Evaluation of air quality forecasting system FORAIR_IT over Europe and Italy at high resolution for year 2017

Mario Adani, Guido Guarnieri, Lina Vitali, Luisella Ciancarella, Ilaria D'Elia, Mihaela Mircea, Maurizio Gualtieri, Andrea Cappelletti, Massimo D'Isidoro, Gino Briganti, Antonio Piersanti, Milena Stracquadanio, Gaia Righini, Felicita Russo, Giuseppe Cremona, Maria Gabriella Villani, and Gabriele Zanini

ENEA, Italian National Agency for New Technologies, Energy and Sustainable Economic Development, Italy

Correspondence: Mario Adani (mario.adani@enea.it)

Abstract. Air pollution represents a global threat leading to large impacts on health and ecosystems and many European areas still show a poor air quality. Many measures and policies have been adopted in the past decades at European, national, regional and even local level and many tools have been developed to tackle this issue. Among these tools, the European Air Quality Directive places more emphasis on the use of models for air quality assessment and management. Within this context, air

5 quality forecasting systems play an important role in supporting decision makers when short-term actions are required to reduce human health risks by limiting population exposure. In this framework, at European level within the Copernicus Atmosphere Monitoring Service (CAMS), the regional air quality models participating to the service provide 4-day daily forecasts of the main atmospheric pollutants concentrations, in the lowest layers of the atmosphere. This work presents the development and the performances evaluation of FORAIR_IT, an high-resolution air quality forecasting system operating at both European (20

10 km) and Italian (4 km) scales. Its skill results are compared with CAMS_50 interim ensemble reanalysis (IRA) ones and a long lasting PM10 exceeding event, occurred in Emilia Romagna region in October 2017, is studied in more detail. Results show similar skill scores between FORAIT_IT and CAMS_50. Comparing the annual average of the monthly Root Mean Square Error Difference (RMSED) between FORAIT_IT first forecast day and CAMS_50, over the European domain the RMSED is 0.6,1.7,1.4 and 7.4 μg/m³ for daily mean PM10 and PM2.5 and daily maximum for NO_2 and O_3, respectively, while over

15 the Italian domain it is -1.2,0.3,-4.3 and 3.8 μg/m³. The importance of increasing model resolution in the region of interest is highlighted by the lower values of RMSED over Italy with respect to Europe. The results obtained by the detailed analysis of the PM10 exceeding event suggests the crucial role of the meteorological forcing in capturing both the timing and the intensity of the exceedances. As far as we know FORAIR_IT is the first forecasting system at high spatial resolution at Italian National level.

1 Introduction

Air pollution has negative impacts on human health and environment (Pope et al., 2019; WHO, 2015) and represents the biggest environmental risk to health (WHO, 2019). Since 1970, the European Union and its Member States have made many efforts to improve air quality by reducing anthropogenic emissions, through the improvement of industry, energy production, residential

and commercial heating, transport, waste treatment and agriculture. In spite of this, the last European Environmental Agency (EEA) report on Air Quality in Europe for 2017 (EEA, 2019) shows that, in many countries, in particular in urban areas, pollution levels continue to exceed the air quality standards stated in the two Air Quality Directives presently in force (EC, 2004, 2008). To tackle this issue, many tools have been developed and, in particular the European Air Quality Directive (EC,

5   2008) placed more emphasis on the use of models since they have the advantages of being cheaper than measurements and covering continuously and simultaneously large areas. In the last decades, real-time air quality forecasting (RT-AQF) systems were developed to predict short-term pollutant concentrations, in the days (1 to 5) after the prediction, and they are facing a great diffusion (Neal et al., 2014; Peng et al., 2018; Ha et al., 2019; Miskell et al., 2019). Knowing in advance the probability of high concentrations episodes allows to take actions, both at policy-administration level and at private and citizen level, for

10   limiting the population exposure, reducing the risks for human health. The first empirical approaches and statistical models were applied to forecast chemical state of the atmosphere in the 1970s (Zhang et al., 2012a, b). Advances in the knowledge of physical and chemical dynamics of pollution and in the capacity of computing facilities enabled the development of 3-dimensional numerical models, the so called chemical transport or air quality models, which evolved up to the present. They have been applied to operational forecast since mid-1990s. An exhaustive review of national operational air quality forecasting

15   systems available in Europe in 2012 was done by Kukkonen et al. (2012) in the framework of the European Cooperation in Science and Technology (COST) ES0602 action. Since 2009, over the European continent, a pre-operational analysis and forecasting system was developed during MACC (Monitoring Atmospheric Composition and Climate) and MACC-II (Monitoring Atmospheric Composition and Climate: Interim Implementation) European projects (Marécal et al., 2015). The forecasting system is operationally producing real time air quality predictions since 2015 (http://macc-raq-op.meteo.fr) in the framework of

20   the Copernicus Atmosphere Monitoring Service (CAMS) and currently includes nine numerical air quality models: CHIMERE (INERIS, France), EMEP (MET Norway), EURAD-IM (University of Cologne, Germany), LOTOS-EUROS (KNMI and TNO, Netherlands), MATCH (SMHI, Sweden), MOCAGE (METEO-FRANCE, France), SILAM (FMI, Finland), GEM-AQ (IEP-NRI, Poland) and DEHM (AARHUS UNIVERSITY, Denmark) (the latter two still not operational at the time of the present work). While this service is providing air quality forecast ensembles over Europe with the spatial resolution of 0.1º (ca. 10 km),

25   the national air quality forecasting systems are maintained or developed in order to predict pollution levels at higher spatial resolutions (up to 1 km) on a national domain. The higher resolutions allow to better capture the local features of weather and emissions, both anthropogenic and natural, and its more detailed representation of atmosphere allows a better understanding of forecast uncertainties due to model formulation and/or input data. Moreover, the operational air quality forecasting systems at national scale, due to higher spatial detail, may get higher skills in reproducing measured concentrations, given that monitoring

30   stations have a spatial representativeness of concentration which is usually limited to small areas. Even if these instruments require substantial efforts and continuous updates to state-of-the-art scientific knowledge, RT-AQF systems have a fundamental relevance in supporting mitigation actions for limiting the harmful effects of high pollution episodes on the surrounding population.

This article presents a new air quality forecasting system FORAIR_IT which provides on a daily basis the concentrations of

35   main pollutants in the three following days over Italy with a spatial resolution of 4km. The system is operational since 2017

and originally it started on the track of QualeAria forecasting system (http://www.qualearia.it; D'Allura et al.), operating since 2007, then taking an autonomous path linked also to the computing infrastructure of ENEA CRESCO. The new forecasting system, whose upgrades are now funded by the Italian Ministry of Environment, is maintained and operated by ENEA (Italian National Agency for New Technologies, Energy and Sustainable Economic Development) completing the modelling tools of

5   national air quality to support the implementation of the Air Quality Directive, as stated by the Italian law (Legislative Decree 155/2010). Section 2 describes the FORAIR_IT atmospheric modelling system and the used setup, while Section 3 shows the evaluation of forecasts at European and national levels for the year 2017. Section 4 discusses extensively a particulate matter (PM) episode in the Emilia Romagna Region, in the heart of Po Valley, an area often experiencing high pollution levels due to high anthropogenic emissions and unfavourable meteorological conditions. The last section summarizes the performances of

10   forecasting system and foreseen developments are discussed.

2   Air quality forecasting system: description and setup

The components of the modelling system are the same as in AMS-MINNI (Atmospheric Modelling System - Italian National Integrated Model to support the international negotiation on atmospheric pollution), whose performances have been evaluated in several studies over Italy (Mircea et al., 2014, 2016; Ciancarella et al., 2016; Vitali et al., 2019) and in the intercomparison

15   exercise EURODELTA-III over Europe (Bessagnet et al., 2014, 2016). The modelling system is based on the chemical transport model FARM - Flexible Air quality Regional Model - (Silibello et al., 2008, 2012; Gariazzo et al., 2007; Kukkonen et al., 2012; Adani et al., 2015), the meteorological driver RAMS - Regional Atmospheric Modelling System - (Cotton et al., 2003), the pre-processor SURFPro - SURFace-atmosphere interface PROcessor - (Arianet, 2011) and the emission processor EMMA - EMission MAnager - (Arianet, 2014).

FORAIR_IT forecasts on daily basis the concentrations in the three following days. Simulations are carried out, in two-way nesting mode, on the two domains shown in Fig. 1. The first one (EU) covers Europe at about 20kmx20km spatial resolution; the second one (IT) covers Italy at 4kmx4km resolution. This study focuses on forecast assessment of four pollutants (PM10, PM2.5, O_3, NO_2) over both domains.
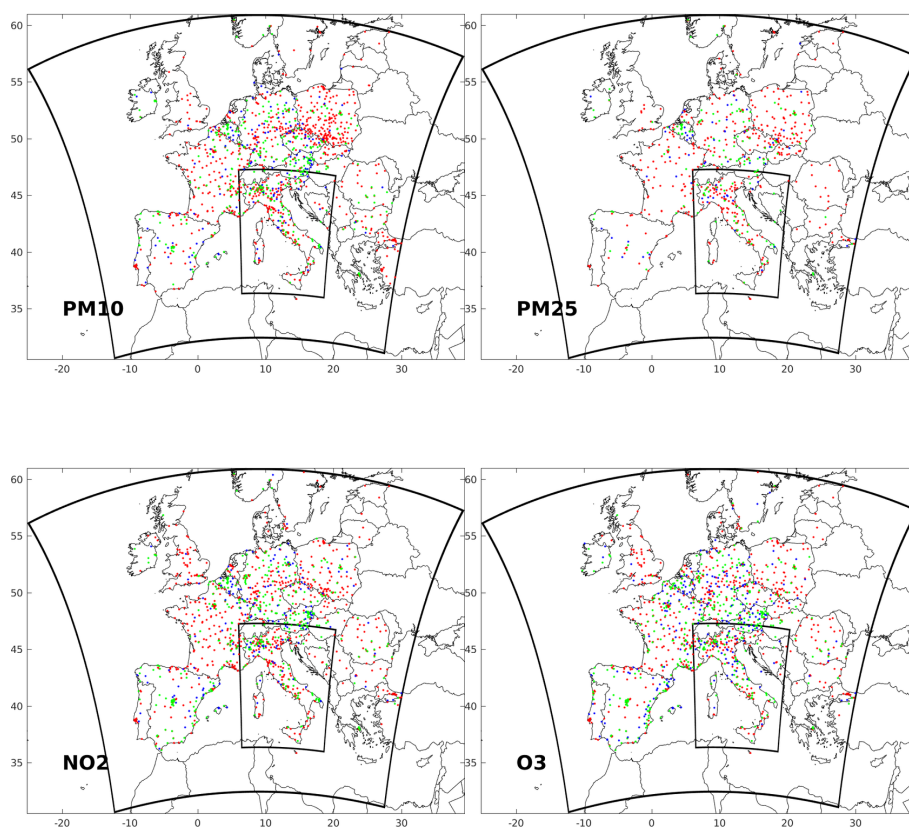
Figure 1. Model domains and 2017 valid observations for each evaluated pollutant. Red, green and blue dots identify urban, suburban and rural background stations respectively.

