developed by Madronich and Weller (1990). The modal aerosol dynamics model MADE (Modal Aerosol Dynamics Model for Europe; Ackermann et al., 1998) is used to provide information on the aerosol size distribution and chemical composition. To solve for the concentrations of the secondary inorganic aerosol components, a FEOM (fully equivalent operational model) version, using the HDMR (high dimensional model representation) technique (Rabitz et al., 1999; Nieradzik, 2005), of an accurate mole-fraction-based thermodynamic model (Friese and Ebel, 2010) is used. The updated SORGAM module (Secondary Organic Aerosol Model; Li et al., 2013) simulates secondary organic aerosol formation. Biogenic emissions are calculated in the EURAD-IM CTM with 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 as 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) the formulation of both background error covariance matrices for the initial states and the 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 memory L-BFGS (Broyden–Fletcher–Goldfarb–Shanno) algorithm described in Nocedal (1980) and Liu and Nocedal (1989) is applied for the minimisation. Following Weaver and Courtier (2001) with the promise of a high flexibility in designing anisotropic and heterogeneous influence radii, a diffusion approach for providing the background error covariance matrices is implemented.

One of the EURAD-IM strengths is that it includes a comprehensive treatment of aerosol dynamics and chemistry. Parameterisations of the formation of secondary particles are temperature dependent for both the inorganic and organic components. However, the complexity of the aerosol components of EURAD-IM is as yet not supported by sufficiently known emission rates of particle types, nor for organic gaseous precursor compounds, especially from biogenic sources. Another strength of the EURAD-IM system is its ability to assimilate chemical data from a wide range of instruments ranging from surface or airborne in situ data to retrievals from several satellites, which are then defining the initial values.

## 2.5 LOTOS-EUROS forecast and analysis system

The 3-D chemistry-transport model LOTOS-EUROS (Schaap et al., 2008) is developed by the Dutch institutes TNO (www.tno.nl), RIVM (www.rivm.nl) and, more recently, KNMI (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 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 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 3 years, three 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's web 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 $SO_2$ to $SO_4$ conversion rate, use of AQMEII (air quality model evaluation international initiative) conventions for the fine/coarse dust assignment, update of resistances for e.g. ozone (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 25 m thick surface layer. The lowest dynamic layer is the mixing layer, followed by two reservoir 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 Walcek (2000). Gas-phase chemistry is described using the TNO CBM-IV scheme (Schaap et al., 2008). Hydrolysis of $N_2O_5$ is described following Schaap et al. (2004). Aerosol chemistry is represented using ISORROPIA-2 (Fountoukis and Nenes, 2007). The aerodynamic resistance is calculated for all land use types separately. Below, cloud scavenging is described 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 parameterization for gases (Wichink Kruit et al., 2012) and the Zhang et al. (2001) parameterization for particles.

Biogenic isoprene emissions are calculated following the mathematical description of the temperature and light dependence of the isoprene emissions, proposed by Guenther et al. (1993), using the actual meteorological data. For land use the CORINE/Smiatek database has been enhanced using the tree species map for Europe made by Koeble and Seufert (2001). Total $PM_{10}$ in the LOTOS-EUROS model is composed of chemically unspecified PM in the fine and coarse modes, black carbon, dust, ammonium, sulfate, nitrate and sea salt (Na in the fine and coarse modes).

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 favour short computation times. For this, the vertical top of the operational model version is limited and covers only the boundary layer and reservoir layers (up to 3.5 km); effectively, the model therefore employs only four dynamic layers. Concentrations from the free troposphere are taken from the global boundary conditions, and therefore fully incorporate the knowledge, assimilations, and validation efforts present in the global model. A major weakness is that secondary organic aerosols are currently not included; instead, a bias correction for total PM is used to account for the missing aerosols. In spite of the limited complexity, the model performs well in simulation of $O_3$ (Curier et al., 2012) and has a skill to forecast the observed variability in $PM_{10}$ (De Ruyter de Wildt et al., 2011). Apart from the relative short run-through time, the strength of the model is in the detailed description of anthropogenic emissions, given the close cooperation with the developers of the TNO-MACC emission inventory; this is for example shown by excellent simulation of boundary layer $NO_2$ (Vlemmix et al., 2015).

## 2.6 MATCH forecast and analysis system

The MATCH 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 turnover 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 described by Van Ulden and Holtslag (1985) for land surfaces, and Burridge and Gadd (1977) for sea surfaces. The boundary layer height is calculated from formulations proposed by Zilitinkevich and Moronov (1996) for the neutral and stable case and from Holtslag et al. (1995) for the convective case. These parameterisations drive the formulations for vertical 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. (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 is the most notable based on the so-called Carter-1 mechanism (Carter, 1996; Langner et al., 1998).

Aerosols are described for four 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).

A 3-D variational data assimilation scheme is used with spectral transformation (Kahnert, 2008). The limitation then is that background covariance structures are described as isotropic and homogeneous, however, not necessarily the same for different wave numbers and derived 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 covariance between spectral components are explicitly handled. The block diagonal elements are the covariance between wave components at different model layers and chemical compounds.

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 (HIgh-Resolution Limited Area Model) data with a coarser vertical resolution, the use of the high-resolution vertical levels from IFS may lead to less accurate chemistry forecasts compared to the HIRLAM version. A weakness is missing SOA and wind-blown dust in the PM description.

## 2.7 MOCAGE forecast and analysis system

The MOCAGE 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 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 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). MOCAGE includes the RACM scheme for tropospheric chemistry (Stockwell et al., 1997) and the REPROBUS scheme for stratospheric chemistry (Lefèvre et al., 1994). Biogenic 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. (2015).

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., 2006; Lahoz et al., 2007). The system, recently renamed VALENTINA, has been used to compute global and regional re-analyses of atmospheric composition in multiple studies (El 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 3DVar with assimilation windows of 1 h length (Jaumouillé et al., 2012), corresponding to the frequency of surface measurements. The assimilation has first been set for surface ozone analyses and in MACC-III it has been extended to surface $NO_2$. The specification of the background and observation errors is done based on the evaluation of historical time series of observations and model values. The horizontal error correlation has a Gaussian shape and its typical length is set to 0.4° for ozone and 0.1° for $NO_2$, to account for the larger variability of $NO_2$ at fine spatial scales. The vertical error correlation length is set to one model grid point for all species ($\sim 100$ m). As a consequence, assimilation increments linked to surface observations are confined in the planetary boundary layer.

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 particular boundary layer–troposphere and

troposphere–stratosphere exchanges, and the time evolution of stratospheric conditions for accurate photolysis rate calculations at the surface. The MOCAGE assimilation system in its MACC configuration produces robust analyses for both $O_3$ and $NO_2$ as illustrated in the annual re-analysis reports (http://www.gmes-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 aerosols have been developed recently and will be included in the next MACC operational version (Guth et al., 2015). This new feature is also used in the current development of $PM_{10}$ assimilation.

## 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, emergencies, composition–climate interactions, and air quality modelling purposes. The model has been applied with resolutions ranging from 1 km up to 3°, incorporates eight chemical and physical transformation modules and covers the troposphere and the stratosphere. The model is publicly available since 2007 and is used as an operational and research tool.

The model has two dynamic cores: Lagrangian (Sofiev et al., 2006), primarily used in emergency-type applications, and Eulerian (Galperin, 2000; Sofiev, 2002) used in atmospheric composition, climate, and air-quality-related applications, including MACC-II. The MACC-II operational SILAM v.5.2 uses the simple dry deposition scheme of Sofiev (2000) for gases and a new approach for aerosols Kouznetsov and Sofiev, 2012), which covers particle sizes from 1 nm up to $\sim 50\,\mu m$ of effective aerodynamic size. The wet deposition scheme used in MACC-II simulations calculates the 3-D removal coefficient and distinguishes between sub- and in-cloud scavenging, large-scale and convective 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 the 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 species-specific size spectra. The aerosol species include primary anthropogenic aerosols, divided into $PM_{2.5}$ and $PM_{10}$, secondary inorganic aerosols (sulfates, nitrates and ammonia), and sea salt aerosols.

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

The data assimilation system of SILAM consists of 3DVar and 4DVar modules (Vira and Sofiev, 2012). The MACC-II near-real time analysis suite uses the 3DVar method and assimilates hourly surface observations of $NO_2$, $O_3$ and $SO_2$. PM observations have been assimilated in re-analysis simulations (Vira and Sofiev, 2015). The 4DVar methodology is utilised in re-analysis mode for pollen.

The model evolution from the MACC-II v.5.2 towards v.5.4, which will become operational in early 2015, includes 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, 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 $PM_{2.5}$, especially in winter when inorganic aerosols are dominant. The current limitation of the model is the secondary organic aerosols formation that makes use of the volatility-based model but it is not yet incorporated in the operational simulations, being tested in research projects. A workaround of this limitation is included in the data assimilation modules, which allow for assimilation of both in situ and remote-sensing measurements of gaseous and particulate species. The module now allows for the PM and aerosol optical depth observations being assimilated into an unspecified particulate matter, which is then treated as inert aerosol, thus compensating for the lack of secondary organic particles.

## 2.9 ENSEMBLE forecast and analysis system

To process the ensemble median, all seven individual models are first interpolated to a common $0.1° \times 0.1°$ horizontal grid. For each grid point, the ensemble model (referred to as ENSEMBLE hereafter) value is calculated as the median value of the individual model forecasts or analyses available. The median is defined as the value having 50 % of individual 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 useful from an operational point of view. The method is also little sensitive if a particular model forecast or analysis is occasionally missing. The performances of the ensemble median are discussed in Sect. 3. For the forecasts, the ENSEMBLE is produced for all levels and all species (core and additional). For the analyses, the individual assimilation systems provide only analyses at the surface level and do not produce analyses for all species yet. At the end of MACC-II, ozone was the only species that was produced by six of the models. For other species, analyses from less than five models were available. This is why the ENSEMBLE analy-

sis in MACC-II was only calculated for ozone. It has been extended to $NO_2$ in 2015 since more models will produce $NO_2$ analyses.

## 3 Evaluation of the performances of the system

### 3.1 General description

The evaluation of the performances of a forecast system is a necessary step for rating its quality and thus proving its usefulness. The MACC-II air quality forecasts are evaluated against the NRT AQ surface monitoring data detailed in Sect. 2.1. Note that this set of data is fully independent of the forecast since the analyses assimilating the NRT AQ data are produced too late to be used to initialise the forecasts. The tools to assess the performances of the analyses 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, 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.

The forecast performances are measured using the five statistical indicators detailed in the Appendix A: the mean bias (MB), the root mean square error (RMSE), the modified normalised 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 by several means:

– on a daily basis with plots of statistical indicators and charts available on the MACC-II regional website (http://macc-raq.gmes-atmosphere.eu/),

– on a 6-month 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-month basis, reports are especially dedicated to the scientific analysis of the forecasts of the seven models and of the ENSEMBLE in the Mediterranean area (http://www.gmes-atmosphere.eu/documents/maccii/deliverables/ens/). The Mediterranean area is recognised 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 data set was not yet in place at the end of MACC-II project.
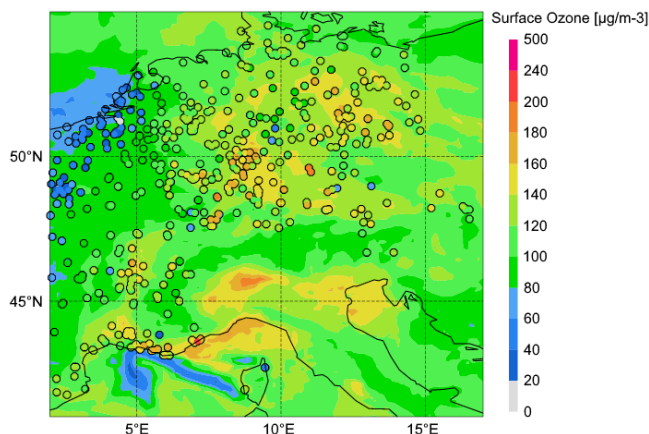


**Figure 2.** Magnified map of ozone concentrations at the surface (in $\mu g\,m^{-3}$) of the 15 h forecast for 10 June 2014 at 15:00 UTC of the ensemble median constructed with the seven model forecasts. NRT AQ observations available (circles) for the same date/time are overplotted on the maps using the same colour scale.

### 3.2 Availability statistics

The MACC-II regional air quality forecasting and analysis system is currently under a pre-operational status that can be seen as the demonstrator of a future operational system. The proper function of the seven model chains and of the ENSEMBLE chain is monitored on working hours only since, at this stage, there had been no funding yet for a 7-day/7-day, 24 h/24 h control. Nevertheless, in their pre-operational configuration the production chains are reliable with availability in time (see Table 2) of the seven individual forecasts and analysis generally above 85 % during MACC-II. During the past year, the production suffered from failures because of the many changes that were applied to the individual and central systems to fit with fully operational standards (data format, file transfer, databases, processing softwares, etc.). With the operationalisation being nearly fully settled, the reliability has been improved since the end of MACC-II (generally above 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 seven models are not available.

### 3.3 Example of the forecast of two ozone episodes between 10 and 13 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 10 and 13 June 2014. A more in-depth analysis of the individual model and of the ENSEMBLE performances is done over longer time periods in Sect. 3.4.

During the case study period, there were two regional areas with high ozone concentrations ($> 120\,mg\,m^{-3}$) occur-

ring at the same time, one over Austria and surrounding regions (south of Germany and Hungary), and one over the south-east of France and the north of Italy. This is illustrated by Fig. 2 displaying the maps of the 15 h forecasts, for the June at 15:00 UTC, of ozone at the surface from the ENSEMBLE together with the available observations. Note that unfortunately there are no observations from Italy available during the time period considered. Even if the comparison is limited by the missing observations, Fig. 2 shows that the ensemble median captures the two ozone episodes.

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 representation of the spread of the seven 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 Europe and made available on the MACC regional website. Here, EPSgrams are calculated and plotted for the same locations as the measurements (Fig. 3) from the forecast started on 10 June at 00:00 UTC. Note that, in Fig. 3, EPSgrams are 3-hourly while observations are hourly. In the observations (left panel), the "Austrian" episode is highest on 10 June in Fechenheim (Germany) and on 11 June in Hallein (Austria) and Sopron (Hungary) with values reaching 200 µg m$^{-3}$ at Sopron. The "French" episode peaks at 250 µg m$^{-3}$ at Sausset (France) with daily maxima over 150 µg m$^{-3}$ from 10 to 13 June for all three stations.

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 the Sopron station there is a main peak in the model on 11 June as in the observations. For Fechenheim, the forecast gives highest concentrations on 10 June as measured and also reproduces the anomaly recorded in the observations in the morning of 11 June. For this "Austrian" ozone episode, the spread between the seven models is very reasonable (generally less than 30 µg m$^{-3}$ for the 25–75 % range), showing the good consistency between the models with slightly larger spread between the forecasts for the highest peak times. There is an exception in Fechenheim on 10 June where the seven models exhibit a large spread. This can be explained by the effect of complex topography combined with specific meteorological conditions that lead to different behaviours of the models which have different horizontal grids and orography.

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 30–50 µg m$^{-3}$ but the maxima of the seven models are nevertheless close to those observed. There are also cases when the ensemble median forecasts higher peaks than measured as in Fechenheim on 10 June. This can also be seen in the map in Fig. 2 where some observations are lower than the ensemble median. During night-time, the ozone median is close to the observed values.

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 seven models are over 120 µg m$^{-3}$ for each of the 4 days forecasted. Nevertheless, at Sausset the very high ozone peak measured on 10 June (over 240 µg m$^{-3}$) is underestimated in all seven forecasts. The very small spread of the models indicates a possible error in the meteorological forecast for this day and/or in the emissions. Sausset is located on the Mediterranean coast, west of Marseille, an industrial city. On this particular day, IFS forecasts an eastern wind with high NO$_x$ from Marseille limiting the daytime production of ozone by the models compared to observations. For Plan d'Aups, there is a very large spread of the models, particularly on 10 June. This can be explained by the effect of sea and land breezes on this date combined with steep orography and the presence of a pollution plume nearby with NO$_x$ titrated ozone that leads to model differences. In this case (Plan d'Aups), models do not reproduce observations. While for the other stations, the behaviour of the ensemble median is good. For St Rémi, the models perform well with a small spread and diurnal variations close to the 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 often lower than observed, but not always, as shown at Sausset on 13 June and St Rémi on 10 June. In the case of underestimation of the ensemble median, the maximum of the individual models are 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 the meteorological forecast.

For further evaluation, Fig. 4 displays the MB, MNMB, RMSE, FGE and $R$ (defined in Appendix A) of ozone of the seven individual models and the ENSEMBLE calculated using the representative observations available over the whole European domain. These statistics are based on seven consecutive 96 h forecasts run every day from 9 to 15 June. This figure shows that there is a spread of the seven models and that the ENSEMBLE 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, exhibit a diurnal cycle with higher correlations and lower RMSE and FGE during daytime (when ozone is high) than during night-time. 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 over 1 week but there is a good consistency with scores based on longer time series, as shown and analysed in detail in Sect. 3.4.

To illustrate the behaviour of the MACC-II system in the case when some of the seven models are missing for the
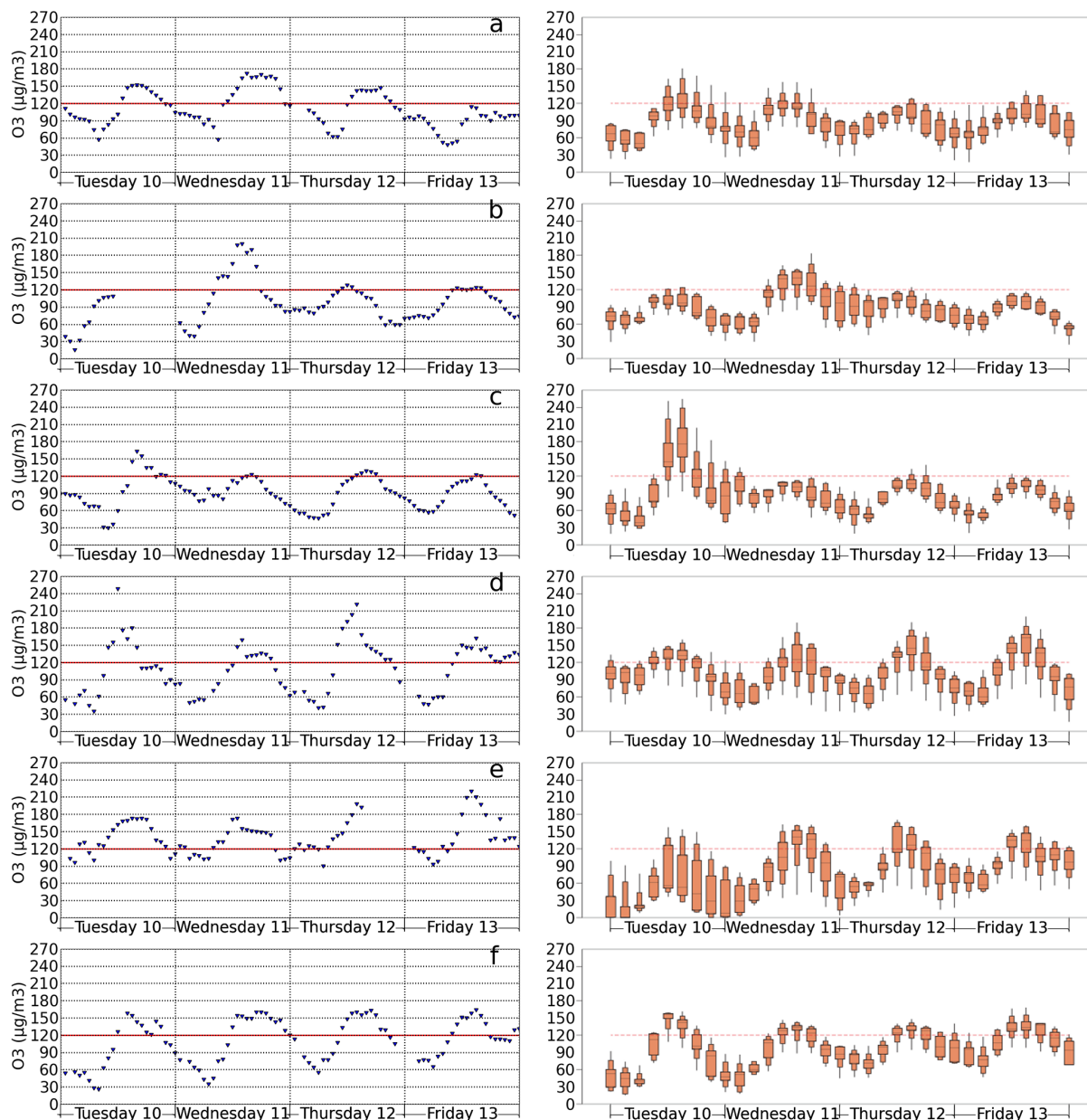
**Figure 3.** Left panels: ozone measurements from surface stations (in µg m$^{-3}$) from 10 June 2014 at 00:00 UTC to 14 June 2014 at 00:00 UTC, located at **(a)** 47.67° N, 13.11° E (Hallein, Austria), **(b)** 47.69° N, 16.58° E (Sopron, Hungary), **(c)** 50.13° N, 8.75° E (Fechenheim, Germany), **(d)** 43.33° N, 5.12° E (Sausset, France), **(e)** 43.34° N, 5.73° E (Plan d'Aups, France) and **(f)** 43.79° N, 4.83° E (St Rémi, France). Right panels: EPSgrams giving median, 90th percentile, 75th percentile, 25th percentile, 10th percentile, minimum and maximum from 3 h outputs of the 96 h forecasts of the seven models from 10 June 2014 at 00:00 UTC to 14 June 2014 at 00:00 UTC. Model outputs are interpolated at the location of the stations shown in the left panel. The red dashed line corresponds to 120 µg m$^{-3}$.

production of the ENSEMBLE, we selected as an example the period of 9–15 June 2014 corresponding to the ozone episodes discussed in Sect. 3.2 and we compared the following ensembles:

– "MEDIAN 7", the operational ensemble method which is the median of the seven models (i.e. ENSEMBLE) as presented in Fig. 4;

– "MEDIAN 5", built on five individual models, after filtering out the best and the worst models, according to the criterion described below;

– "MEDIAN 3", built on three individual models, after filtering out the two best and the two worst models, according to the criterion described below;

**Surface Ozone Forecast**
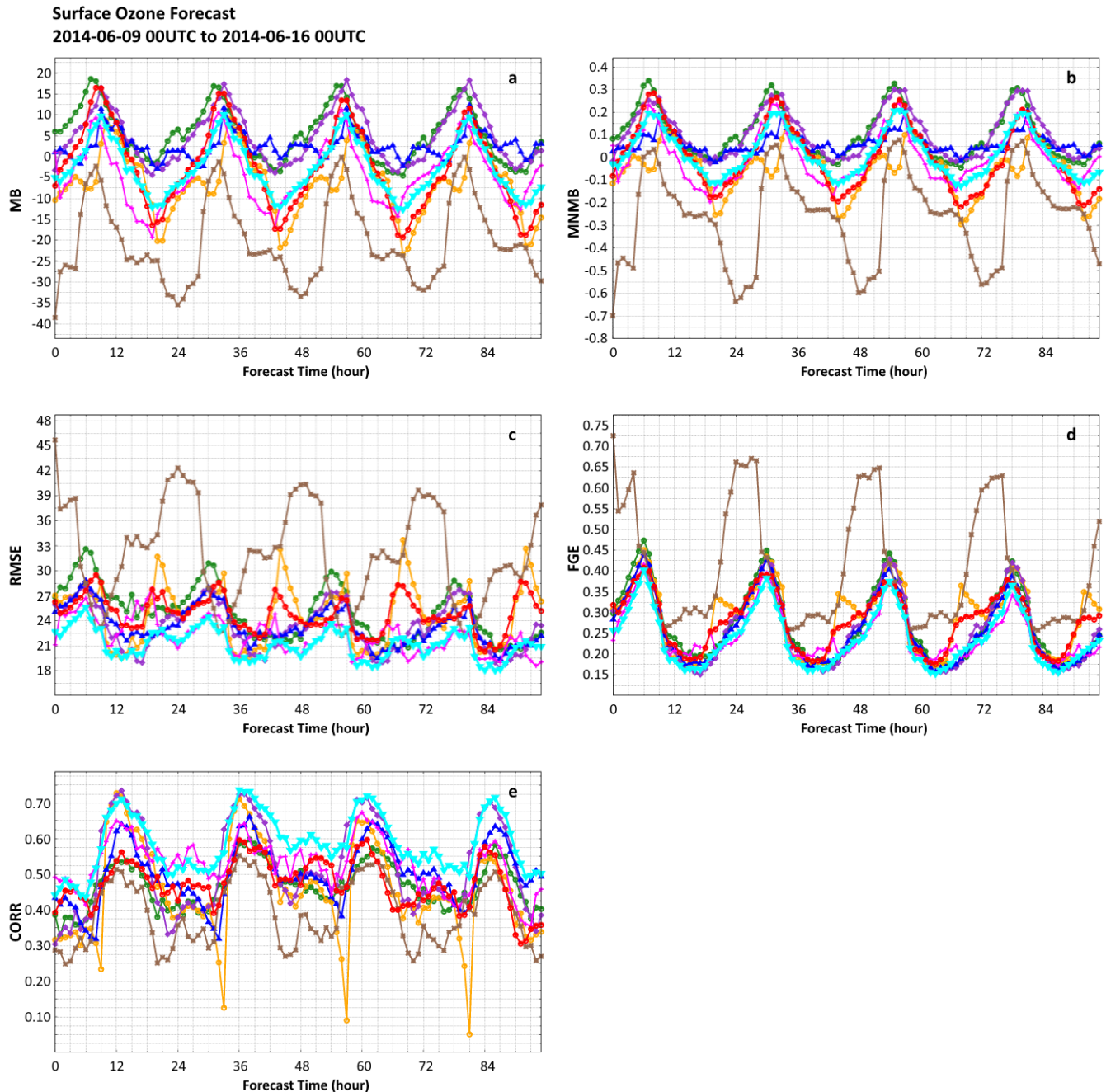**2014-06-09 00UTC to 2014-06-16 00UTC**



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

– "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 seven individual models from worst to best is measured by their RMSE over the seven days of the verification between 12:00

and 18:00 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 MNMB 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 in Fig. 4c and the worst in brown.