5   RAMS is a prognostic non-hydrostatic meteorological model, based on terrain following vertical coordinate system, multiple grid two-way nesting scheme and nudging techniques. In FORAIR_IT application it operates with a polar stereographic grid centred over Italy and 35 unevenly spaced levels in sigma coordinates. Convection has been parametrized following Kuo (1974) on the coarse grid and explicitly solved on the fine one.

SURFPro is a diagnostic meteorological post-processor for the calculation of turbulent parameters. Different schemes are
10   available to compute Planetary Boundary Layer (PBL) scaling parameters and horizontal and vertical diffusion coefficients. Concerning mixing height evaluation, the maximum value between outcomes of Maul (1980) version of Carson (1973) encroachment method and mechanical mixing height by Venkatram (1980) has been used for daytime estimate. Instead, for night-time mixing height, the minimum value has been chosen between Nieuwstadt (1981) extension of Zilitinkevich (1972) formula and Venkatram (1980) empirical relationship results. The vertical diffusivity has been parameterized following Lange
15   (1989), as implemented in CMAQ-Models3 (Byun et al., 1999), while the horizontal one is the sum of the values computed using Smagorinsky (1963) formulation and an empirical function dependent on local stability class and wind speed.

The chemical transport model FARM is a three-dimensional Eulerian model. It includes transport (advection and turbulent diffusion), gas phase chemistry by means of SAPRC-99 chemical scheme (Carter, 2000) and in-cloud chemistry. Aerosol processes are simulated by means of AERO3 (Binkowski, 1999; Binkowski and Roselle, 2003) for aerosol dynamics, ISOR-
20   ROPIA (Nenes et al., 1998; Fountoukis and Nenes, 2007) for aerosol inorganic chemistry and SORGAM (Schell et al., 2001) for secondary organic aerosol production. Wet and dry removal processes are also taken into account. The former is computed according to EMEP model approach (EMEP, 2003), distinguishing between in-cloud and below-cloud removal, the latter is evaluated by means of the well-known resistance model (Seinfeld and Pandis, 1998). It runs with automatic time step that is internally computed to satisfy the Courant-Friedrichs-Lewy (CFL) criteria.
25   Within FORAIR_IT, FARM has been implemented using 16 unevenly spaced Z-levels up to 10000 m, with the first one centred at 20 m above the ground.

The forecast system needs chemical and meteorological boundary conditions as external real-time input data: chemical data are downloaded from the global service (C-IFS) of Copernicus Atmospheric Service (CAMS, http://atmosphere.copernicus.eu) using ftp protocol at 15 UTC corresponding to the 00Z run with a spatial resolution of 80 km; meteorological data come from
30   GFS (Global Forecast System) meteorological forecast of NCEP (U.S. meteorological system) and are downloaded using the same protocol at 19:30 UTC corresponding to 18Z runwith a spatial resolution of 100 km. The fields used for meteorological boundary conditions are the same used for initializing meteorological model that has 6 hours of spin up, which is a good compromise for having adjusted meteorological fields and air quality forecast output on time. Initial conditions for FARM simulation come from the previous day forecast.
35   Emission data include anthropogenic, biogenic, sea salt and resuspended soil dust emissions. Anthropogenic emissions are referred to the year 2010 and come from TNO-MACIII (MACIII, 2016) and ISPRA (ISPRA, 2013) inventory dataset, covering Europe and Italy respectively. The most relevant point sources (stack emissions from large power plants, refineries, steelworks, industrial facilities) on the Italian domain are explicitly treated, with stack and flue gas-PM parameters. The emission inventories are mapped onto FARM species and disaggregated in time and space using the emission processor EMMA to produce

https://doi.org/10.5194/gmd-2020-54

5   three-dimensional gridded emission fields. About natural emissions, the MEGAN (Model of Emissions of Gases and Aerosols from Nature) model (Guenther et al., 2006) is used to produce biogenic emissions; sea salt emissions are function of wind speed and relative humidity, following Zhang et al. (2005); dust emissions take into account erosion and resuspension processes. They are parameterized, according to Vautard et al. (2005), through a formulation depending on air density, wind speed, convective velocity scale, soil type and soil moisture by means of a climatic index. The measurement data used for model validation

10 have been provided by the European Air Quality Portal (https://aqportal.discomap.eea.europa.eu/requirements/data-flows/), of European Environmental Agency (EEA, https://www.eea.europa.eu) which covers most of the European Member States. In particular primary validated assessment data – measurements (E1a data set) with the best quality flag have been used for the validation. The PM2.5 composition observations that are used in section 4 were kindly provided by the Environmental Regional Protection Agency of Emilia Romagna (ARPAE). According to European Air Quality Directive 2008 (EC, 2008), a minimum data availability of 75% is required to consider a monitoring station suitable for validation. Over the Italian territory, only daily mean PM10 and PM2.5 datasets were available. Only background stations were used as adequately spatially representative, given FORAIR_IT spatial resolution.

5   In Fig. 1 the observational network in 2017 year is shown for each pollutant and station area. In particular red, green and blue dots identify urban, suburban and rural stations respectively.

Table 1. Number of valid background stations for the year 2017 used for model evaluation per area.

	PM10	PM2.5	NO$_2$	O$_3$
Rural	363	154	384	414
Suburban	362	153	383	413
Urban	934	467	835	688

Table 1 shows the number of valid background stations used for the forecast evaluation per pollutant and station area.

In order to compare FORAIR_IT forecast with state of the art air quality models, the multi-model ensemble interim re-analysis of regional CAMS products (hereafter CAMS_50, http://macc-raq-op.meteo.fr) have been used. The skills of the two modelling systems in representing pollutant concentrations were compared with respect to the observations. Anyway, in the comparison with CAMS_50 we have to keep in mind that statistically multi-model ensemble should have lower error than the single members and that some of the EEA observations are not independent with respect to CAMS_50, since they have been assimilated by each member of the ensemble (https://policy.atmosphere.copernicus.eu/Reports.php).

FORAIR_IT operational chain starts at 20:00 UTC preparing inputs for meteorological model RAMS. Once they are ready, RAMS starts running with 384 cores and, meanwhile, the preprocessing of chemical boundary conditions and anthropogenic emissions are produced. Once RAMS has produced 3 days physical atmospheric forecast, the post-processing and the turbulent parameters are computed so that all the inputs needed by the chemical transport model FARM are ready. All the single core processes are contemporary submitted for both domains. FARM starts running with 384 cores for the first day using the initial condition of the first day forecast of the previous forecasting cycle. Once the first forecast day is completed, the necessary plots for the first day are produced with scalar job. Meanwhile, the second day of the chemical forecast is submitted. The procedure is cycling for the 3 forecasting days. The operational chain ends between 02UTC and and 04UTC depending on the computational load of the High Performance Computing facility CRESCO (Computational RESsearch Centre on COmplex systems) (Ponti et al., 2014). The whole operational chain is automatically launched via software utility cron. At present all forecast plots and movies are published on the web-site http://www.afs.enea.it/project/ha_forecast/ and a subset of them on https://impatti.sostenibilita.enea.it/research/activity/FORAIR-IT.

3   Evaluation of FORAIR_IT European and national forecast

Main statistical parameters are presented in the following figures for the pollutants of interest: PM10,PM2.5,NO$_2$ and O$_3$. The group of panels *a* of Figs. from 2 to 5 and group of panels *b* of Figs. 4 and 5 present the same skills for each pollutant, within a common layout. Labels *a* and *b* refer to skillscores computed over European and Italian domain, respectively.

The skill scores considered are RMSE, Modified Mean Bias and Correlation Coefficient (CORR) in panels *a2(b2)*, *a3(b3)* and *a4(b4)* as a function of forecast day (D0, D1, D2),respectively, whose definition is reported in Appendix A1 of the Supplementary Material.

The panels *a1* and *b1* show monthly Root Mean Square Error (RMSE) values computed for all valid monitoring stations. Values are refferred to daily maximum (for O_3 and NO_2) or daily mean (PM10 and PM2.5). Black and red lines correspond to
10 the first forecast day (D0) for FORAIR_IT and CAMS_50 Interim reanalysis (IRA), respectively.

The panels *a2*, *a3 a4* and *b2*, *b3 b4* show statistical scores of hourly pollutant concentrations as a function of forecast time for NO_2 and O_3 in Figs. 4 and 5. For PM10 and PM2.5 the panels from *b2* to *b4* of Figs. 2 and 3 show the skill scores on daily basis since Italian aerosol measurements are available as daily average concentrations.

The skill score presented are Root Mean Square Error (RMSE), Modified Mean Bias (MMB, see Appendix A1 for details)
15 and Correlation Coefficient (CORR) in panels *a2(b2)*, *a3(b3)* and *a4(b4)* as a function of forecast day (D0, D1, D2),respectively.

In all the figures, black lines and blue shaded regions refer to FORAIR_IT results: they represent the median value and the 25th and 75th percentiles of the distribution of the values computed, at a given forecast time on the year-long time series, for all the valid background monitoring stations. Blue lines represents the median values on the year long time series for CAMS_50 skills and they are compared with the FORAIR_IT first forecast day (D0), since it is the most analogous representation of the
20 same time series given the intrinsic difference between the two realizations. The grey shaded boxes represent the 25th and 75th percentiles; the line inside the boxes represents the median and the ends of the whiskers the 5th and the 95th percentiles. The average values are also displayed with cross markers.

Tables from 2 to 5 show RMSE, BIAS and CORR as function of forecast day and station area for all the pollutants taken into account (PM10, PM, NO_2, O_3). Averaged observed concentration as function of station area are shown too. Averaged skills
5 for both European and Italian domains are presented here. Appendix A2 (Supplementary Material) shows the same scores separately for European and Italian domain.