**Surface Ozone Forecast**
**2014-06-09 00UTC to 2014-06-16 00UTC**



**Figure 5.** Statistical indicators (see Appendix A) for ozone as a function of the forecast time (in hours) 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 9 to 15 June 2014 over the MACC-II European domain. **(a)** MB (in $\mu g\,m^{-3}$), **(b)** MNMB, **(c)** RMSE (in $\mu g\,m^{-3}$), **(d)** FGE and **(e)** correlation.

Results of the sensitivity experiments are shown in Fig. 5. This figure confirms that the ensemble median (MEDIAN 7) using all seven models performs generally better than the best model on all statistical indicators. When only five models (excluding the best and the worst) are available to calculate the ensemble, all scores show only very slight differences with the ENSEMBLE (MEDIAN 7) based on seven models.

Going to only three models to calculate the ensemble (MEDIAN 3) leads to statistical indicators degraded compared to the ensemble from seven (MEDIAN 7) or five (MEDIAN 5) models but performs generally better than the best model (1BEST). This indicates that using an ensemble of models, even if reduced, is more useful than using a single model event of very good quality. This also shows that with five

models available (which may happen in case of problems of production of two of the seven models), the ensemble median is still robust compared to observations.

In our tests we disregarded the worst (or two 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 Sect. 3.4 over a longer time period.

## 3.4 Statistical performances of the forecasts on a seasonal basis

In addition to the production of daily skill scores, statistical indicators are calculated for ozone, $NO_2$ and $PM_{10}$ at the surface on a seasonal basis since September 2009 for each of the seven models and for the ENSEMBLE. These skill scores and the analysis of their seasonal and year-to-year evolutions are presented in 6-month reports, each including two seasons. The model's statistical indicators are calculated against measurements from the European AQ surface station network available in NRT and selected as detailed in Sect. 2.1. So far, the data provision in NRT is not fully operational. Therefore, there is some variability with time of the number of data available and of their location. Also, the spatial coverage of the surface AQ network in Europe is very inhomogeneous with a high density of stations in France, Germany, UK, Belgium and the Netherlands. Thus, the statistical indicators are more representative 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 (1 June– 1 September 2014) and $PM_{10}$ in winter (1 December 2013– 1 March 2014). We do not show scores for previous years since the use of a different set of surface observations from one year to another does not allow for a fair comparison of the model skills.

MB, MNMB, RMSE, FGE and $R$ for ozone in summer 2014 from the seven models and the ENSEMBLE are shown in Fig. 6. One main feature, which is common to all models including the ENSEMBLE, is that there is no day-to-day degradation of MB, MNMB, and FGE indicators from Day0 to Day3 and a slight increase of the RMSE around 15:00 UTC (about $1 \mu g m^{-3}$). 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 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 previous years. Correlations give a measure of the ability of each model to fit the time variations of ozone regardless of concentration biases. Therefore, correlations are more sensitive to the meteorological forecast

skills than MB, MNMB, RMSE and FGE. Nevertheless, the decrease of the correlation with forecast day is slow.

In Fig. 6 there is a marked diurnal cycle of all statistical indicators for all models which leads to a similar diurnal cycle in the ENSEMBLE scores. MNMB, FGE and $R$ show best performances peaking at 15:00 UTC and worst peaking at 06:00 UTC for each of the 4 days of the forecast. This means that all models are able to simulate the ozone daytime photochemistry 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 statistical indicators can be at least partly explained by uncertainties in the diurnal cycle of the emissions of ozone precursors used in the individual models. This is illustrated with CHIMERE correlation at night, which is better than most of the other models. CHIMERE has developed diurnal factors for traffic emissions based on an objective analysis of $NO_2$ measurements in the different countries in Europe, which improves ozone titration at night (Menut et al., 2012). 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 layer in the LOTOS-EUROS simulations is described with a single model level, with a diurnal variation in the boundary layer height obtained 3-hourly from the ECMWF forecasts. This differs from the description of vertical mixing in the other models and may be responsible for the low correlation feature at around 09:00 UTC. MATCH shows the largest diurnal variability that can be partly related to a combination of chemistry, deposition and the vertical resolution, where the latter is inherited from the IFS model with a rather shallow lowest model layer ($\sim 20$ m). The ozone depletion processes at the surface appear too strong and not enough 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's skill.

Figure 6 also shows that there is generally a positive bias (both in MB and MNMB) of the ENSEMBLE for each of the 4 days of the forecast except around the end of the afternoon when there is a slight ozone underestimation. The ENSEMBLE MB varies from $-6$ to $15.5 \mu g m^{-3}$. This is consistent with the behaviour of the individual models, most of them having a positive bias on 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 (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 new scheme explicitly accounting for the leaf area index is being tested in a pre-operational regime.

The normalised indicators (MNMB and FGE) of all seven models vary from $-0.45$ to $0.45$ and from $0.16$ to $0.59$, respectively. This means that all models performed fairly well

**Surface Ozone Forecast**
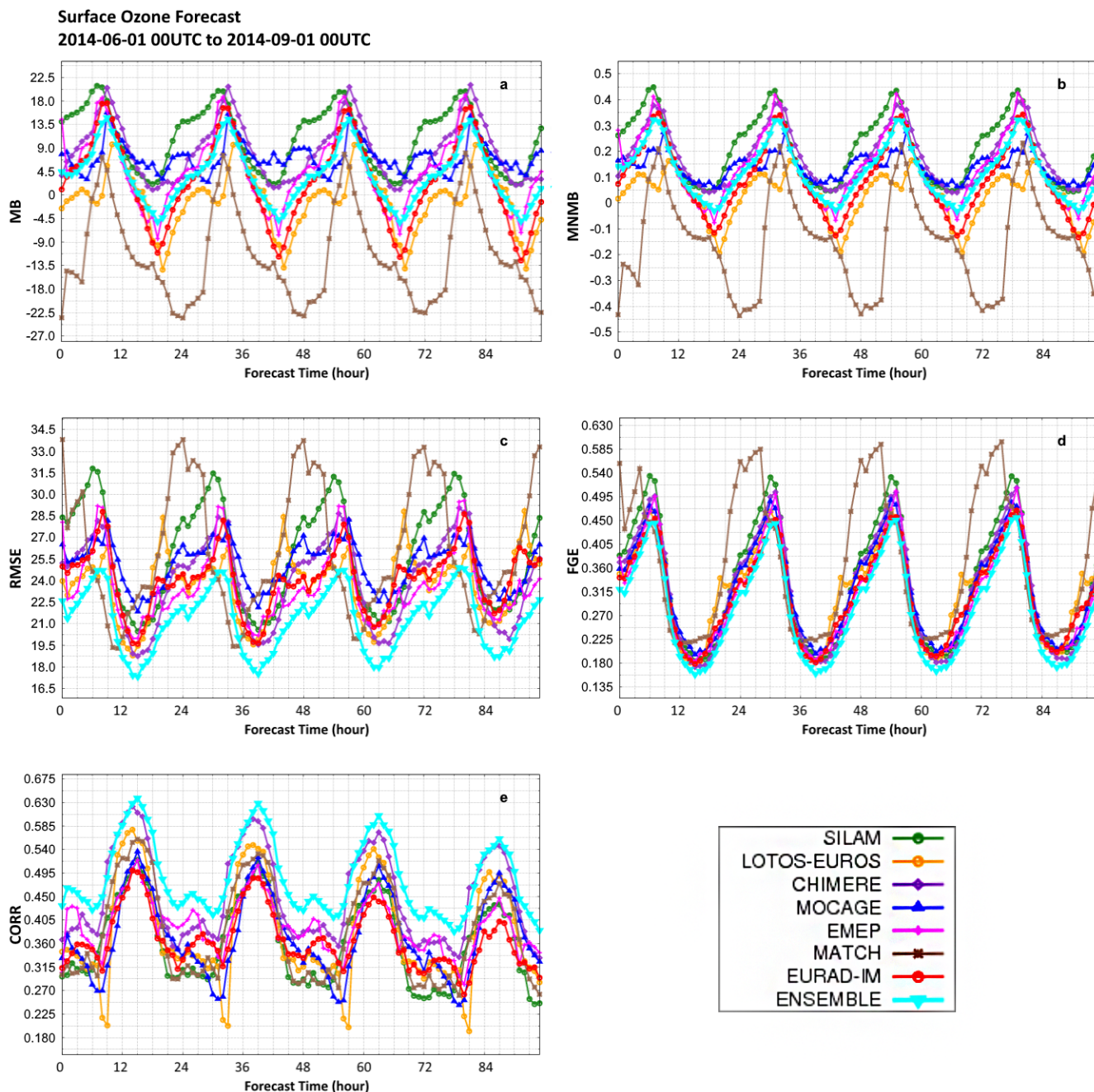**2014-06-01 00UTC to 2014-09-01 00UTC**



**Figure 6.** Statistical indicators (see Appendix A) for ozone as a function of the forecast time (in hours) for the seven models and the ENSEMBLE compared to the hourly surface station measurements available for the period from 1 June at 00:00 UTC to 1 September at 00:00 UTC over the MACC-II European domain for 2014: **(a)** MB (in $\mu g\,m^{-3}$), **(b)** MNMB, **(c)** RMSE (in $\mu g\,m^{-3}$), **(d)** FGE and **(e)** correlation.

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, consistent 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 parameterisations). The

ENSEMBLE gives generally better scores for ozone than any of the individual models.

In Sect. 3.3, we made a first investigation of the robustness of the ensemble median method with regard to the number of models available on a case study of 1 week. To go a step further, we ran a series of tests by removing one or more models in the calculation of the ensemble median over the 3 months of summer 2014. The removal is done randomly on each of the daily forecasts. Figure 7 shows the statistical results against observations for the ENSEMBLE (seven models) and the other ensemble medians calculated by removing randomly one, two, three or four models. For MB, MNMB and FGE, there is hardly any difference between all ensembles. Only RMSE and $R$ (correlation) give significant changes. As expected, decreasing the number of models used in the ensemble tends to degrade its performances. Using six models gives a RMSE and $R$ close to that of the full ensemble based on seven models. The scores for ensembles with four and five models are close to each other but are degraded compared to when seven or six models are used. When only three models are used, RMSE and $R$ are worse compared to the other configurations by $\sim 0.5\,\mu g\,m^{-3}$ and $\sim 0.05$, respectively. This shows that the multi-model ENSEMBLE at the end of MACC-II, which is based on the median of seven models, is robust even if two or three models are unavailable. These results are consistent with the results discussed in Sect. 3.3 that were calculated over 1 week and with a different method for the model removal.

Figure 8 shows the seven models and the ENSEMBLE scores for $PM_{10}$ for the last winter (2013–2014) of MACC-II. Since there were much fewer observations available at 00:00 UTC compared to other times of the day, the values given at the forecast times of 0, 24, 48, 72 and 96 h show a specific behaviour that is not analysed since it is not typical. Among the seven models, MATCH $PM_{10}$ scores 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 components that build up $PM_{10}$ in the model. Secondary organic aerosols are not yet included and there should be an underestimate, rather than an overestimate of $PM_{10}$ by MATCH. The poor correlations in the period are partly related to a too strong signal from sea salt. The verification for 2015 shows clearly that correction of these errors has improved the simulation of $PM_{10}$. There is also a positive bias ($\sim 3.5\,\mu g\,m^{-3}$ on average for MB and 0.3 for MNMB) for CHIMERE. This could be due to the specific setup in CHIMERE being more efficient in the detection of $PM_{10}$ threshold exceedances during wintertime. It is achieved with a correction for lowering the wind over urban areas and with the modulation of the emissions from domestic heating to account for the impact of extremely low temperatures occurring during a cold surge for instance. The other extreme models that exhibit a negative bias of about $-6\,\mu g\,m^{-3}$ on average for MB and $-0.37$ for MNMB are MOCAGE and SILAM. For MOCAGE, this

is due to the lack of secondary aerosols in the model. Although secondary aerosols are not dominant in $PM_{10}$ in winter, there is an expected contribution of sulfates mainly in eastern Europe. Substantial bias of SILAM $PM_{10}$ is caused by the missing SOA. However, the model showed the highest spatial correlation with the $PM_{10}$ observations, which largely follows from the detailed treatment of sea salt emission and transport. As a result, SILAM also showed among the lowest RMSEs. LOTOS-EUROS, EURAD-IM, and EMEP have a smaller negative MB and MNMB. LOTOS-EUROS has an advanced treatment of the SIA fraction but the 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 comprehensive treatment of SIA and SOA. The ENSEMBLE MB and MNMB both indicate a low bias related to the fact that five of the seven models have a negative bias. RMSE and FGE (Fig. 8c, d) are consistent with bias scores with largest values for MATCH and MOCAGE.

In Fig. 8, MB, MNMB, RMSE and FGE are best during daytime (generally around 06:00–07:00 and 15:00 UTC) with diurnal variations fairly similar for all models. This is related to the fact that $PM_{10}$ is dominated by primary anthropogenic emissions of black and organic carbon which are prescribed in all models by the same TNO inventories and which have maxima in the morning and in the afternoon. The worst MB, MNMB, RMSE and FGE are at night, as for ozone. This may be linked to uncertainties in the boundary layer height at night, in vertical diffusion 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 skill. As for ozone, this decrease is likely 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, a bit higher than 0.4 at maximum. This is due to the lack of certain types of aerosols (SIA and/or SOA) in some models but also likely to uncertainties in the diurnal cycle of the 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 anthropogenic emissions of aerosols in Europe is generally modelled as two fixed peaks that do not fully take into account the differences in habits between the countries in Europe. Also, sea salts contribute to $PM_{10}$ on the western side of Europe with emissions and scavenging depending closely on meteorological conditions and therefore directly affected by meteorological uncertainties.