3.1 PM10
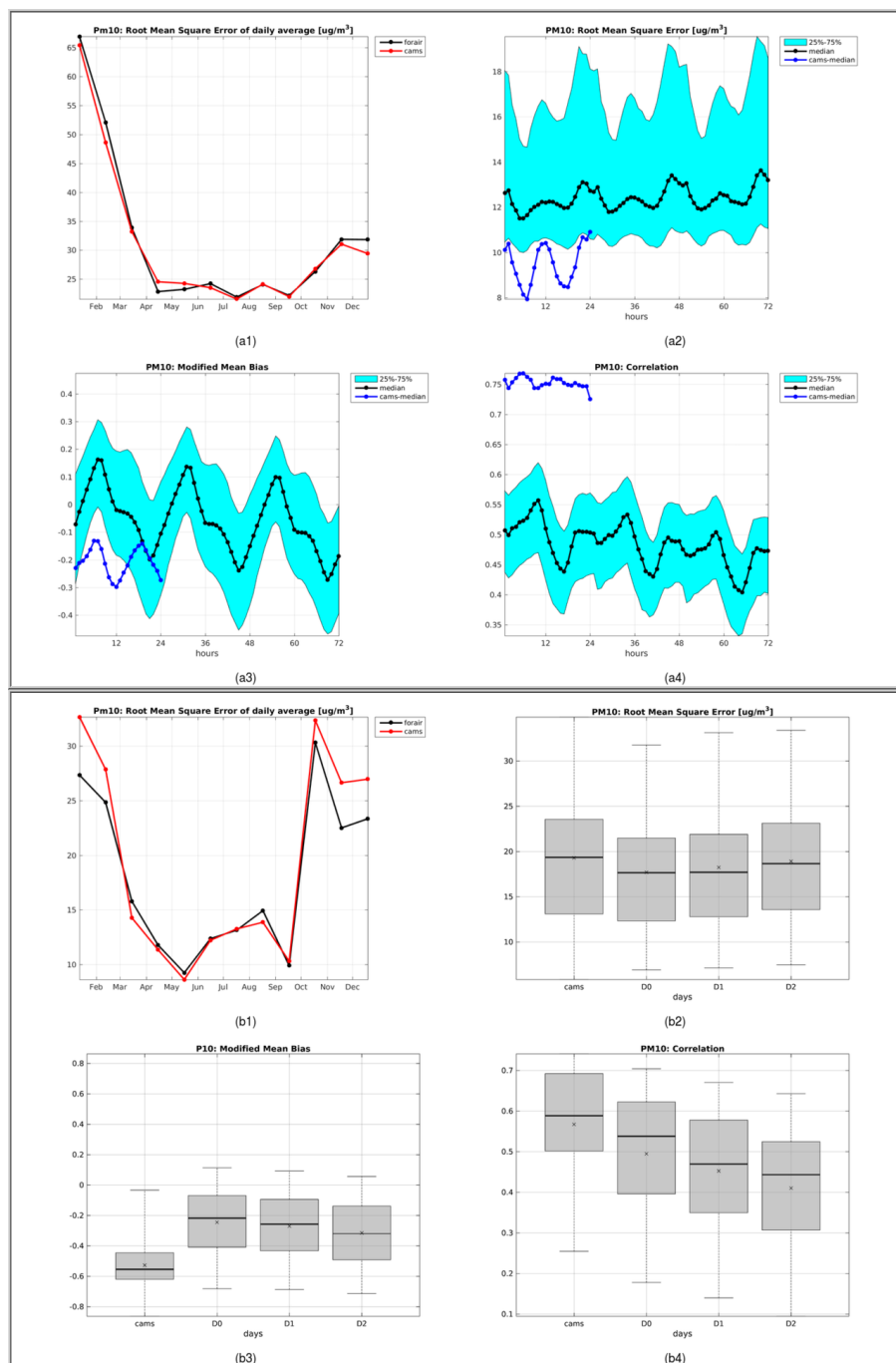
Figure 2. Forecast skill scores for PM10. Panels *a* and *b* refers to European and Italian domains, respectively.

Fig. 2 *a* and *b*, shows the performance skill score for PM10 over Europe and Italy respectively.

Monthly RMSE results, computed on daily averages for CAMS_50 and FORAIR_IT D0, are shown in Fig. 2 *a1*, on
European domain. The performances, for both realizations, show a marked seasonal dependency, with higher values from
November to March and lower ones from May to October. During January and February, RMSE reaches the highest values
for both FORAIR_IT and CAMS_50. During 2017 winter season, CAMS_50 has experienced the worst skills in the last three
years,2015-2018 (Collin et al., 2018) This may be explained by an anomalous aerosol concentration event occurred in Eastern
Europe, which was mainly driven by a significant long-range transport contribution (Tarasson et al., 2018; Fagerli et al., 2019).
Hourly RMSE values (Fig. 2 *a2*), as a function of forecast time, present a daily cycle with two minima (one in the morning
and one in the afternoon) and two maxima (around midnight and midday, respectively). This behaviour can be observed for
CAMS_50 results too but with a better defined bell-shape daily cycle. Moreover, in CAMS_50 cycle, midday maximum is as
high as midnight one, whereas FORAIR_IT daily cycle has midnight maximum substantially higher than midday one, espe-
cially for the median and the 75th percentile. RMSE performances remain almost unchanged as function of forecast time, even
if a slightly positive trend can be observed and may be due to the decrease of accuracy in describing atmospheric forcing and
boundary conditions. Overall RMSE CAMS_50 results are better than FORAIR_IT ones; anyway ca. 3 µg/m^3 mean distance
between median values may be considered an acceptable outcome considering the inherent nature diversity of the two realiza-
tions.

MMB values (Fig. 2 *a3*) show that FORAIR_IT generally overestimates PM10 concentration values during the morning and
underestimates them during the rest of the day. Moreover, MMB median values vary as a function of forecast time: overesti-
mation decreases and underestimation increases. Unlike FORAIR_IT, CAMS_50 underestimates PM10 concentrations all day
long. Consistently with RMSE values, the highest biases are observed at midnight and midday, while the lowest underestima-
tions are observed at 7 and 19 hours. Furthermore, whereas FORAIR_IT shows a 24-hour cycle of MMB skills, a 12-hour less
accentuated cycle characterizes CAMS_50 results. This difference between the daily cycles of MMB may be due to differences
in daily disaggregation of the emissions used for CAMS_50 ensemble models and FORAIR_IT, which is based on EMEP tem-
poral profiles.

CORR (Fig. 2 *a4*) ranges from 0.6 in the late morning of the first day of forecast to ca. 0.35 in the afternoon of the third day.
As MMB, CORR decreases as a function of the forecast time and it is the statistical metric most affected by the forecast length.
CAMS_50 shows a better CORR than FORAIR_IT and this may be due to the assimilation of some of the observations. In
summary, over the European domain, FORAIR_IT performance is comparable with CAMS_50 one on monthly basis RMSE,
it shows worse skills for hourly RMSE and CORR as fuction of forecast time. FORAIR_IT has a well-defined daily cycle for
CORR but not for RMSE. MMB skills are generally better than CAMS_50: bias errors oscillate around 0 during the 24 hours
while CAMS_50 shows a systematic negative bias. Concerning model performances on Italian domain, RMSE values, com-
puted on daily averages for FORAIR_IT first forecast day D0 and CAMS_50 (Fig. 2 *b1*), show a marked monthly dependency,
with the highest values in January (27 µg/m^3) and October (30 µg/m^3), and a relative maximum in August (15 µg/m^3). The
overall monthly RMSE behaviour is consistent with CAMS_50 and the FORAIR_IT skills do not differ too much as function
of forecast time (Fig. 2 *b2*). As already highlighted on European domain, MMB skills (Fig. 2 *b3*) are better for FORAIR_IT

than for CAMS_50 whereas Correlation performances (Fig. 2 *b4*) are worse. In any case, on the Italian domain the difference between the two model performances is less accentuated. The MMB outcomes show that FORAIR_IT underestimates the

10   PM10 concentrations over Italy and the underestimation increases during the forecasting period, reaching the highest value in the third day. Again the Correlation performances decrease as function of forecast time. The difference of the mean of monthly RMSEs for FORAIR_IT first day of forecast (D0) and CAMS_50, hereafter RMSED, introduced for a concise comparison indicator of the two realizations, is computed as described in Appendix A1. Since the optimal values of RMSE are reached when they are close to 0, lower are the values of the RMSED better is the performance of one realization respect to another.

15   PM10 RMSED values are 0.6 and -1.2 µg/m^3 for European and Italian domain, respectively. The negative value highlights the better perfomances of FORAIR_IT with respect to CAMS_50 on the Italian domain with respect to the European one, meaning how a higher resolution of the computational grid in the region of interest increases the forecast performances. Moreover the higher consistency between FORAIR_IT and CAMS_50 results is probably related to the fact that in CAMS_50 simulations the observations over Italy were not assimilated since they were not available on hourly basis.

Table 2 shows RMSE, BIAS and CORR per forecast day (D0, D1, D2) and station area (rural, suburban, urban). As expected all the skills deteriorate as function of forecast time. However, the accuracy reduction is quite limited suggesting that the prediction could be extended to more days in advance. Highest errors (RMSE and BIAS) are found for background urban

5   stations despite the higher correlation. Overall PM10 performances are consistent with the results obtained by other air quality forecasting systems (e.g. Honoré et al., 2008; Žabkar et al., 2015).

Table 2. PM10 EU+IT. RMSE, BIAS, CORR and annual mean values of observations as function of forecast days and station areas.