The ENSEMBLE scores are best compared to the seven individual models for RMSE and FGE. For MB and MNMB, the EURAD-IM and EMEP models perform better than the ENSEMBLE. This shows that there are compensating positive/negative biases in EURAD-IM and EMEP that are removed when RMSE and FGE scores are considered. The EN-
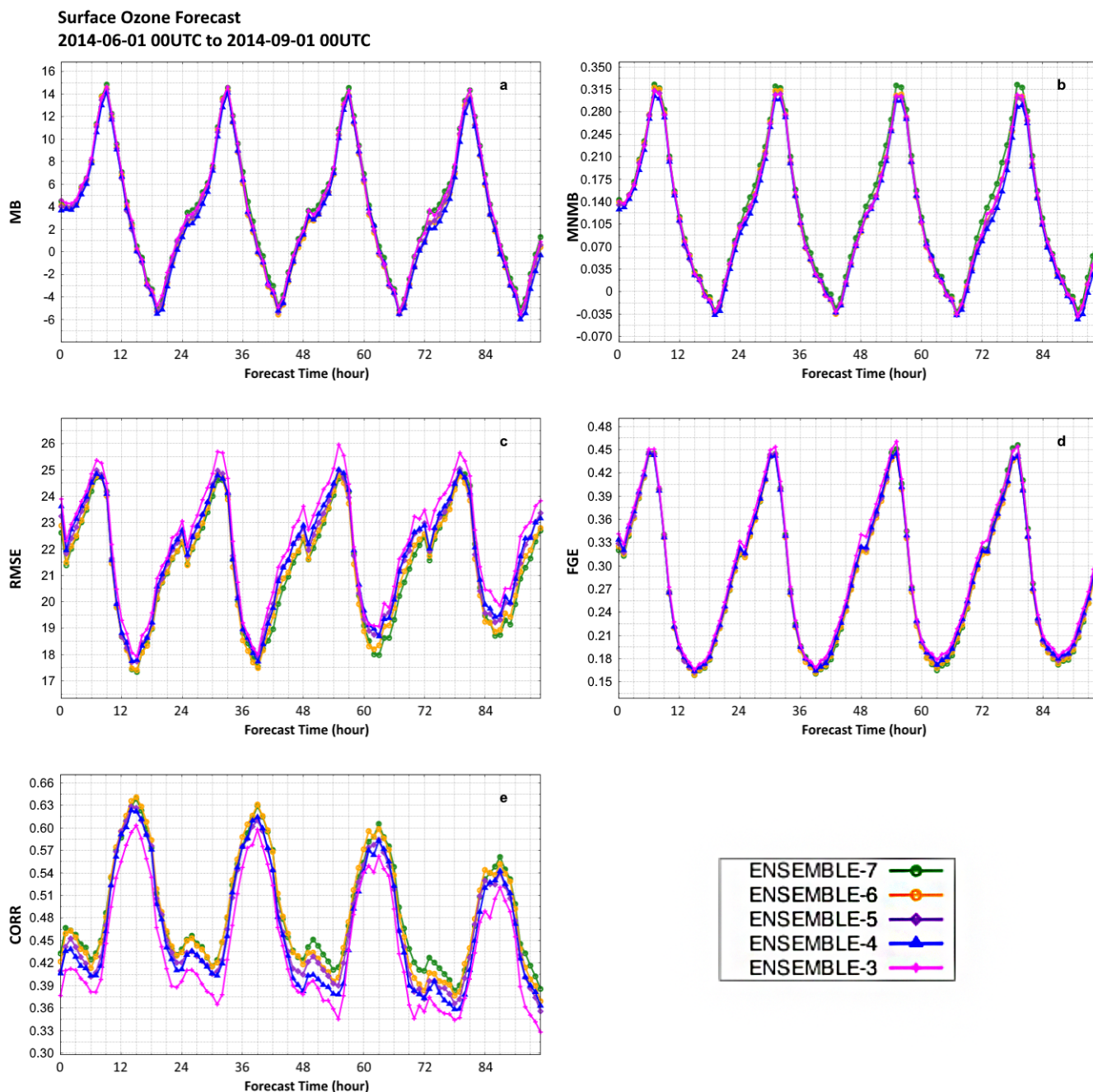
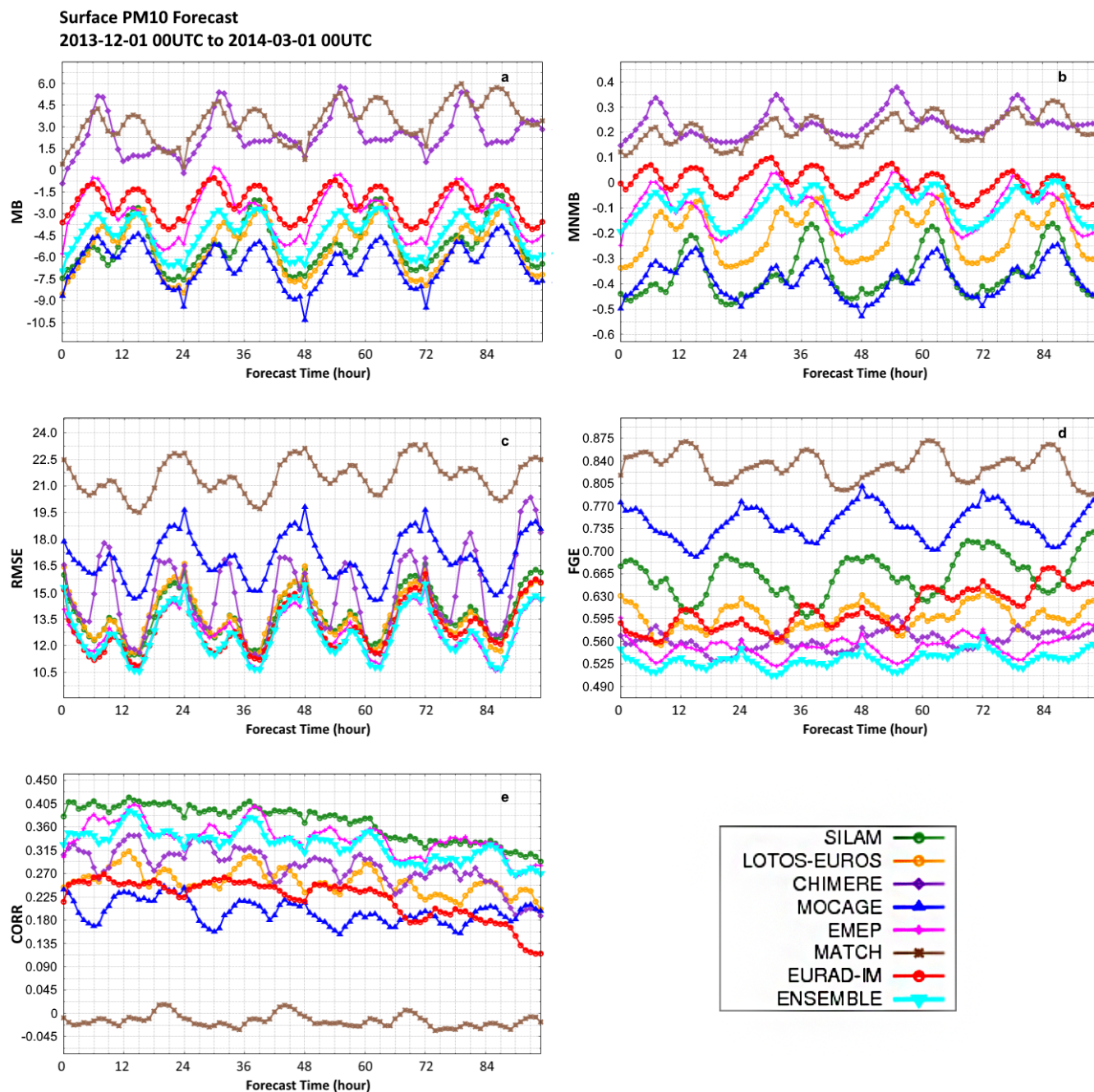**Figure 7.** Statistical indicators (see Appendix A) for ozone as a function of the forecast time (in hours) for an ensemble of between seven and three models compared to the hourly surface station measurements available for the period from 1 June 2014 at 00:00 UTC to 1 September 2014 at 00:00 UTC over the MACC-II European domain: **(a)** MB (in $\mu g\,m^{-3}$), **(b)** MNMB, **(c)** RMSE (in $\mu g\,m^{-3}$), **(d)** FGE and **(e)** correlation.

SEMBLE correlation 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 modelling 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 forecast models are optimized to run in short times. This constrains the level of detail of aerosol processes that can be afforded.

**Figure 8.** Statistical indicators (see Appendix A) for $PM_{10}$ as a function of the forecast time (in hours) for the seven models and the ENSEMBLE compared to the hourly surface station measurements available for the period from 1 December 2013 at 00:00 UTC to 1 March 2014 at 00:00 UTC over the MACC-II European domain: **(a)** MB (in $\mu g\,m^{-3}$), **(b)** MNMB, **(c)** RMSE (in $\mu g\,m^{-3}$), **(d)** FGE and **(e)** correlation.

## 3.5 Example of the specific evaluation for the Mediterranean area

Within the European continent, the Mediterranean area is characterized by special features – high emission densities due to concentration of human activities in surrounding coastal areas, intense photochemistry, high background pollution, small-scale meteorology – that make air quality forecasting specially challenging. This is why work has been specifically carried out to evaluate the seven models and the ENSEMBLE in this region. This is complementary to the systematic daily and seasonal evaluation performed over the whole European continent. Its aim is not about scor-

ing the system but to achieve better scientific understanding of the behaviour of the seven models and the ENSEMBLE in the Mediterranean region. This work is based, firstly, on two high-resolution models run daily over eastern (Greece) and western (Spain) Mediterranean areas and surface station measurements that are not used in the operational MACC evaluation and, secondly, on scientific analyses of case studies.

For the eastern Mediterranean area, the LAP-AUTH forecasting system is run daily. It consists of the Weather Research and Forecasting mesoscale meteorological model (WRF version 3.2) (Skamarock et al., 2008) and the chemistry-transport model Comprehensive Air quality Model with extensions (CAMx version 5.30) (ENVIRON, 2010). The anthropogenic 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 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 NEMO (Natural Emission MOdel) (Markakis et al., 2009). Wind erosion dust, sea salt and biogenic NMVOC emissions are calculated using the WRF model meteorology. The air quality forecasting system derives meteorological initial and boundary conditions from the operational 12:00 UTC forecast of ECMWF while chemical boundary conditions derived from the IFS–MOZART global model forecast and replaced by C-IFS from September 2014. The domain of the WRF–CAMx implementation is the south-eastern Europe/eastern Mediterranean region 18–30° N, 34.9–44.5° E. The grid resolution is $10 \times 10\,\text{km}^2$. The air quality modelling system runs on a daily basis in order to produce 72 h air quality forecasts. For the verification, the WRF-CAMx, the ENSEMBLE and the seven models are compared with available air quality data from the GMEECC (Greek Ministry of Environment Energy and Climatic Change) air pollution monitoring network as well as from the background station of Finokalia, operated by the University of Crete (Greece).

AEMET (the Spanish Meteorological State Agency) runs daily a version of the MOCAGE (Josse et al., 2004) model at 0.05° horizontal resolution in the western Mediterranean coast, over a 48 h time range using the ENSEMBLE forecasts as chemical boundary conditions. Meteorological forcings for the high-resolution domain come from an operational HIRLAM run every 6 h at AEMET (Navascues et al., 2013). Emissions over land in this domain come from the GEMS-TNO inventory (Visschedijk et al., 2007). The domain is 44–36° N, 5° W–5° E. The ENSEMBLE has been compared to the AEMET forecasts and to observations from EMEP/GAW Spanish stations 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 been an advantage, has been selected and analysed. These comparisons show the

high variability of results between model forecasts depending on the location, time and day, whereas, sometimes, model forecast agreement is quite noticeable.

We are presenting here a brief summary of the analysis of the case study that occurred between 15 and 18 July 2013, when high values of ozone were measured in many Spanish air quality monitoring stations due to very strong solar radiation and high temperatures together with persistent anticyclonic conditions and very weak pressure gradients. Ozone concentrations at the surface above $140\,\mu\text{g m}^{-3}$ were not rare at the stations used in this period and values above $120\,\mu\text{g m}^{-3}$ were common. Figure 9 shows two maps with the 18 July 2013 ENSEMBLE and AEMET model at H+18 forecasts and the observations overplotted 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 smooth to capture all the small-scale features occurring in reality because of its horizontal resolution ($\sim 15\,\text{km}$). As an example, we can look at Fig. 9 in which the Madrid area has been magnified to observe how ozone values between 100 and $160\,\mu\text{g m}^{-3}$ were measured by different air quality networks (belonging to the Madrid Regional Authorities and Madrid City Council) whereas in the ENSEMBLE forecasts all the concentrations lie in the $100–120\,\mu\text{g m}^{-3}$ interval. In the same period, the AEMET forecasts provide values in this area with a higher spread, between 100 and $160\,\mu\text{g m}^{-3}$, which fit better to observations. Something similar can be observed in the eastern Spain area, also magnified in the same figure.