	RMSE			BIAS			CORR			OBS
Forecast Day	D0	D1	D2	D0	D1	D2	D0	D1	D2	
Rural	14.9	15.0	15.2	-4.7	-5.0	-5.5	0.39	0.38	0.37	16.6
Suburban	18.9	19.1	19.4	-8.0	-8.3	-8.8	0.33	0.32	0.30	20.8
Urban	26.6	26.8	27.0	-12.1	-12.5	-13.0	0.27	0.26	0.25	25.4

3.2 PM2.5
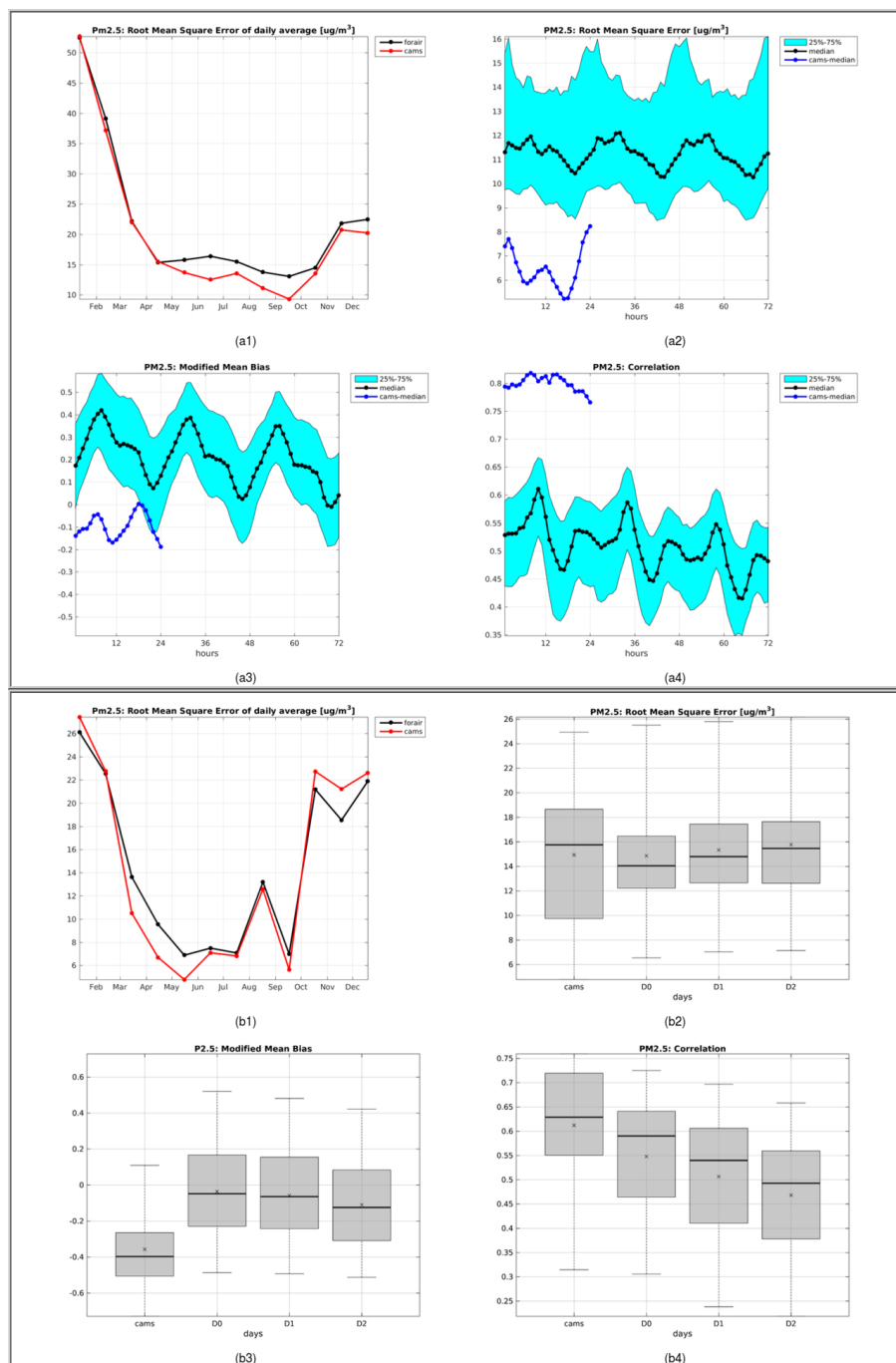
Figure 3. Forecast skill for PM2.5. Panels *a* and *b* refers to European and Italian domains, respectively.

Group of panels *a* and *b* of Fig. 3 shows skill scores for PM2.5 computed over Europe and Italy respectively.

10   Monthly RMSE values, computed on European domain using daily averages of CAMS_50 results and of FORAIR_IT first days of forecast (panel *a1* in Fig. 3), show a marked seasonal dependency, with higher values from November to March and lower ones from April to October. The highest RMSE value is reached in January (about 52 μg/m^3) consistently with PM10 trends, as discussed above. In general, FORAIR_IT skills are very similar to CAMS_50 ones in the first months of the year, whereas they are slightly worse from May to the end of the year. Overall, the maximum difference between FORAIR_IT and

15   CAMS_50 RMSE values is about 4 μg/m^3 and is observed in September.

Concerning RMSE as function of forecast time (panel *a2* in Fig. 3), FORAIR_IT PM2.5 skills behaviour is quite similar to PM10 one. FORAIR_IT and CAMS_50 simulations present average RMSE values of about 11 μg/m^3 and 7 μg/m^3 respectively, with FORAIR_IT daily oscillations (less than 2 μg/m^3) lower than CAMS_50 ones (about 2.5 μg/m^3). Again the reason of the different daily cycles of the two models performances may be found in the different temporal profiles used for emission

20   disaggregation.

This hypothesis is confirmed by the differences in the daily cycles of MMB values (panel *a3* in Fig. 3). Daily cycle of FORAIR_IT performances presents the highest MMB value in the morning, with a single peak at hour 8, and the lowest one at hour 23; on the contrary CAMS_50 performances cycle exhibits a bimodal pattern with a reduced oscillation. In general, CAMS_50 underestimates PM2.5 concentrations (of about -0.1) while FORAIR_IT overestimates it (of about 0.25) with a

25   slight decrease of the overall bias along with the forecast time.

The CORR as function of the forecast time (panel *a4* in Fig. 3) exhibits an accentuated daily cycle for FORAIR_IT simulation that is not present for CAMS_50 outcomes. CORR daily cycle shows the highest values at 9 and the lowest ones at 17 of each forecast day. A slight decrease of FORAIR_IT CORR values along with the forecast time is also observed: the highest median value is about 0.6 during the first forecast day and the lowest value is about 0.47 during the last one. This behaviour is quite

30   similar to PM10 one, as well as the fact that CORR is the metric where the differences between FORAIR_IT and CAMS_50 performances are the highest. Mean FORAIR_IT CORR for the first forecast day is about 0.52 whereas CAMS_50 one is higher than 0.8. CORR is probably the metric most affected by the assimilation process; indeed, as already observed for PM10 outcomes, differences between FORAIR_IT and CAMS_50 correlation performances are not so evident, when computed over the Italian domain (panel *b4* in Fig. 3).

35   In general, most of the FORAIR_IT skills show better results when computed on the Italian domain (group of panels *b* in Fig. 3), regarding both the comparison with observations and the inter comparison with CAMS_50 performances.

Monthly RMSE values computed using first day forecast daily averages (panel *b1* in Fig. 3) show the same seasonal cycle, already observed on the European domain, but the values are lower: the highest error is found in January (26 μg/m^3) and the lowest one in May and September (7 μg/m^3). Overall FORAIR_IT and CAMS_50 outcomes present the same pattern.

5   FORAIR_IT skills are slightly better than CAMS_50 ones during winter months, whereas they are slightly worse in spring.

FORAIR_IT RMSE as function of forecast day (panel *b2* in Fig. 3) exhibits a slight increase along with the forecast time, however also for the last day of forecast the median value of FORAIR_IT RMSE is still lower than CAMS_50 one.

Also for the MMB skills (panel *b3* in Fig. 3) FORAIR_IT performs better than CAMS_50: MMB median value is very close

to 0 even if a slight degradation appears along the forecast days.

10   CAMS_50 performs better than FORAIR_IT only with respect to correlation (panel *b4* in Fig. 3). Anyway differences are not so evident: the median values of CAMS_50 and FORAIR_IT correlation coefficients for the first day forecast are respectively 0.62 and 0.59. Again a degradation of FORAIR_IT performances is observed along with the forecast time, more accentuated compared to the other metric behaviour. Overall, the comparison between FORAIR_IT and CAMS_50 performances over Italy shows similar results with respect to the European domain. RMSED values for PM25 are 1.7 and 0.3 μg/m^3 over European and

15   Italian domain, respectively. The better skills obtained over the Italian domain with respect to European one are probably due to the increase of spatial resolution. In addition, the Italian measurements, differently from the European ones, were not used for assimilation within CAMS_50 simulation; indeed on Italian domain, validation and assimilation dataset are independent for both FORAIR_IT and CAMS_50 realizations.

Table 3 shows that model performances don't change so much as function of forecast day. Correlation skills slightly deteriorate along with the forecast time. On the contrary, BIAS tends to be reduced from D0 to D2. This behaviour is common for all the pollutants overestimated by the model (see NO$_2$ hereafter). The positive and negative BIASes observed respectively for PM2.5 and PM10 may indicate that coarse material emissions are underestimated and they are only partially compensated by

5   overestimation of PM2.5 emissions and chemical production of the secondary fraction. The highest RMSE values are found in urban stations, the worst BIASes in rural stations whereas the lowest correlation coefficients are observed at suburban sites.

Table 3. PM2.5 EU+IT. RMSE, BIAS, Correlation and annual mean values of observations as function of forecast days and station areas.

	RMSE			BIAS			CORR			OBS
Forecast Day	D0	D1	D2	D0	D1	D2	D0	D1	D2	
Rural	10.6	10.6	10.7	1.6	1.2	0.7	0.54	0.53	0.51	11.6
Suburban	12.5	12.6	12.8	0.2	-0.2	-0.8	0.50	0.48	0.46	14.6
Urban	14.7	14.9	15.2	-0.7	-1.2	-1.8	0.52	0.51	0.48	16.5