Illustrations of the results of the ENSEMBLE and AEMET models compared to three EMEP stations are given in Figs. 10–12 for ozone during summer. For Cap de Creus (Fig. 10), there is a good agreement between observations and models, but the two models have a wider 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-eastern coast of Spain. Therefore, ozone formed in 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 Mahón (Fig. 11), both models fit fairly well the observations but the observations have generally a wider diurnal cycle in concentrations, contrarily to Cap de Creus. The Mahón station is located in the small island of Menorca in the Balearic Archipelago. It is sometimes exposed to the pollution produced by ships entering the port early in the morning, leading to high $NO_x$ conditions and low ozone. This could explain the observed decrease of ozone at this time. This local effect is not captured by the two models. The San Pablo de Montes measurements (Fig. 12) show a different behaviour with generally higher concentrations than the ENSEMBLE and AEMET models. The discrepancy is particularly important during July. This can be explained by an underestimation of the local isoprene emissions in the two models. Isoprene concentrations measured at San Pablo de los Montes exhibit
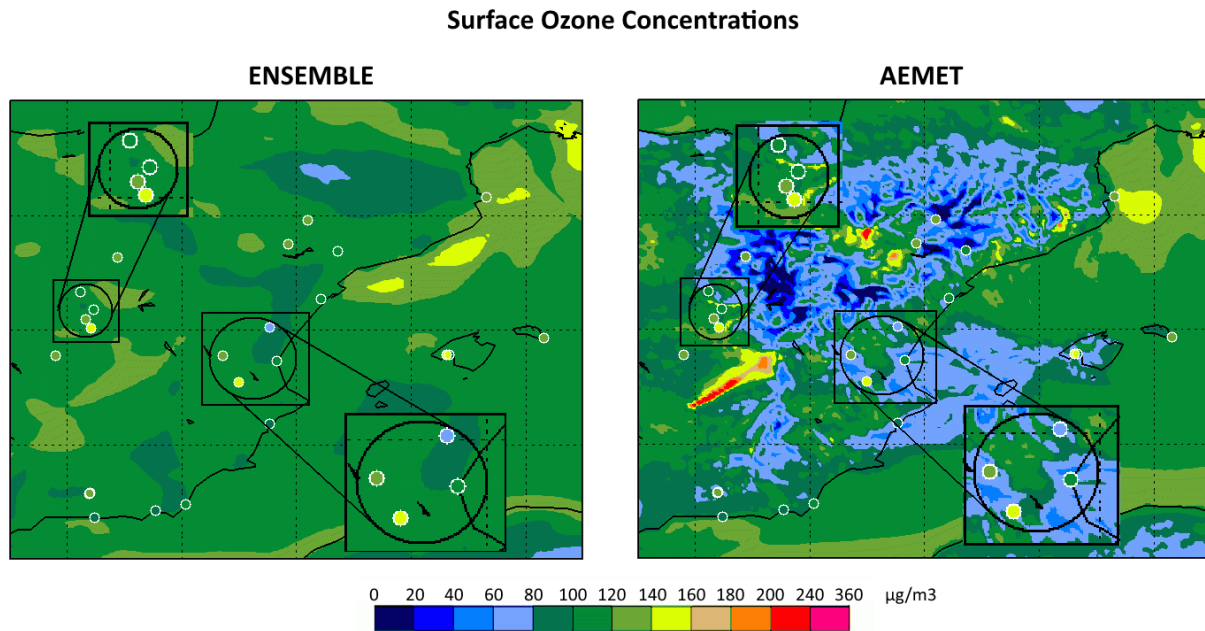
**Surface Ozone Concentrations**



**Figure 9.** ENSEMBLE (left) and AEMET (right) surface ozone concentrations (in $\mu g\, m^{-3}$) from a forecast (H+18) started on 18 July 2013 at 00:00 UTC 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.

higher values than other sites in the same area because of the oak vegetation surrounding the station.

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 seven 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.

Another product we have started to generate at the end of the project is the behaviour of forecasts of the seven models together with the ENSEMBLE and the AEMET forecasts 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 the ES10 station, which is located at Cap de Creus in the north-eastern corner of Spain (42.32° N, 3.32° E). We observe that the spread between the seven model forecasts in the H+24–H+48 forecast periods from 9 April 2014 is fairly low with most of the members producing similar forecasts. It changes quickly 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 the members at other locations on the same day and forecast time. More generally, this pattern with very different spreads (ranging from low to high) depends on the case studies: day, time period and

location. The analysis of the spread between different model forecasts in the same period can help modellers understand how their models behave in the Mediterranean area.

## 4 Conclusion and future developments

In this paper, we give an overview of the current state and performances of the forecasting system for European air quality that was put in place in the framework of the MACC project and continued during the MACC-II project and now in the MACC-III project. Its strength comes from the fact that it is based on an ensemble of seven state-of-the-art chemistry-transport models (CHIMERE, EMEP, EURAD-IM, LOTOS-EUROS, MOCAGE, MATCH, SILAM) that are developed and run by recognised institutes in Europe. It also relies on good quality inputs for meteorological forcings, emissions and chemical boundary conditions. It provides daily 4-day forecasts for six major pollutants ($O_3$, $NO_2$, $SO_2$, CO, $PM_{10}$ and $PM_{2.5}$) and birch pollen during the pollen season, as well as additional species for downscaling air quality modelling purposes. The production also includes hourly analysis for the previous day. Daily statistical performances of the forecasts against available European air quality monitoring stations are processed daily, weekly and every 3 months, giving an objective assessment of the products to users. They are also used to monitor the seasonal and yearly evolutions of the forecast scores.

Because of the resolution of the seven models (10–20 km), this system is not designed and does not attempt to fore-
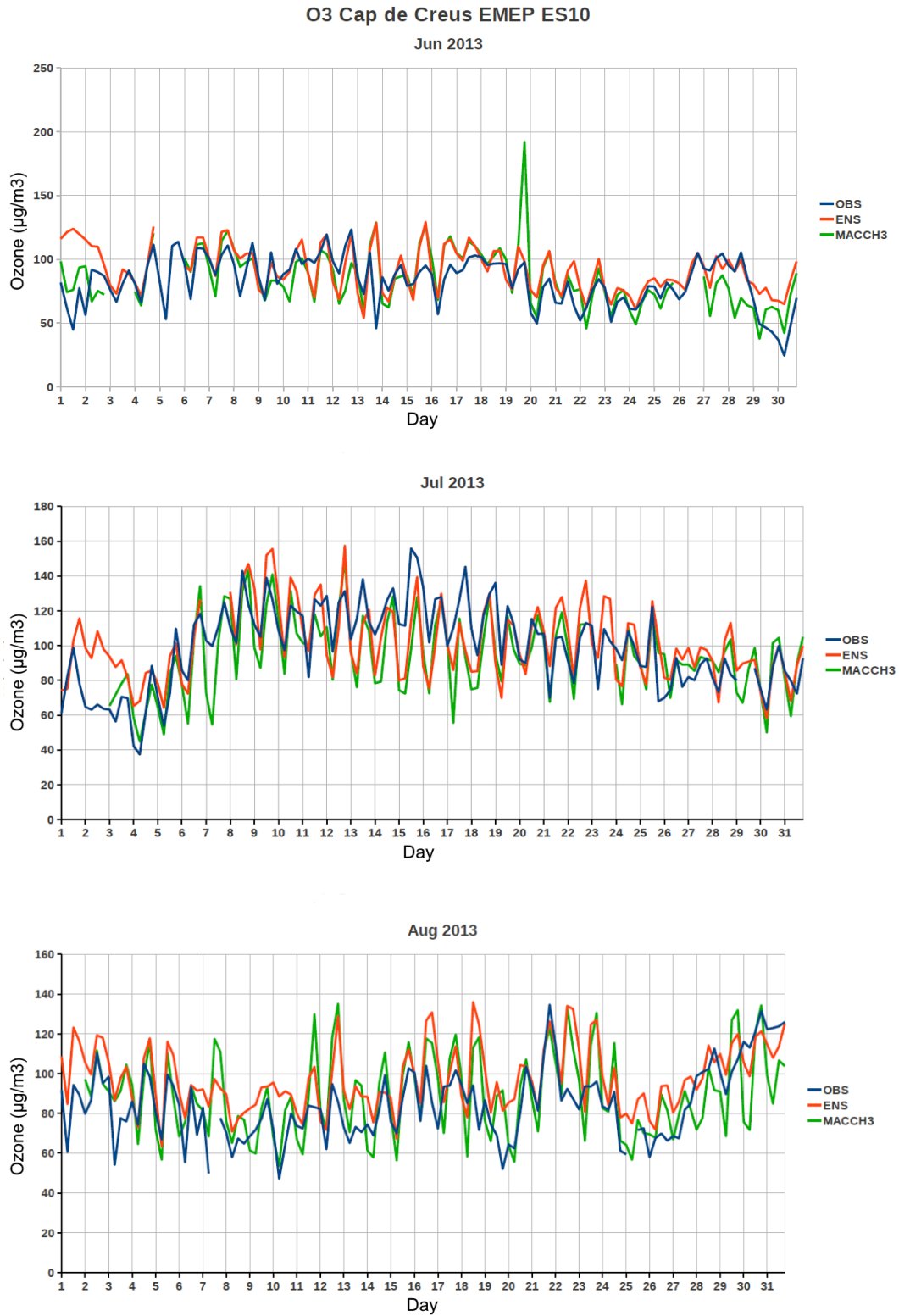
**Figure 10.** Ozone concentrations (in µg m$^{-3}$) as a function of days at Cap de Creus (42.32° N, 3.32° W). (Top panel) for June 2013, (middle panel) for July 2013 and (bottom panel) for August 2013. The blue line is for EMEP observations, the red line is for the ENSEMBLE and the green line for the AEMET model.
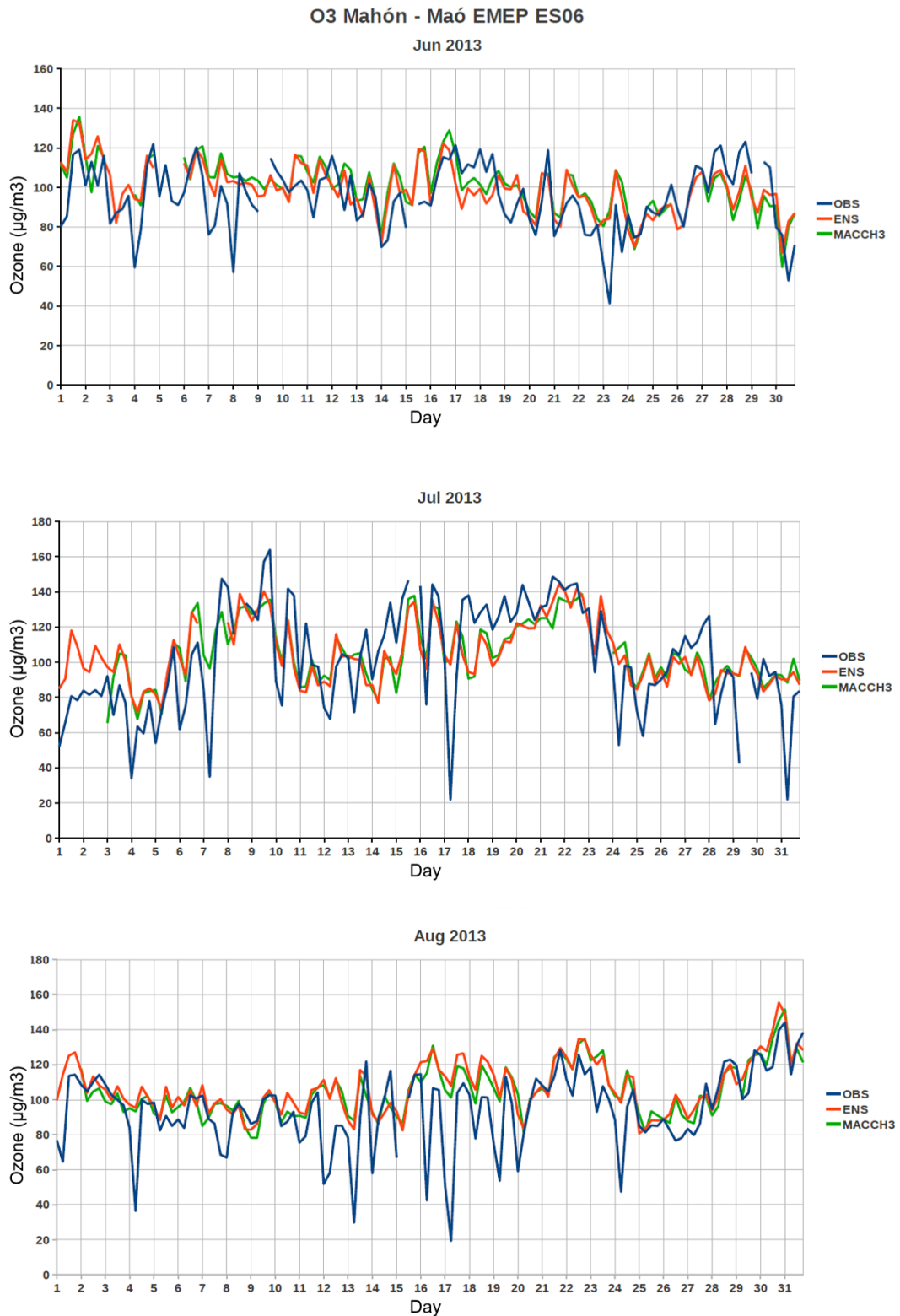
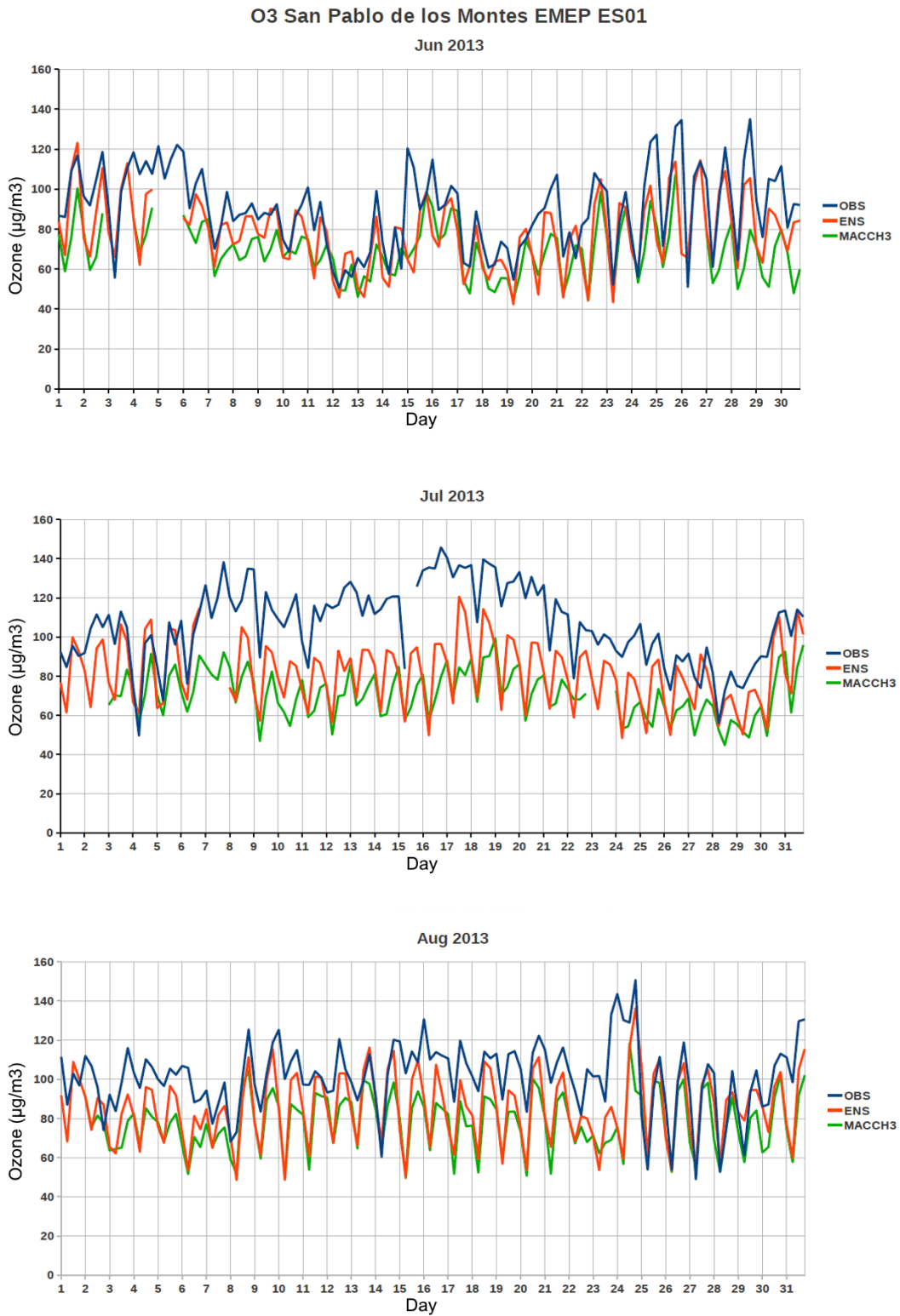**Figure 11.** As Fig. 10 but for Mahón (39.867° N, 4.32° W).