3.3 Nitrogen dioxide
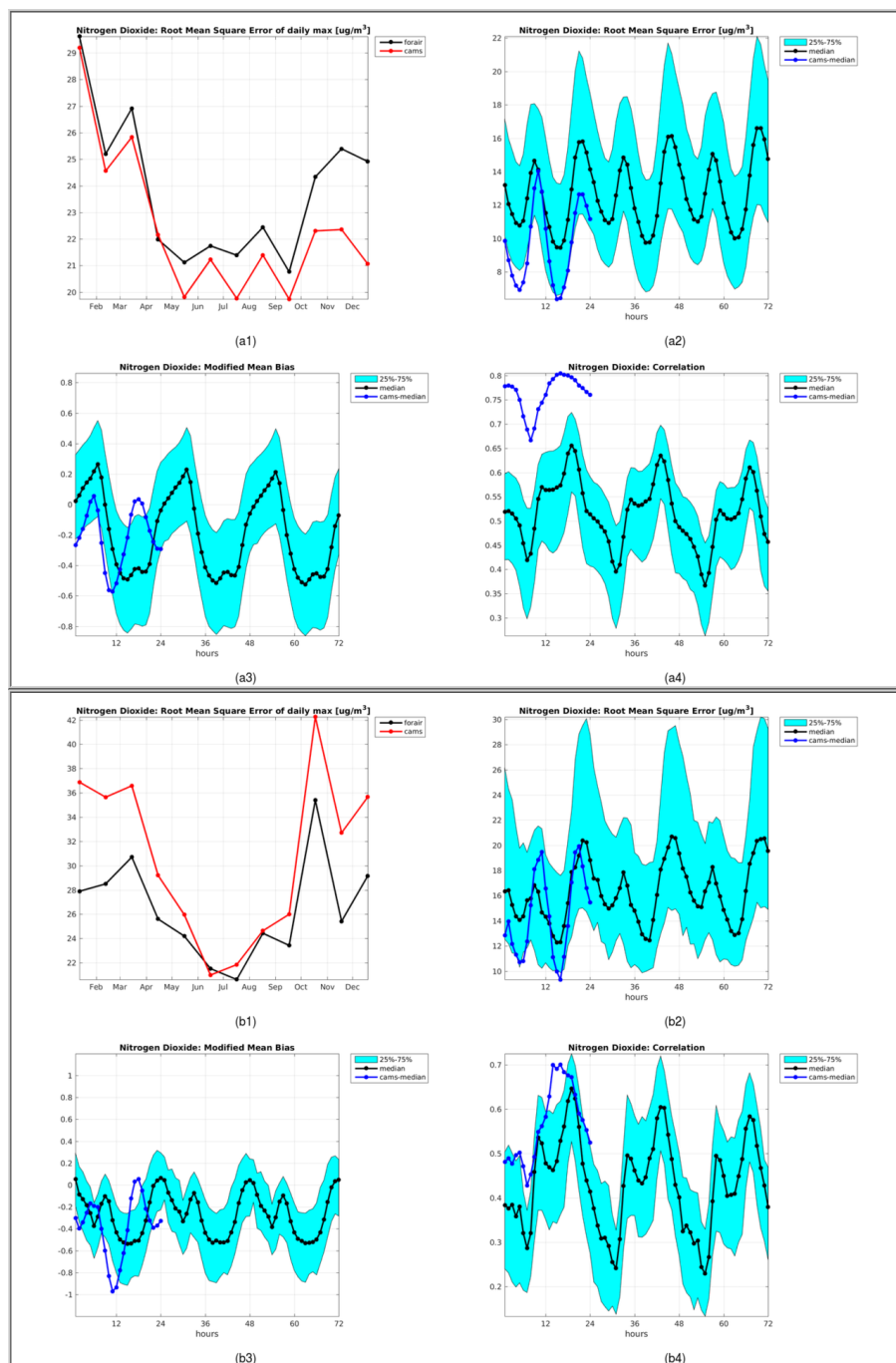
Figure 4. Forecast skill for NO$_2$. Panels *a* and *b* refer to European and Italian domains, respectively.

Panels *a* and *b* of Fig. 4 show skill scores for Nitrogen Dioxide (NO_2), computed over Europe and Italy respectively.

10   On the European domain, the RMSE as function of months for FORAIR_IT D0 realization (panel *a1* in Fig. 4) has been computed on NO_2 daily maximum, without taking into account that observation and forecast time series may have reported maximum values at different time instants. Both FORAIR_IT and CAMS_50 simulations present a seasonal cycle for NO_2 skills: overall, RMSE of NO_2 daily maximum is higher in winter and lower in summer. Even if the two simulations show similar seasonal cycles, RMSE values are systematically higher for FORAIR_IT simulation than for CAMS_50 ones.

15   Hourly RMSE values (panel *a2* in Fig. 4) show very similar daily cycles for FORAIR_IT and CAMS_50 simulations, with two peaks, one in the morning and one in the evening. The highest value is in the morning for CAMS_50, and in the evening for FORAIR_IT; moreover, in this case, a slight increase of FORAIT_IT RMSE values from D0 to D2 is also observed.

MMB median values (panel *a3* in Fig. 4) range from -0.5 to 0.3 for FORAIR_IT, with a clear daily cycle: overestimation during the first hours of the day and underestimation after midday are observed. The shape of the daily cycle is different

20   between FORAIR_IT and CAMS_50 since the latter shows a bimodal pattern: at hours 5 and 18, MMB values are close to 0, whereas the highest underestimation (-0.6) is reported around midday. Again the reason of the different daily cycles of the two models performances may be related to the different temporal profiles used for emission disaggregation.

Concerning correlation skills (panel *a4* in Fig. 4), the highest values occur in the evening and the lowest ones in the morning; moreover, performances generally deteriorate along with the forecast time.

25   The combined decrease in correlation and increase of bias in the first hours of the morning may be related to excessive tritation of the O_3. This hypothesis is in agreement with the O_3 skillscore presented in the next subsection. Even if CAMS_50 correlation values are always higher than FORAIR_IT ones, the shape of the curves is quite similar apart from a slight advance of CAMS_50 maximum.

On Italian domain, the seasonal cycle of monthly RMSE values of NO_2 daily maximum (panel *b1* in Fig. 4) is similar

30   to the pattern observed on European domain. However, RMSE values are generally higher on Italian domain, especially for CAMS_50 simulation. Overall, CAMS_50 values are always higher than FORAIR_IT ones, except for June. The shapes of FORAIR_IT and CAMS_50 curves are very similar with a maximum in October. Again errors are higher in winter and lower in summer when NO_2 concentrations are usually lower too.

Also hourly RMSE skills (panel *b2* in Fig. 4) are quite in agreement with performances on European domain. In particular,

35   the same daily cycle is observed, with both the morning and the evening peaks. However values are higher, especially as far as the evening peak is concerned. FORAIR_IT daily cycle is consistent with CAMS_50 one, even if the latter presents a wider range of variation. In particular CAMS_50 has two similar maxima in the morning and in the afternoon (about 20 $\mu g/m^3$) and FORAIR_IT has a lower maxima in the morning (16 $\mu g/m^3$) and similar to CAMS_50 maxima in the afternoon. As far as the minima are concern, CAMS_50 shows better skill than FORAIR_IT with morning values reaching 11 $\mu g/m^3$ and 14 $\mu g/m^3$ and

5   afternoon values about 10 $\mu g/m^3$ and 12 $\mu g/m^3$ for CAMS_50 and FORAIR_IT respectively.

MMB skills (panel *b3* in Fig. 4) in FORAIR_IT and CAMS_50 realizations show different daily cycles, as it occurs for PM10, PM2.5 and NO_2 over Europe. Furthermore, here the two curves are quite anti-correlated, a part from a small shift. Overall MMB skills show that underestimation occurrences are more frequent than overestimation ones, especially for

CAMS_50. Performances trends over the three days forecast period do not show any degradations, indicating that forecast
10   might be extended further without losing much accuracy.

According to correlation skills (panel *b4* in Fig. 4), the same daily cycles of the performances occur as observed on European domain, with the highest values in the evening and the lowest ones in the morning, together with a slow decrease along with the forecast time. FORAIR_IT correlation coefficient values are quite in agreement with CAMS_50 ones but they are usually lower. Again as described for previous pollutants, the differences between FORAIR_IT and CAMS_50 in correlation skills are

15   reduced over Italian domain with respect to European one. NO_2 RMSEDs are 1.4 and -4.3μg/m^3 over European and Italian domain, respectively. Particularly positive performances are achieved for NO_2 since a negative value is obtained over Italy despite the lack of data assimilation in FORAIR_IT.

Table 4 shows RMSE, BIAS and CORR for NO_2 as function of station area and forecast day. Overall NO_2 BIAS is positive for rural stations and negative for suburban and urban ones. For NO_2 it is evident that forecasted concentration values tend to decrease passing from D0 to D2: indeed positive BIASes decrease and negative ones increase. This behaviour, already observed for PM10 and PM2.5, may be due to either meteorological conditions, being more dispersive along with the forecast time, or to

5   global boundary conditions, underestimating more in D2 than in D0. Generally, the best performances in terms of all the metrics are found for rural stations and again the overall scores suggest that the forecast may be acceptable for more than three days.

Table 4. NO$_2$ EU+IT. RMSE, BIAS, Correlation and annual mean values of observations as function of forecast days and station areas.

	RMSE			BIAS			CORR			OBS
Forecast Day	D0	D1	D2	D0	D1	D2	D0	D1	D2	
Rural	10.7	10.7	10.8	1.4	1.1	0.9	0.56	0.54	0.52	9.5
Suburban	16.1	16.3	16.5	-3.7	-4.0	-4.2	0.50	0.48	0.47	18.2
Urban	18.7	19.0	19.2	-6.5	-6.8	-7.0	0.50	0.48	0.47	21.8

3.4 Ozone
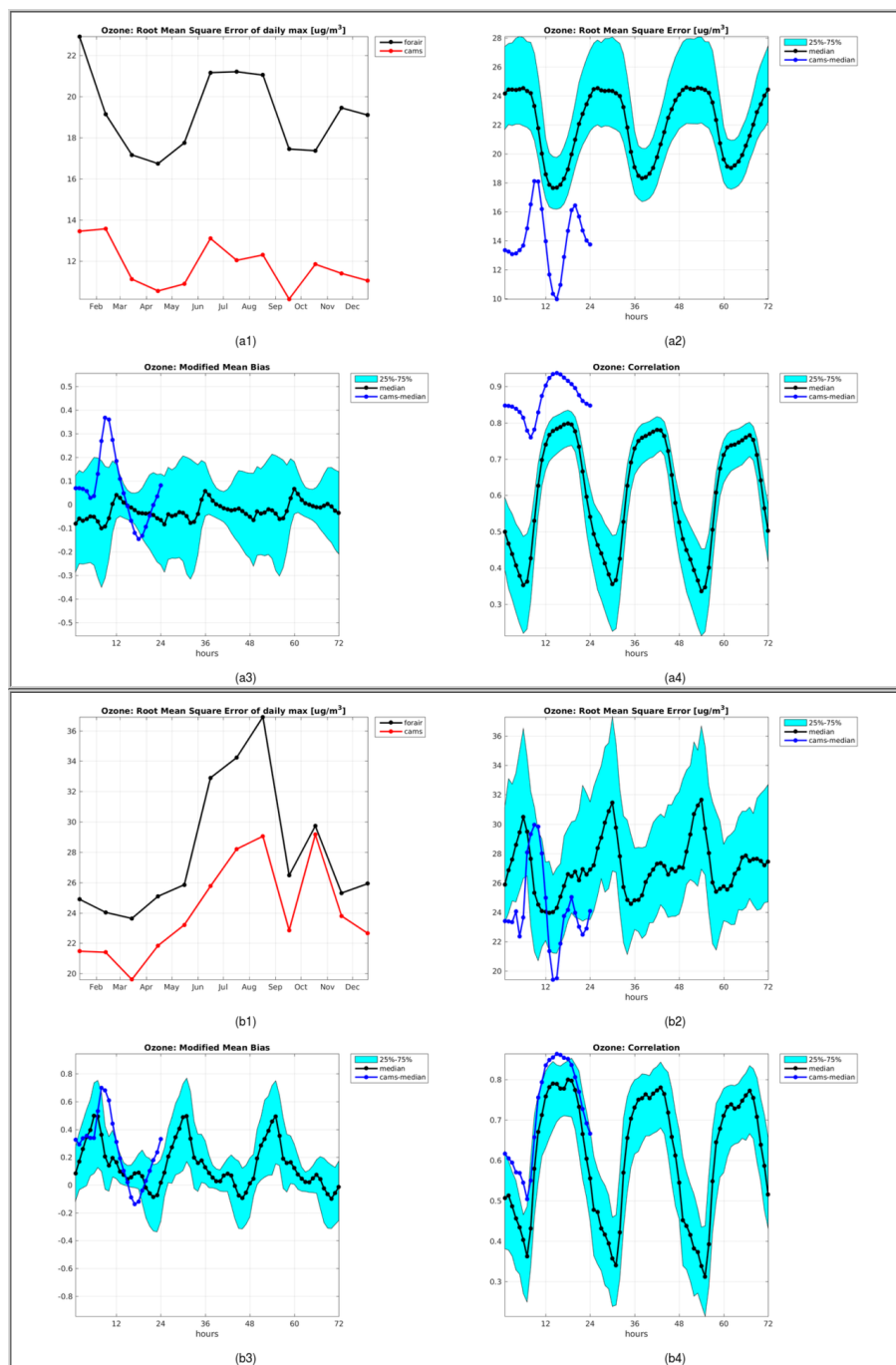
Figure 5. Forecast skill for O_3. Panels *a* and *b* refer to European and Italian domains, respectively.

Panels *a* and *b* of Fig. 5 show Ozone (O_3) skill scores computed over Europe and Italy respectively.

10   Concerning the European domain, RMSE outcomes for O_3 D0 daily maximum values (panel *a1* in Fig. 5) show a clear seasonal cycle with minimum values in spring and autumn and maximum ones in summer and winter. FORAIR_IT and CAMS_50 present a similar behaviour, even if FORAIR_IT RMSE values are much higher and with a more accentuated seasonal cycle with respect to CAMS_50. The RMSE peak values occur not surprising in summer, when O_3 concentrations are usually higher. Conversely the winter one is quite unusual; anyway it is present both in FORAIR_IT and CAMS_50 outcomes and it may be

15   related to high concentrations of chemical reactive compounds in Eastern Europe that caused anomalous high concentrations of NO_2 and particulate matter, in the first months of the year.

Hourly RMSE values as function of forecast time (panel *a2* in Fig. 5) are consistent with the outcomes on D0 daily maximum. In particular, FORAIR_IT error is higher than CAMS_50 one. The two simulations report the lowest RMSE values at the same time instant. On the contrary, nighttime behaviour is quite different: FORAIR_IT reports RMSE highest values while

20   CAMS_50 presents a relative minimum. From D0 to D2 no evident performances degradation is found especially concerning maxima values, whereas the minima ones are slightly increasing.

Contrary to RMSE, MMB outcomes (panel *a3* in Fig. 5) are quite good. FORAIR_IT shows a less accentuated daily cycle compared to CAMS_50 one. Moreover MMB values are very close to 0 and quite stable during the three forecast days. Overall FORAIR_IT overestimates O_3 in some sites and underestimates it in other ones and the time series of the MMB median val-

25   ues present the highest overestimation values around midday. The high RMSE and the negative MMB during nighttime may enforce the hypothesis of excessive tritation of O_3.

Since RMSE is composed of two terms (the BIAS component and the Centered Root Mean Square Error, CRMSE), simultaneous bad RMSE and good MMB skills mean high CRMSE values. This latter metric summarizes in turn the model capacity of capturing the observed standard deviation and the correlation between the simulated and the observed concentrations. For

30   both European and Italian domains, the amplitude of the standard deviations is quite well captured since the median of the skill variance (ratio between the observed and the simulated standard deviations) is 0.98 for Europe and 0.82 for Italy. Therefore, the highest error is related to the correlation: in particular, this term explains about the 93% (on European domain) and 98% (on Italian domain) of the error associated to CRMSE. However, for some stations problems arise also in capturing the large observed daily excursion since the 5% of the stations has skill variance less than 0.7.

35   Correlation coefficient values are shown in panel *a4* of Fig. 5 where a slight decrease of performances can be noticed along with the forecast time.

Over the Italian domain the seasonal cycle of RMSE values of O_3 daily maximum (panel *b1* in Fig. 5) presents the highest values only in summer as expected. This outcome confirms the hypothesis of anomalous concentration values over Eastern Europe as the reason of the performances deterioration in this area during the winter. Again, FORAIR_IT values are higher

5   than CAMS_50 ones but the differences are less accentuated over Italy than over Europe.

Daily cycles of hourly RMSE values (panel *b2* in Fig. 5) present the best FORAIR_IT performances around midday and the worst ones during the early morning. A slight shift is observed for CAMS_50 cycle, especially as far as early morning maximum value is concerned. MMB outcomes (panel *b3* in Fig. 5) show a general overestimation and a similar trend of

FORAIR_IT and CAMS_50 performances; the similar skill is confirmed by the correlation values patterns (panel *b4* in Fig.

10 5). Moreover, both these skills show a slight deterioration along with the forecast time.

In conclusion, the comparison between FORAIR_IT and CAMS_50 performances over Italian domain shows similar results with respect to European one, even if differences are less accentuated, probably due to a better representations of the processes thanks to the increase of spatial resolution on Italian FORAIR_IT domain.This is also evident from RMSED values that are 7.4 and 3.8 μg/m^3 over European and Italian domain, respectively. Table 5 shows the model skills for O_3 daily maximum as function

5 of forecast day and station area. Again, the skills do not degrade much along the forecast days and best results are found for rural stations in terms of RMSE, and suburban one in terms of BIAS. Contrary to what is observed for the other pollutants, here negative BIASes decrease along with the forecast time. This outcome, meaning that O_3 forecasted concentration values tend to increase passing from D0 to D2, is expected because, being O_3 a secondary pollutant, its precursors decrease.

Table 5. O_3 EU+IT. RMSE, BIAS, Correlation and annual mean values of observations as function of forecast days and station areas.

	RMSE			BIAS			CORR			OBS
Forecast Day	D0	D1	D2	D0	D1	D2	D0	D1	D2	
Rural	23.9	24.2	24.6	-5.8	-5.6	-5.5	0.65	0.64	0.62	62.6
Suburban	24.8	25.1	25.3	2.3	2.5	2.6	0.67	0.66	0.65	53.5
Urban	24.2	24.5	24.9	3.1	3.4	3.5	0.67	0.65	0.64	50.5

4 PM episode on Po valley

Air pollution in Italy is one of the main environmental issues due to the high number of days with daily mean PM10 concentration exceeding the limit of 50 μg/m^3 (EEA, 2019). High levels of airborne PM are also related to health issues. It is therefore pivotal to have the possibility to predict with accuracy in space and time the expected PM concentration in order to activate possible reduction strategies before the actual occurrence of exceedances. The Po Valley is one of the main hot spot for air pollution in Italy and Europe. This area is conditioned by unfavourable meteorological conditions (air stagnation, thermal high) that determine a low boundary layer height during the cold season (autumn - winter) (Bigi and Ghemaldi, 2016). During October 2017, a 10 day outbreak of high airborne PM pollution was recorded in Emilia Romagna (ER), a region in the hearth of the Po Valley. In order to understand the accuracy of the model in predicting such events, the PM monitoring stations in ER have been isolated by the databased presented in Fig. 1 and analyzed separately. Black markers on the upper panel of Fig. 6 show the sampling points where both PM10 and PM2.5 daily average were measured by Regional Agency for Environmental Protection (ARPAE) monitoring network. Square markers (both black and white) are locations where the PM2.5 composition is also available. Observed and first day forecast average and spatial standard deviation are computed at black marker sampling points both for PM10 and PM2.5. In the lower panel of the same figure the observed (dash line) and simulated (continuous line) concentrations for PM10 (magenta) and PM2.5 (cyan) are shown. Shaded area represents the PM10 standard deviations computed among the stations. It is evident that FORAIR_IT was able to forecast the trend of the air pollution event. The observed PM10 and PM2.5 air concentrations are well reconstructed by the model, in particular from October 8th to October 17th both in term of absolute values and correlation between model simulation and observations. During this period, PM10 concentration underestimation is partly compensated by an overestimation of PM2.5, while from 17th October to the 28th the modelling system starts to underestimate also PM2.5, and consequently, the overall agreement between observed and predicted PM concentrations worsens.
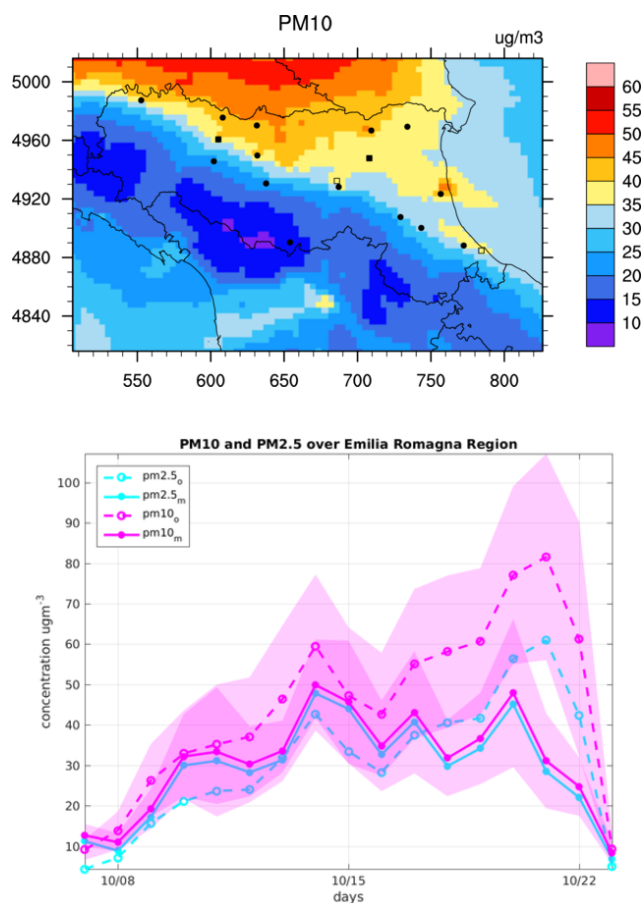
Figure 6. Upper panel: Colours are averaged PM10 concetrations over the period 9th-22nd October. Black markers are stations that measure both PM10 and PM2.5. Square markers are stations that measures PM2.5 compositions. Lower panel: PM10 (magenta) and PM2.5 (cyan) observed (dash line) and simulated (continuous line) daily concentrations computed at black markers. Shaded areas represent the observed and simulated standard deviations of PM10 concentration among the stations.