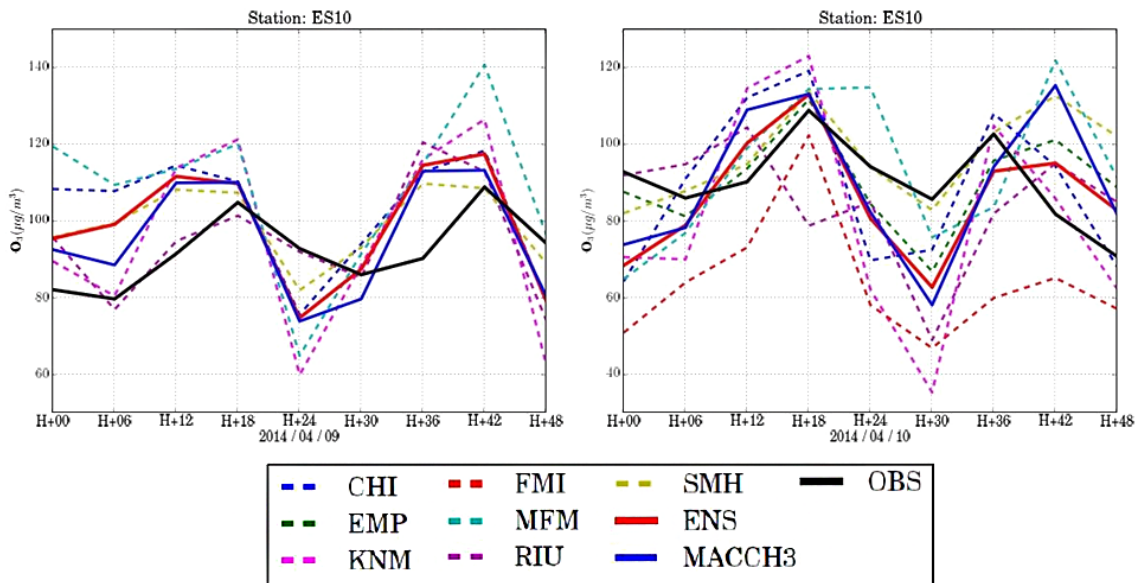**Figure 12.** As Fig. 10 but for San Pablo de los Montes (39.55° N, 4.35° E).

**Figure 13.** Ozone concentrations (in µg m$^{-3}$) from a 48 h forecast of ENSEMBLE, AEMET and the seven individual models at the ES10 EMEP air quality station which is located at Cap de Creus in the north-eastern corner of Spain (42.32° N, 3.32° E). The forecast started on 9 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 and AEMET models and observations, respectively.

cast 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 10 to 13 June 2014. On a seasonal basis, the seven models show good statistical performances for ozone in summer 2014 and the ensemble median outperforms any of the individual models. The normalised indicators, which 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 ozone peak is underestimated by the ENSEMBLE by about 4 µg m$^{-3}$ on average during summer 2014. The underestimation is larger during ozone episodes, such as during 9–15 June 2014, when the ENSEMBLE underestimates ozone daily maxima on average by about 10 µg m$^{-3}$. Comparing to local surface station measurements within the pollution episode areas, the ENSEMBLE low bias is often between 30 and 50 µg m$^{-3}$ but one of the models is often close to the observed values. For PM$_{10}$, there is a negative bias of the ENSEMBLE with MB ∼ −4.5 µg m$^{-3}$ and MNMB ∼ −0.1, on average. The ENSEMBLE FGE is larger than for ozone (∼ 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 seven models for the last winter of the MACC-II period (winter 2013–2014). This is related to different levels of complexity in the representation of aerosols in the models. PM consists of regulatory pollutants that are difficult to forecast mainly because of uncertainties in primary

aerosol emissions, our partial knowledge on secondary organic aerosol processes and the constraint of timely production that prevents using very sophisticated representations of secondary aerosols. A possible improvement to the ensemble median performances at low cost 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 evaluation of the seven models and of the ENSEMBLE is also done for the Mediterranean region because of its specificities (emissions, population, topography, meteorology, photochemistry). Another important point to note is that major efforts have been put during MACC-II towards the full operationalisation of the system in order to improve its robustness.

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 users' feedbacks. In the very short term, the ENSEMBLE analysis that is only provided for ozone will be extended to NO$_2$ from January 2015 and verification statistics with independent data will be produced. A shift to an earlier ENSEMBLE analysis production time, at 11:00 UTC, is also planned from early 2015 following the users' recommendation. One planned change in the mid-term will be to have all individual models run at a ∼ 10 km horizontal resolution. This should improve the performances of the system compared to observations. Also, the regional production benefits and will continue to benefit from the evolutions and improvements of the global production, such as from the use of the

newly operated C-IFS (fully 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 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 evaluation is done on the basis of the available NRT measurements from the European AQ monitoring network. Stations used are selected to be as representative as possible of the model horizontal resolution, by retaining only classes 1–5 from the Joly and Peuch (2012) classification. There is ongoing work to improve the station selection, still based on Joly and Peuch's classification, by determining the best class ranges to be used individually for each of the six pollutants ($O_3$, $NO_2$, $SO_2$, CO, $PM_{10}$, $PM_{2.5}$) based on pollutant lifetime.

Continuous research is pursued to improve the seven individual models and their assimilation systems. In particular, there is an important effort for the use of new satellite data or combinations of satellite data with surface measurements in the assimilation systems. Also, there is ongoing work on ensemble methods in order to extract as much value as possible from the seven model forecasts. Alternative methods to the median are currently tested: 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 alternative methods applied to the MACC-II multi-model ensemble will be the subject of a forthcoming paper. Another goal in MACC-II was the start of research and developments for the modelling of $CO_2$ in the regional models in view of potential future high-resolution surface $CO_2$ flux inversion products over Europe. This work will be pursued.

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 times 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 allow for earlier production of the analyses with the goal of using the analyses as the initial state for the forecasts.

Other studies will be conducted on the possibility to provide complementary indicators such as the exceedances in ozone or $PM_{10}$. In future, the production could be extended to types of pollens other than birch. There are currently some developments to test olive, grass and ambrosia pollens based on work done at the Finnish Meteorological Institute. Also, the 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 of aerosols.

## Appendix A: Statistical indicators

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

The mean bias captures the average deviations between two data sets and is defined as

$$MB = \frac{1}{N} \sum_i (f_i - o_i),$$

where $f_i$ and $o_i$ are the forecast value at the observation location and the observation value, respectively.

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 squaring operation. Especially in cases where prominent outliers occur, the usefulness of the RMSE is questionable and the interpretation becomes more difficult. MB and RMSE are not dimensionless variables but have the same dimension as the modelled/observed quantity and require knowledge of typical mean values. By scaling the MB and RMSE to the observations, these metrics can be made relative, dimensionless, and hence more appropriate for use as a score. This is relevant when comparing the bias and RMSE of atmospheric species whose concentrations can vary by orders of magnitude. This is why the modified normalised mean bias (MNMB) and the fractional gross error (FGE) are also used. MNMB is defined as

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

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

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

$$R = \frac{\frac{1}{N} \sum_i (f_i - \bar{f})(o_i - \bar{o})}{\sigma_f \sigma_o},$$

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

# References

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.

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, 18075–18084, 2001.

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, doi:10.5194/gmd-8-171-2015, 2015.

Andersson-Sköld, Y. and Simpson, D.: Comparison of the chemical schemes of the EMEP MSC-W and IVL photochemical trajectory models, Atmos. Environ., 33, 1111–1129, 1999.

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., Solomos, S., Sørensen, B., Tsegas, G., Vignati, E., Vogel, B., and Zhang, Y.: Online coupled regional meteorology chemistry models in Europe: current status and prospects, Atmos. 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 multi-component data assimilation experiment directed to sulphur dioxide and sulphate over Europe, Atmos. Environ., 43, 1622–1631, doi:10.1016/j.atmosenv.2008.12.005, 2008.

Barré, J., El Amraoui, L., Ricaud, P., Lahoz, W. A., Attié, J.-L., Peuch, V.-H., Josse, B., and Marécal, V.: Diagnosing the

transition layer at extratropical latitudes using MLS $O_3$ and MOPITT CO analyses, Atmos. Chem. Phys., 13, 7225–7240, doi:10.5194/acp-13-7225-2013, 2013.

Barré, J., Peuch, V.-H., Lahoz, W. A., Attié, J.-L., Josse, B., Piacentini, A., Emerenko, M., Dufour, G., Nédélec, P., von Clarmann, T., and El Amraoui, L.: Combined data assimilation of ozone tropospheric columns and stratospheric profiles in a high-resolution CTM, Q. J. Roy. Meteorol. Soc., 140, 966–981, doi:10.1002/qj.2176, 2014.

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, doi:10.5194/acp-10-10067-2010, 2010.

Benedetti, A., Morcrette, J.-J., Boucher, O., Dethof, A., Engelen, R. J., Fisher, M., Flentje, H., Huneeus, N., Jones, L., Kaiser, J. W., Kinne, S., Mangold, A., Razinger, M., Simmons, A. J., and Suttie, M.: Aerosol analysis and forecast in the European Centre for Medium-Range Weather Forecasts Integrated Forecast System. 2. Data assimilation: J. Geophys. Res., 114 , D13205, doi:10.1029/2008JD011115, 2009.

Bergström, R., Denier van der Gon, H. A. C., Prévôt, A. S. H., Yttri, K. E., and Simpson, D.: Modelling of organic aerosols over Europe (2002–2007) using a volatility basis set (VBS) 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.

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, Atmos. Environ., 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, J. Atmos. Chem., 61, 175–202, 2009.

Bott, A.: A positive definite advection scheme obtained by nonlinear renormalization of the advective fluxes, Mon. Weather Rev., 117, 1006–1016, 1989.