10   It is interesting also to highlight that during the considered period the variability of the observations increases, as evidenced by the increase in their standard deviation, while the model variation does not reflect this high variability. This could reflect local features in some meteorological variables that the model is not perfectly able to forecast (although the model well reproduces the absence of precipitation recorded over the studied period) or the effects of other variables, such as emission sources, that do not allow the accurate reconstruction of local processes. In order to understand the differences between observed and simulated

15   variability the PM10 average concentration over the period 9th-22nd October was analyzed. The model average forecast (upper panel 6) shows a typical regional distribution of PM with higher concentrations expected close to the Po Valley (in the north of ER region) and lower concentration in the south where the Appennini mountains chain begins. The model predictions highlight also an area with a strong gradient of PM10 concentration (around 15 μg/m^3 in few kilometers) stretching from NW to SE in ER region. This high gradient area may be particularly challenging for the model and specific analysis were then conducted to

20   analyze its variability over the period under consideration. Daily PM10 concentration averages were computed for the period 7th - 24th of October in order to evaluate the position of this gradient area over ER. Fig. A1 (in supplementary material) shows that, during the considered days, the advection of PM10 from the NW of Po Valley is highly variable with also significant inter day variability. In particular the highest PM10 concentrations are found over Lombardy region and they are transported toward ER region. The concentration plume changes several time its directions, from NW to SE, then to E and then again to

25   SE. Consequently the highest gradient area shifted several times its position over ER with usually the north east part of ER highly affected by this pollutant intrusion. The peculiar topographical conformation of the Po Valley, surrounded by complex orography, could have lead the meteorological model to forecast too intense zonal wind (or incorrect wind directions) during the last part of the period that shifted the plume signal too much north, therefore avoiding an actual intrusion of significant PM concentration and determining the underestimation here described. This condition highlights the importance of capturing

30   the spatial and temporal variability of the meteorological conditions in forecasting model for air pollutants. Although this may be the reason of less variability in the model than in the observations, it does not explain model underestimation of PM10 and overestimation of PM2.5 concentrations. In order to further analyze model underestimation of PM10, Fig. 7 shows the reconstruction over the year of coarse PM10 (namely PM10-PM2.5), in comparison with the relative observed values. The daily averages of observed coarse PM10 (red line) were compared with the model daily averages (black line) computed on the same sampling points (black markers Fig. 6) Relative contribution of boundary intrusion along with soil and sea salt resuspension is also shown with black, blue and cyan dots, respectively. Dots are shown only when the values are one standard deviation greater than the annual average, that is when a significant event occurred. Similarly to the period of the forecast specifically analyzed,

5   PM10 is always underestimated (yearly mean concentration of observed values equal to 8 μg/m^3 versus 3 μg/m^3 from the modelled ones). Significantly, both sea salt and soil resuspension did not contribute significantly to the highest concentration of PM modelled. Indeed the peaks of concentration from the model simulation were mostly related to the intrusion of air masses from the boundary of the domain (Fig. 7, black dots).
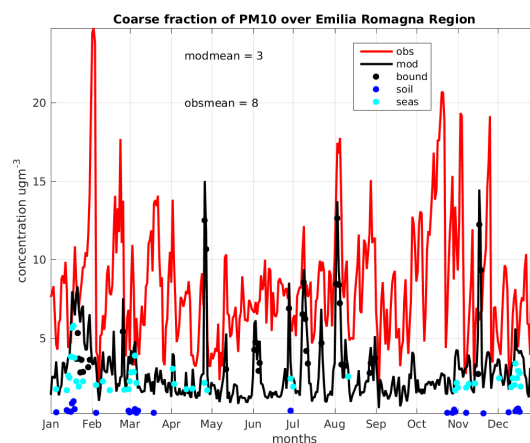
Figure 7. Annual distribution of modelled and observed ER regional daily PM10 coarse fraction mean values. Simulated (black line) and observed (red line) coarse fraction of PM10 are reported. Black dots correspond to the coarse fraction of PM10 coming from the boundary conditions, cyan dots to the sea salt component and blue dots to the soil and mineral dust component from the model output. Dots are reported only if the daily value is one standard deviation greater than the annual mean of the relative specific component.

10   Several causes may contribute to model underestimation of the coarse fraction of PM10: underestimation in the anthropogenic emission inventory, lack of processes related to generation of organic coarse fraction (i.e. degradation of plant, pollens etc), excessive dry deposition and lack of processes related to the formation of the coarse fraction of nitrate that may be relevant in certain condition (Uno et al., 2017). Trying to quantify the underestimation of the coarse fraction of PM10 in the anthropogenic emission inventory, among the different causes, if we consider a closed box of the same surface of the ER region (where this emission is about 3300 ton/y), with the same thickness of the first model layer (40 m) within which no removal,

5   advection and transformation processes happen, the ER coarse fraction of PM10 will contribute with a concentration of 0.4 μg/m^3. Thus we may conclude that the underestimation in the anthropogenic emission inventory plays an important role in the overall underestimation of the coarse fraction of PM10 concentration Despite the error associated to PM10, some discrepancy between observations and model estimate is found also in PM2.5, therefore the chemical composition of PM2.5 has been analysed where available (square markers of Fig. 6)
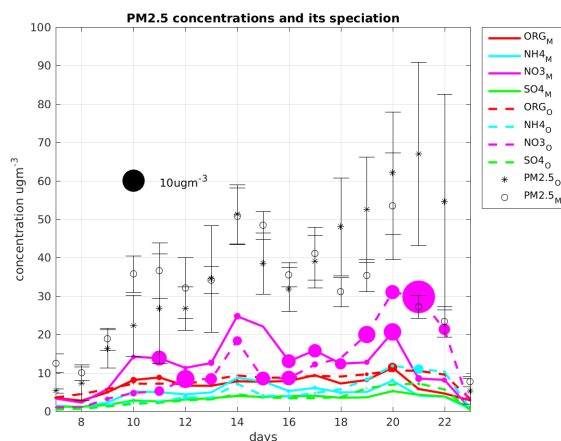
Figure 8. PM2.5 daily mass concentration (observed - black stars - and modelled - white dots) for the period 7th - 23rd of October. Average PM2.5 daily mass is reported with the one standard deviation. Chemical PM2.5 speciation from data of the air monitoring network was obtained (dashed lines, magenta for nitrate, green for sulphate, cyan for ammonium ions and red for organic compounds) and compared with the same output variable from the model (full lines). The standard deviation among the different measuring stations is reported, for each chemical specie, by the dimension of the full coloured markers.

10   The observed and modelled PM2.5 composition is presented in Fig. 8. The main discrepancy between modelled and observed chemical species is the concentration of nitrogen species, especially nitrates . Indeed NO_3^- and partly NH_4^+ are overestimated until the 17th, determining a consequent overestimation of fine PM mass, while during the last period of simulation nitrate and ammonium ions are underestimated, resulting in a significant underestimation of fine PM mass. The highest NO_3^-/NH_4^+ underestimation from the model is on October 21st where the observed values showed the maximal variability

15   among the sampling stations. This mismatch between observed and predicted values determines the highest differences between observed and modelled NO_3^-/NH_4^+ and PM2.5 concentrations (21st and 22nd of October). Eventually starting from the 22nd and during the 23rd moderate precipitations associated with high wind speed occurred over ER, mainly in the NE area (ARPAE https://www.arpae.it/cms3/documenti/_cerca_doc/meteo/radar/rapporti/Rapporto_meteo_20171022-23.pdf) determining the reported drop of observed and modelled concentrations on the 23rd of October. The mismatch between observed

20   and modelled nitrogen species needs further investigations. From these analyses the importance of the capability of the model to capture the actual spatial and temporal variability in meteorological conditions results clear. However, despite the incomplete agreement in the absolute values between the model and the regional monitoring network, FORAIR_IT was able to predict the first peak of air pollution associated with stable meteorological conditions.

5   Summary and conclusions

25 Despite many efforts in reducing anthropogenic emission some exceedances of the European Directive standard limit value are still recorded. In Italy, one of the main environmental concern, especially in the Po valley, are exceedances of the daily mean PM10 limits. Therefore, the forecasting of such events may be crucial for protecting both envirnoment and human health by applying special mitigation procedures. In the present work, the FORAIR_IT air quality forecasting system has been presented. It runs on daily basis and computes the main pollutant concentrations up to 3 days, with 20kmx20km resolution over Europe

30 and 4kmx4km resolution over the Italian domain. Overall, model skill score are presented for NO_2, O_3, PM10 and PM2.5 along with a detailed analysis of a three weeks lasting high level pollution event in Emilia Romagna. In the Appendix 4 of the Supplementary Material, the capability of the model in reproducing the exceedances is also presented and compared with CAMS_50. Major findings are: i) the model skill scores computed over the Italian domain are generally better than European one, highlightening the importance of both model resolution and local emission inventory; ii) model behaviour is consistent with CAMS_50 reanalysis ensemble product for almost all the species. Significantly, monthly RMSE skill scores are very similar both as absolute values and annual cycle comparing CAMS_50 and FORAIR_IT outputs for airbone particulates and NO_2. However, for hourly model statistics, CAMS_50 seems to perform better especially for the correlation with respect to

5 observations, which is probably related to data assimilation procedure applied on single members of the ensemble. Ozone skill scores suggest that some developments in FORAIR_IT are still necessary in order to have comparable skill with CAMS, in term of absolute values both of daily maximum and daily cycle. Concerning the importance of future developments, as the degradation of model skill from D0 to D2 is not significant, the possibility to extend the forecast window is very feasible. Eventually, the analysis of the specific ER event shows that the capacity of capturing both the large scale processes and the

10 small scale ones is crucial in order to capture the variability and the timing of the exceedances. In conclusion, FORAIR_IT may be considered an up to date forecasting system whose skills are comparable with the ones presented in Copernicus (CAMS) program, and as far as we know it is the first forecasting system at high spatial resolution proposed at Italian National level. Furthermore, FORAIR_IT may be a useful tool to be used by the policy makers in order to apply extraordinary procedure to prevent/mitigate exceedances. However, as forecasting systems require continuous developments and updates to release

15 increasingly accurate forecasts, the following development are under investigation:

- the substitution of the meteorological model RAMS with Weather Research and Forecasting Model (WRF);

- the addition of fire emission from Global Fire Assimilation System (GFAS, https://atmosphere.copernicus.eu/global-fire-emissions);

- the operationally evaluation of the first day of forecast with one day delay with respect to the production;

20 - the assimilation of insitu chemical measurements in order to better initialize the forecast;

Appendix A: Supplementary Material

A1   Metrics definition

The forecast performances are measured using the same statistical indicators that are routinely applied to CAMS_50 ensemble members in the quarterly evaluation report. They are Root Mean Square Error (RMSE), BIAS, Correlation and Modified Mean