Bousserez, N., Attié, J.-L., Peuch, V.-H., Michou, M., Pfister, G., Edwards, D., Avery, M., Sachse, G., Browell, E., and Ferrare, E.: Evaluation of MOCAGE chemistry and transport model during the ICARTT/ITOP experiment, J. Geophys. Res., 112, D120S42, doi:10.1029/2006JD007595, 2007.

Burridge, D. M. and Gadd, A. J.: The Meteorological Office Operational 10-level numerical weather prediction model, Sci. Pap., 34, UK Meteorological Office, 1977.

Carter, W. P. L: Condensed atmospheric photo oxidation mechanism for isoprene, Atmos. Environ., 30, 4275–4290, 1996.

Colette, A., Granier, C., Hodnebrog, Ø., Jakobs, H., Maurizi, A., Nyiri, A., Bessagnet, B., D'Angiola, A., D'Isidoro, M., Gauss, M., Meleux, F., Memmesheimer, M., Mieville, A., Rouïl, 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, 11657–11678, doi:10.5194/acp-11-11657-2011, 2011.

Curier, R.L., Timmermans, R., Calabretta-Jongen, S., Eskes, H., Segers, A., Swart, D., and Schaap, M.: Improving ozone forecasts over Europe by synergistic use of the LOTOS-EUROS chemical transport model and in-situ measurements, Atmos. Environ., 60, 217–226, doi:10.1016/j.atmosenv.2012.06.017, 2012.

Cuvelier, C., Thunis, P., Vautard, R., Amann, M., Bessagnet B., Bedogni, M., Berkowicz, R., Brandt, J., Brocheton, F., Builtjes, P., Coppalle, A., Denby, B., Douros, G., Graf, A., Hellmuth, O., 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., Vignati, E., Volta, M., White, L., Wind, P., and Zuber A.: CityDelta: A model intercomparison study to explore the impact of emission reductions in European cities in 2010, Atmos. Environ., 41, 189–207, doi:10.1016/j.atmosenv.2006.07.036, 2007.

Delle Monache, L., Deng, X., Zhou, Y., and Stull, R.: Ozone ensemble forecasts: 1. A new ensemble design, J. Geophys. Res., 111, D05307, doi:10.1029/2005JD006310, 2006.

De Ruyter de Wildt, M., Eskes, H., Manders, A., Sauter, F., Schaap, M., Swart, D., and van Velthoven, P.: Six-Day PM10 Air Quality Forecasts For The Netherlands With The Chemistry Transport Model Lotos-Euros, Atmos. Environ., 45, 5586–5594, doi:10.1016/j.atmosenv.2011.04.049, 2011.

Dufour, A., Amodei, M., Ancellet, G., and Peuch, V.-H.: Observed and modelled "chemical weather" during ESCOMPTE, Atmos. Res., 74, 161–189, 2004.

Ebel, A., Jakobs, H., Memmesheimer, M., Elben, H., and Feldmann, H.: Numerical forecast of air pollution: advances and problems, Advances in Air Pollution Modeling for Environmental Security, Springer, doi:10.1007/1-4020-3351-6_14, 2005.

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 $O_3$ and $N_2O$ measurements: $N_2O$ as a dynamical tracer. Q. J. R. Meteorol. Soc., 134, 217–228, doi:10.1002/qj.191, 2008.

Elbern, H., Strunk, A., Schmidt, H., and Talagrand, O.: Emission rate and chemical state estimation by 4-dimensional variational inversion, Atmos. Chem. Phys., 7, 3749–3769, doi:10.5194/acp-7-3749-2007, 2007.

Elbern, H., Strunk, A., Friese, E., and Nieradzik, L.: Assessment of Source/Receptor Relations by Inverse Modelling and Chemical Data Assimilation, in: Persistent Pollution Past, Present and Future School of Environmental Research – Helmholtz-Zentrum Geesthacht, edited by: Quante, M., Ebinghaus, R., and Flöser, G., 1st Edn., ISBN 978-3-642-17420-9, 2011.

Emili, E., Barret, B., Massart, S., Le Flochmoen, E., Piacentini, A., El Amraoui, L., 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, doi:10.5194/acp-14-177-2014, 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. Pollut., 154, 448–461, 2008.

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, doi:10.5194/gmd-2-253-2009, 2009.

Flemming, J., Huijnen, V., Arteta, J., Bechtold, P., Beljaars, A., Blechschmidt, A.-M., Diamantakis, M., Engelen, R. J., Gaudel, A., Inness, A., Jones, L., Josse, B., Katragkou, E., Marecal, 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., 8, 975–1003, doi:10.5194/gmd-8-975-2015, 2015.

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.

Fountoukis, C. and Nenes, A.: ISORROPIA II: a computationally efficient thermodynamic equilibrium model for $K^+$–$Ca^{2+}$–$Mg^{2+}$–$NH_4^+$–$Na^+$–$SO_4^{2-}$–$NO_3^-$–$Cl^-$–$H_2O$ aerosols, Atmos. Chem. Phys., 7, 4639–4659, doi:10.5194/acp-7-4639-2007, 2007.

Friese, E. and Ebel, A.: Temperature dependent thermodynamic model of the system $H^+$-$NH_4^+$-$Na^+$-$SO_4^{2-}$-$NO_3^-$-$Cl^{-?}$-$H_2O$, J. Phys. Chem. A., 114, 11595–11631, 2010.

Fuhrer, J. and Booker, F.: Ecological issues related to ozone: agricultural issues, Environ. Int., 29, 141–154, 2003.

Galmarini, S., Kioutsioukis, I., and Solazzo, E.: E pluribus unum*: ensemble air quality predictions, Atmos. Chem. Phys., 13, 7153–7182, doi:10.5194/acp-13-7153-2013, 2013.

Galperin, M. V.: The Approaches to Correct Computation of Airborne Pollution Advection, in: Problems of Ecological Monitoring and Ecosystem Modelling, XVII, Gidrometeoizdat, St. Petersburg, 54–68, 2000 (in Russian).

Gaubert, B., Coman, A., Foret, G., Meleux, F., Ung, A., Rouil, L., Ionescu, A., Candau, Y., 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, doi:10.5194/gmd-7-283-2014, 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, doi:10.5194/acp-6-5445-2006, 2006.

Geiger, H., Barnes, I., Bejan, I., Benter, T., and Spittler, M.: The tropospheric degradation of isoprene: an updated module for the regional atmospheric chemistry mechanism, Atmos. Environ., 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 merging multiple satellite fire products, Biogeosciences, 7, 1171–1186, doi:10.5194/bg-7-1171-2010, 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.-Atmos., 91, 14367–14376, 1986.

Guenther, A., Karl, T., Harley, P., Wiedinmyer, C., Palmer, P. I., 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, doi:10.5194/acp-6-3181-2006, 2006.

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, 12609–12617, doi:10.1029/93JD00527, 1993.

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. Geophys. Res., 100, 8873–8892, doi:10.1029/94JD02950, 1995.

Guenther, A. B., Jiang, X., Heald, C. L., Sakulyanontvittaya, T., Duhl, T., Emmons, L. K., and Wang, X.: The Model of Emissions of Gases and Aerosols from Nature version 2.1 (MEGAN2.1): an extended and updated framework for modeling biogenic emissions, Geosci. Model Dev., 5, 1471–1492, doi:10.5194/gmd-5-1471-2012, 2012.

Guth, J., Josse, B., Marécal, V., and Joly, M.: Simulating Secondary Inorganic Aerosols using the chemistry transport model MOCAGE version R2.15.0, Geosci. Model Dev. Discuss., 8, 3593–3651, doi:10.5194/gmdd-8-3593-2015, 2015.

Hass, H., Jakobs, H. J., and Memmesheimer, M.: Analysis of a regional model (EURAD) near surface gas concentration predictions using observations from networks, Meteorol. Atmos. Phys., 57, 173–200, 1995.

Heimann, M. and Keeling, C. D.: A three-dimensional model of $CO_2$ tramsport based on observed winds, Model description and simulated trace experiment, in: Aspects of Climate Variability in the Pacific and Western Americas, edited by: Peterson, D. H., American Geophysical Union, Washington, DC, 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 Consortium: Toward a Monitoring and Forecasting System For Atmospheric Composition: The GEMS Project. Bull. Amer. Meteor. Soc., 89, 1147–1164, doi:10.1175/2008BAMS2355.1, 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, Bound.-Lay. Meteorol., 36, 201–209, 1986.

Holtslag, A. A. M., van Meigaard, E., and De Rooy, W. C: A comparison of boundary layer diffusion schemes in unstable conditions over land, Bound.-Lay. Meteorol., 76, 69–95, 1995.

Honoré, C., Rouil, L., Vautard, R., Beekmann, M., Bessagnet, B., Dufour, A., Elichegaray, C., Flaud, J.-M., Malherbe, L., Meleux, F., Menut, L., Martin, D., Peuch, A., Peuch, V.-H., 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.

Huijnen, V., Williams, J., van Weele, M., van Noije, T., Krol, M., Dentener, F., Segers, A., Houweling, S., Peters, W., de Laat, J., Boersma, F., Bergamaschi, P., van Velthoven, P., Le Sager, P., Eskes, H., Alkemade, F., Scheele, R., Nédélec, P., and Pätz, H.-W.: The global chemistry transport model TM5: description and evaluation of the tropospheric chemistry version 3.0, Geosci. Model Dev., 3, 445-473, doi:10.5194/gmd-3-445-2010, 2010.

Inness, A., Baier, F., Benedetti, A., Bouarar, I., Chabrillat, S., Clark, H., Clerbaux, C., Coheur, P., Engelen, R. J., Errera, Q., Flemming, J., George, M., Granier, C., Hadji-Lazaro, J., Huijnen, V., Hurtmans, D., Jones, L., Kaiser, J. W., Kapsomenakis, J., Lefever, K., Leitão, J., Razinger, M., Richter, A., Schultz, M. G., Simmons, A. J., Suttie, M., Stein, O., Thépaut, J.-N., Thouret, V., Vrekoussis, M., Zerefos, C., and the MACC team: 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 time-dependent background error covariance matrix on air quality analysis, Geosci. Model Dev., 5, 1075–1090, doi:10.5194/gmd-5-1075-2012, 2012.

Joly, M. and Peuch, V.-H.: Objective Classification of air quality monitoring sites over Europe, Atmos. Environ., 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.

Kaiser, J. W., Heil, A., Andreae, M. O., Benedetti, A., Chubarova, N., Jones, L., Morcrette, J.-J., Razinger, M., Schultz, M. G., Suttie, M., and van der Werf, G. R.: Biomass burning emissions estimated with a global fire assimilation system based on observed fire radiative power, Biogeosciences, 9, 527–554, doi:10.5194/bg-9-527-2012, 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, edited by: Burrows, J., Platt, U., and Borrell, P., Springer, doi:10.1007/978-3-642-14791-3, 2011.

Kioutsioukis, I. and Galmarini, S.: *De praeceptis ferendis*: good practice in multi-model ensembles, Atmos. Chem. Phys., 14, 11791–11815, doi:10.5194/acp-14-11791-2014, 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, 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.

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.

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, doi:10.5194/acp-14-10963-2014, 2014.

Kukkonen, J., Olsson, T., Schultz, D. M., Baklanov, A., Klein, T., Miranda, A. I., Monteiro, A., Hirtl, M., Tarvainen, V., Boy, M., Peuch, V.-H., Poupkou, A., Kioutsioukis, I., Finardi, S., Sofiev, M., Sokhi, R., Lehtinen, K. E. J., Karatzas, K., San José, R., Astitha, M., Kallos, G., Schaap, M., Reimer, E., Jakobs, H., and Eben, K.: A review of operational, regional-scale, chemical weather forecasting models in Europe, Atmos. Chem. Phys., 12, 1–87, doi:10.5194/acp-12-1-2012, 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. Environ., 92, 348–358, 2014.

Lahoz, W. A., Geer, A. J., Bekki, S., Bormann, N., Ceccherini, S., Elbern, H., Errera, Q., Eskes, H. J., Fonteyn, D., Jackson, D. R., Khattatov, B., Marchand, M., Massart, S., Peuch, V.-H., Rharmili, S., Ridolfi, M., Segers, A., Talagrand, O., Thornton, H. E., Vik, A. F., and von Clarmann, T.: The Assimilation of Envisat data (ASSET) project, Atmos. Chem. Phys., 7, 1773–1796, doi:10.5194/acp-7-1773-2007, 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 1994 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.-Atmos., 99, 8183–8195, 1994.

Li, Y. P., Elbern, H., Lu, K. D., Friese, E., Kiendler-Scharr, A., Mentel, Th. F., Wang, X. S., 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, 6289–6304, doi:10.5194/acp-13-6289-2013, 2013.

Liu, D. C. and Nocedal, J.: On the limited memory BFGS method for large scale optimization, Math. Programming, 45, 503–528, 1989.

Louis, J.-F.: A parametric model of vertical eddy fluxes in the atmosphere, Bound.-Lay. Meteorol., 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.-Atmos., 105, 22255–22267, 2000.

Markakis, K., Giannaros, T., Poupkou, A., Liora, N., Melas, D., Sofiev, M. and Soares, J.: Evaluating the impact of particle emissions from natural sources in the Balkan region, European Aerosol Conference 2009, 6–9 September 2009, Karlsruhe, Germany, 2009.

Markakis, K., Katragkou, E., Poupkou, A. and Melas, D.: MOSESS: A new emission model for the compilation of model-ready emission inventories. Application in a coal mining area in Northern Greece, Environ. Model. Assess., 18, 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, 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. Environ. Pollut., 22, 108–132, 2004.

Menut, L. and Bessagnet, B.: Atmospheric composition forecasting in Europe, Ann. Geophys., 28, 61–74, doi:10.5194/angeo-28-61-2010, 2010.

Menut, L., Goussebaile, A., Bessagnet, B., Khvorostiyanov, D., and Ung, A.: Impact of realistic hourly emissions profiles on modelled air pollutants concentrations, Atmos. Environ., 49, 233–244, doi:10.1016/j.atmosenv.2011.11.057, 2012.

Menut, L., Bessagnet, B., Khvorostyanov, D., Beekmann, M., Blond, N., Colette, A., Coll, I., Curci, G., Foret, G., Hodzic, A., Mailler, S., Meleux, F., Monge, J.-L., Pison, I., Siour, G., 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, 2013a.

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. Geophys. Res.-Atmos., 118, 6505–6520, doi:10.1002/jgrd.50313, 2013b.

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, Environ. Pollut., 162, 176–189, 2012.

Monteiro, A., Ribeiro, I., Tchepel, O., Sá, E., Ferreira, J., Carvalho, A., Martins, V., Strunk, A., Galmarini, S., Elbern, H., Schaap, M., Builtjes, P., Miranda, A. I., and Borrego, C.: Bias Correction Techniques to Improve Air Quality Ensemble Predictions: Focus on O$_3$ and PM Over Portugal, Environ. Model. Assess., 18, 533–546, doi:10.1007/s10666-013-9358-2, 2013.

Morcrette, J.-J., Boucher, O., Jones, L., Salmond, D., Bechtold, P., Beljaars, A., Benedetti, A., Bonet, A., Kaiser, J. W., Razinger, M., Schulz, M., Serrar, S., Simmons, A. J., Sofiev, M., Suttie, M., Tompkins, A. M., and Untch, A.: Aerosol analysis and forecast in the European Centre for Medium-Range Weather Forecasts Integrated Forecast System. 1. Forward modelling: J. Geophys. Res., 114, D06206, doi:10.1029/2008JD011235, 2009.

Navascues, B., Calvo, J., Morales, G., Santos, C., Callado, C., Cansado, A., Cuxart, J., Diez, M., del Rio, P., Escriba, P.,

Garcia-Colombo, O., García-Moya, J. A., Geijo, C., Gutier-rez, E., Hortal, M., Martinez, I., Orfila, B., Parodi, J. A., Ro-driguez, E., Sánchez-Arriola, J., Santos-Atienza, I., and Simarro, J.: Long term verification of HIRLAM and ECMWF forecasts over Southern Europe. History and perspectives of Numerical Weather Prediction at AEMET, Atmos. Res., 125–126, 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, 587–598, 2005.

Nieradzik, L. P.: Application of a high dimensional model rep-resentation on the atmospheric aerosol module MADE of the EURAD-CTM, Master Thesis, Institut für Geophysik und Me-teorologie der Universität zu Köln, 2005.

Nocedal, J.: Updating quasi-Newton matrices with limited storage, Math. Comput., 35, 773–782, 1980.

Parrish, D. F. and Derber, J. C.: The national meteorological center's spectral statistical-interpolation analysis system, Mon. Weather Rev., 120, 1747–1763, 1992.

Petroff, A. and Zhang, L.: Development and validation of a size-resolved particle dry deposition scheme for application in aerosol transport models, Geosci. Model Dev., 3, 753–769, doi:10.5194/gmd-3-753-2010, 2010.

Potempski, S., Galmarini, S., Riccio, A., and Giunta, G.: Bayesian model averaging for emergency response atmospheric disper-sion 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.

Poupkou, A., Giannaros, T., Markakis, K., Kioutsioukis, I., Curci, G., Melas, D., and Zerefos, C.: A model for European Bio-genic Volatile Organic Compound emissions: Software devel-opment and first validation, Environ. Model. Softw., 25, 1845–1856, doi:10.1016/j.envsoft.2010.05.004, 2010.

Rabitz, H. and Alis, Ö. F.: 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 evalu-ation 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 transport model, 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 multi-scale air quality (CMAQ) modeling system, EPA 600/R-99-030, EPA, 1999.

Rouïl, 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 Europe, Bull. Am. Meteor. Soc., 90, 73–83, doi:10.1175/2008BAMS2390.1, 2009.

Sandu, A. and Sander, R.: Technical note: Simulating chemical systems in Fortran90 and Matlab with the Kinetic PreProcessor KPP-2.1, Atmos. Chem. Phys., 6, 187–195, doi:10.5194/acp-6-187-2006, 2006.

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. Environ., 37, 5083–5096, 2003.

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., 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., and 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, 2005.

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.

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 evalua-tion 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.

Simpson, D., Winiwarter, W., Börjesson, G., Cinderby, S., Ferreiro, A., Guenther, A., Hewitt, C. N., Janson, R., Khalil, M. A. K., Owen, S., Pierce, T. T., Puxbaum, H., Shearer, M., Skiba, U., Steinbrecher, R., Tarrason, L., and Oquist, M. G.: Inventorying emissions from nature in Europe, J. Geophys. Res.-Atmos., 104, 8113–8152, 1999.

Simpson, D., Fagerli, H., Jonson, J. E., Tsyro, S., Wind, P., and Tuovinen, J.-P.: Transboundary Acidification, Eutrophica-tion and Ground Level Ozone in Europe, Part 1: Unified EMEP Model Description, EMEP Report 1/2003, 2003.

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

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Duda, M. G., Huang, X. Y., Wang, W., and Powers, J. G.: A description of the advanced researcher WRF version 3. NCAR Technical Note, NCAR/TN-475+STR, June 2008, Boulder, Col-orado, USA, 125 pp, 2008.

Sofiev, M.: A model for the evaluation of long-term airborne pollu-tion transport at regional and continental scales, Atmos. Environ., 34, 2481–2493, 2000.

Sofiev, M.: Extended resistance analogy for construction of the ver-tical diffusion scheme for dispersion models, J. Geophys. Res., 107, ACH10-1–ACH10-8 doi:10.1029/2001JD001233, 2002.

Sofiev, M., Siljamo, P., Valkama, I., Ilvonen, M., and Kukko-nen, J.: A dispersion modelling system SILAM and its eval-

uation against ETEX data, Atmos. Environ., 40, 674–685, doi:10.1016/j.atmosenv.2005.09.069, 2006.

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: NATO Science for Piece and Security Serties C: Environmental Security, edited by: Borrego, C. and Miranda, A. I., Air Pollution Modelling and Its Application, XIX. SPRINGER-VERLAG BERLIN, 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., 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, D21302, doi:10.1029/2010JD014713, 2011.

Sofiev, M., Berger, U., Prank, M., Vira, J., Arteta, J., Belmonte, J., Bergmann, K.-C., Chéroux, F., Elbern, H., Friese, E., Galan, C., Gehrig, R., Khvorostyanov, D., Kranenburg, R., Kumar, U., Marécal, V., Meleux, F., Menut, L., Pessi, A.-M., Robertson, L., Ritenberga, O., Rodinkova, V., Saarto, A., Segers, A., Severova, E., Sauliene, I., Siljamo, P., Steensen, B. M., Teinemaa, E., Thibaudon, M., and Peuch, V.-H.: MACC regional multi-model ensemble simulations of birch pollen dispersion in Europe, Atmos. Chem. Phys., 15, 8115–8130, doi:10.5194/acp-15-8115-2015, 2015.

Solazzo, E., Bianconi, R., Vautard, R., Appel, K. W., Moran, M. D., Hogrefe, C., Bessagnet, B., Brandt, J., Christensen, J. H., Chemel, C., Coll, I., Denier van der Gon, H., Ferreira, J., Forkel, R., Francis, X. V., Grell, 5 G., Grossi, P., Hansen, A. B., Jericevic, A., Kraljevic, L., Miranda, A. I., Nopmongcol, U., Pirovano, G., Prank, M., Riccio, A., Sartelet, K. N., Schaap, M., Silver, J. D., Sokhi, R. S., Vira, J., Werhahn, J., Wolke, R., Yarwood, G., Zhang, J., Rao, S. T., and Galmarini, S.: Model evalution and ensemble modelling of surface-level ozone in Europe and North America in the context of AQMEII, Atmos. Environ., 53, 60–74, 2012a.

Solazzo, E., Bianconi, R., Pirovano, G., Matthias, V., Vautard, R., Moran, M. D., Wyat Appel, K., Bessagnet, B., Brandt, J., Christensen, J. H., Chemel, C., Coll, I., Ferreira, J., Forkel, R., Francis, X. V., Grell, G., Grossi, P., Hansen, A. B., Miranda, A. I., Nopmongcol, U., Prank, M., Sartelet, K. N., Schaap, M., Silver, J. D., Sokhi, R. S., Vira, J., Werhahn, J., Wolke, R., Yarwood, G., Zhang, J., Rao, S. T., and Galmarini, S.: Operation model evaluation for particulate matter in Europe and North America in the context of AQMEII, Atmos. Environ., 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, J. Int. Environ. Sci., 9, 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 $PM_{10}$ concentrations, Atmos. Env., 42, 4567–4588, doi:10.1016/j.atmosenv.2008.01.068, 2008.

Stockwell, W. R., Kirchner, F., Kuhn, M., and Seefeld, S.: A new mechanism for regional atmospheric chemistry modeling, J. Geophys. Res.-Atmos., 102, 25847–25879, 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, 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.

Turquety, S., Menut, L., Bessagnet, B., Anav, A., Viovy, N., Maignan, F., and Wooster, M.: APIFLAME v1.0: high-resolution fire emission model and application to the Euro-Mediterranean region, Geosci. Model Dev., 7, 587–612, doi:10.5194/gmd-7-587-2014, 2014.

Van Loon, M., Vautard, R., Schaap, M., Bergstrom, R., Bessagnet, B., Brandtn J., Builtjes, P. J. H, Christensen, J. H., Cuvelier, C., Graff, A., Jonson, J. E., Krol, M., Langner, J., Roberts, P., Rouil, L., Stern, R., Tarrason, L., Thunis, P., Vignati, E., and White, L.: Evaluation of long-term ozone simulations from seven regional air quality models and their ensemble, Atmos. Env., 41, 2083–2097, doi:10.1016/j.atmosenv.2006.10.073, 2007.

Van Ulden, A. P. and Holtslag, A. A. M.: Estimation of atmospheric boundary layer parameters fort diffusion applications, J. Climate. Appl. Met., 24, 1196–1207, 1975.

Vautard, R., Builtjes, P. H. J., Thunis, P., Cuvelier, C., 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. Environ., 41, 173–188, 2007.

Vira, J. and Sofiev, M.: On variational data assimilation for estimating the model initial conditions and emission fluxes for short-term forecasting of SOx concentrations, Atmos. Environ., 46, 318–328, doi:10.1016/j.atmosenv.2011.09.066, 2012.

Vira, J. and Sofiev, M.: Assimilation of surface $NO_2$ and $O_3$ observations into the SILAM chemistry transport model, Geosci. Model Dev., 8, 191–203, doi:10.5194/gmd-8-191-2015, 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 2007-A-R0233/B, 2007.

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-15-1313-2015, 2015.

Walcek, C. J.: Minor flux adjustment near mixing ratio extremes for simplified yet highly accurate monotonic calculation of tracer advection, J. Geophy. Res., 105, 9335–9348, 2000.

Weaver, A. and Courtier, P.: Correlation modeling on the sphere using a generalized diffusion equation, Q. J. R. Meteorol. Soc., 127, 1815–1846, 2001.

Wesely, M.: Parameterization of surface resistances to gaseous dry deposition in regional-scale numerical models, Atmos. Environ., 23, 1293–1304, 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.

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 exchange, Biogeosciences, 9, 5261–5277, doi:10.5194/bg-9-5261-2012, 2012.

Williams, J. E., van Velthoven, P. F. J., and Brenninkmeijer, C. A. M.: Quantifying the uncertainty in simulating global tropospheric composition due to the variability in global emission estimates of Biogenic Volatile Organic Compounds, Atmos. Chem. Phys., 13, 2857–2891, doi:10.5194/acp-13-2857-2013, 2013.

Williamson, D. L. and Rasch, R. P.: Two-Dimensional Semi-Lagrangian Transport with Shape-Preserving Interpolation, Am. Meteorol. Soc., 117, 102–129, 1989.

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, doi:10.5194/acp-3-2067-2003, 2003.

Zhang, L., Gong, S., Padro, J., and Barrie, L.: A size-segregated particle dry deposition scheme for an atmospheric aerosol module, Atmos. Environ., 35, 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.

Zyryanov, D., Foret, G., Eremenko, M., Beekmann, M., Cammas, J.-P., D'Isidoro, M., Elbern, H., Flemming, J., Friese, E., Kioutsioutkis, I., Maurizi, A., Melas, D., Meleux, F., Menut, L., Moinat, P., Peuch, V.-H., Poupkou, A., Razinger, M., Schultz, M., Stein, O., Suttie, A. M., Valdebenito, A., Zerefos, C., Dufour, G., Bergametti, G., and Flaud, J.-M.: 3-D evaluation of tropospheric ozone simulations by an ensemble of regional Chemistry Transport Model, Atmos. Chem. Phys., 12, 3219–3240, doi:10.5194/acp-12-3219-2012, 2012.