25 Bias (MMB). All of them have advantages and disadvantages but may contribute in understanding the lack of the modelling system in reproducing observed values. In the following equations, f stands for the forecast values and O for the observed ones. RMSE is not a dimensionless variable and requires knowledge of typical mean values to be interpreted, being also strongly dominated by the largest values, due to the squaring operation.

$$RMSE = \sqrt{\frac{\sum_{t=1}^{N}(f_t - O_t)^2}{N}} \tag{A1}$$

BIAS gives information of systematic error of the model but, as well as the RMSE, depends on typical mean values of the

5 pollutants taken into account.

$$BIAS = \frac{\sum_{t=1}^{N}(f_t - O_t)}{N} \tag{A2}$$

Correlation is a statistical dimensionless indicator that measure the extent to which two or more variables fluctuate together.

$$CORR = \frac{\sum_{t=1}^{N}(f_t - \bar{f})(O_t - \bar{O})}{\sqrt{\sum_{t=1}^{N}(f_t - \bar{f})^2}\sqrt{\sum_{t=1}^{N}(O_t - \bar{O})^2}} \tag{A3}$$

Modified Mean Bias (MMB) is a statistical dimensionless metrics which has the advantage to measure the forecast bias bounded by the values -2 and +2.

$$MMB = \frac{2}{N} \sum_{t=1}^{N} \left(\frac{f_t - O_t}{f_t + O_t} \right) \tag{A4}$$

Eventually, the Root Mean Square Difference (RMSED) has been introduced to compare the results between FORAIR_IT
5 and CAMS_50 simulations.

$$RMSED = \left(\sqrt{\frac{\sum_{t=1}^{N} (f_t - O_t)^2}{N}} \right)_{FORAIR\_IT} - \left(\sqrt{\frac{\sum_{t=1}^{N} (f_t - O_t)^2}{N}} \right)_{CAMS\_50} \tag{A5}$$

A2 Supplementary Tables

Table A1. PM10 EU. RMSE, BIAS, Correlation and annual mean values of observations as function of forecast days and station areas.

	RMSE			BIAS			CORR			OBS
Forecast Day	D0	D1	D2	D0	D1	D2	D0	D1	D2	
Rural	10.8	10.8	11.0	0.4	-0.1	-0.7	0.51	0.51	0.49	14.9
Suburban	13.5	13.6	13.9	-2.5	-2.9	-3.5	0.49	0.48	0.46	18.6
Urban	16.8	17.0	17.4	-4.5	-5.0	-5.6	0.53	0.52	0.50	21.2

Table A2. PM10 IT. RMSE, BIAS, Correlation and annual mean values of observations as function of forecast days and station areas.

	RMSE			BIAS			CORR			OBS
Forecast Day	D0	D1	D2	D0	D1	D2	D0	D1	D2	
Rural	18.4	18.8	19.4	-6.7	-7.0	-7.6	0.52	0.49	0.46	20.5
Suburban	20.1	20.6	21.4	-8.1	-8.4	-9.4	0.54	0.51	0.48	27.4
Urban	19.3	19.9	20.6	-5.8	-6.1	-7.0	0.49	0.46	0.43	26.9

Table A3. PM2.5 EU. RMSE, BIAS, Correlation and annual mean values of observations as function of forecast days and station areas.

	RMSE			BIAS			CORR			OBS
Forecast Day	D0	D1	D2	D0	D1	D2	D0	D1	D2	
Rural	10.5	10.4	10.4	2.8	2.4	1.8	0.53	0.52	0.51	10.9
Suburban	11.8	11.8	11.9	2.1	1.5	0.9	0.50	0.49	0.47	13.1
Urban	12.6	12.7	12.9	0.6	0.1	-0.5	0.55	0.54	0.52	14.1

Table A4. PM2.5 IT. RMSE, BIAS, Correlation and annual mean values of observations as function of forecast days and station areas.

	RMSE			BIAS			CORR			OBS
Forecast Day	D0	D1	D2	D0	D1	D2	D0	D1	D2	
Rural	15.2	15.6	16.2	-5.3	-5.2	-5.7	0.57	0.53	0.49	17.3
Suburban	18.0	18.5	19.1	-4.2	-4.6	-5.5	0.54	0.52	0.49	21.9
Urban	15.3	15.8	16.0	0.6	0.4	-0.4	0.56	0.54	0.50	18.0

Table A5. NO_2 EU. RMSE, BIAS, Correlation and annual mean values of observations as function of forecast days and station areas.

	RMSE			BIAS			CORR			OBS
Forecast Day	D0	D1	D2	D0	D1	D2	D0	D1	D2	
Rural	10.7	10.7	10.8	2.6	2.3	2.1	0.54	0.52	0.50	8.7
Suburban	15.9	16.0	16.3	-3.1	-3.5	-3.7	0.48	0.47	0.45	18.0
Urban	16.7	17.0	17.2	-5.3	-5.8	-0.3	0.53	0.52	0.50	20.6

Table A6. NO$_2$ IT. RMSE, BIAS, Correlation and annual mean values of observations as function of forecast days and station areas.

	RMSE			BIAS			CORR			OBS
Forecast Day	D0	D1	D2	D0	D1	D2	D0	D1	D2	
Rural	12.4	12.6	12.6	-1.4	-1.4	-1.6	0.54	0.53	0.51	11.5
Suburban	21.3	21.6	21.8	-5.1	-5.2	-5.5	0.52	0.50	0.49	24.7
Urban	21.5	22.0	22.2	-4.2	-4.2	-4.5	0.51	0.49	0.48	26.5

Table A7. O_3 EU. RMSE, BIAS, Correlation and annual mean values of observations as function of forecast days and station areas.

	RMSE			BIAS			CORR			OBS
Forecast Day	D0	D1	D2	D0	D1	D2	D0	D1	D2	
Rural	23.3	23.5	23.9	-6.6	-6.3	-6.0	0.66	0.65	0.64	61.0
Suburban	23.5	23.7	24.0	1.1	1.5	1.7	0.69	0.68	0.67	52.3
Urban	22.8	23.0	23.4	2.4	2.9	3.3	0.69	0.68	0.67	48.7

Table A8. O_3 IT. RMSE, BIAS, Correlation and annual mean values of observations as function of forecast days and station areas.

	RMSE			BIAS			CORR			OBS
Forecast Day	D0	D1	D2	D0	D1	D2	D0	D1	D2	
Rural	27.9	28.4	28.7	2.6	2.1	1.6	0.65	0.63	0.62	63.8
Suburban	29.0	29.7	29.9	8.7	8.5	8.0	0.70	0.69	0.68	51.8
Urban	28.5	29.2	29.7	4.0	3.6	3.0	0.69	0.67	0.66	55.4

A3 Model Evolution of PM10 episode
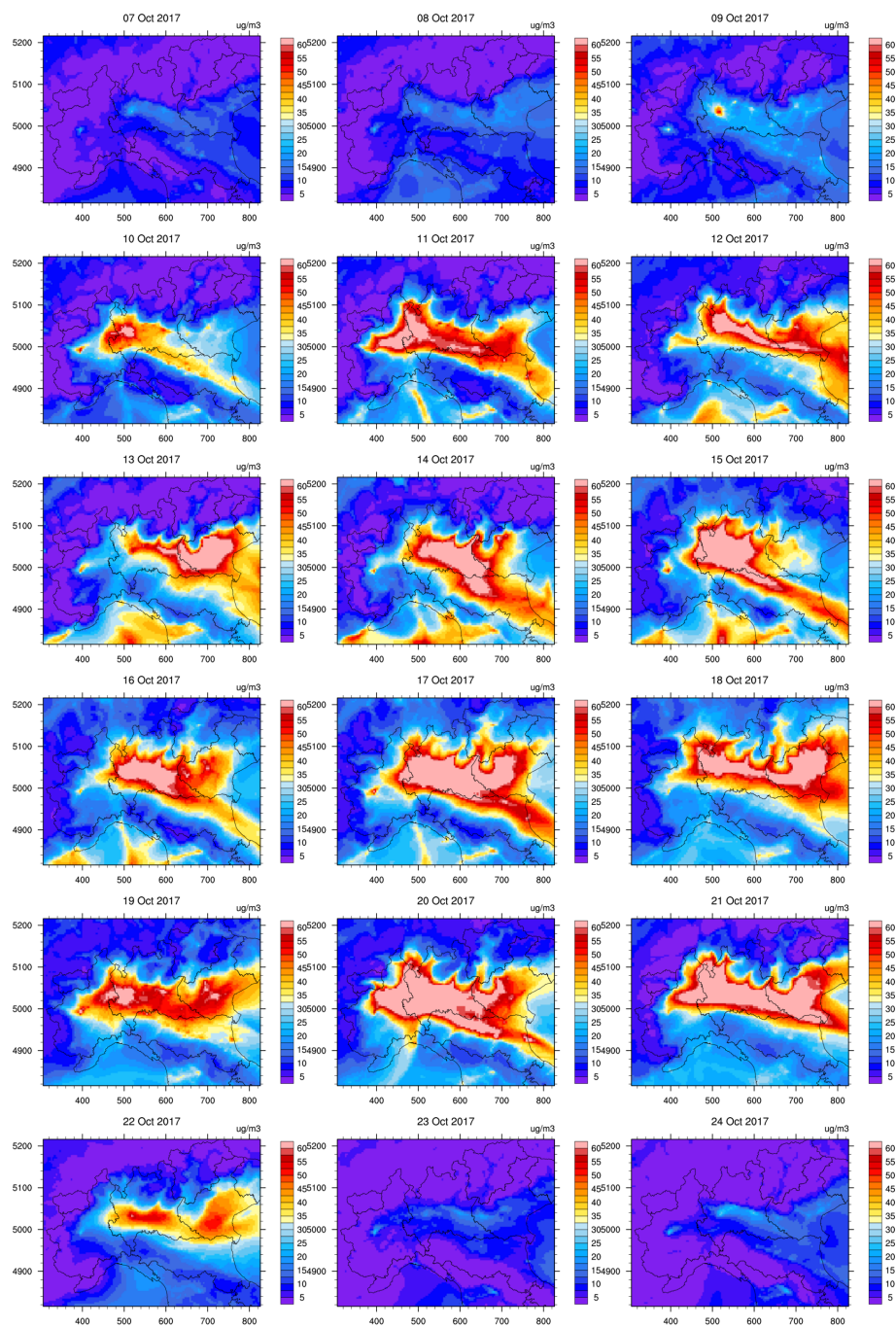
Figure A1. Daily mean concentration of PM10

A4   Evaluation in reproducing the exceedances
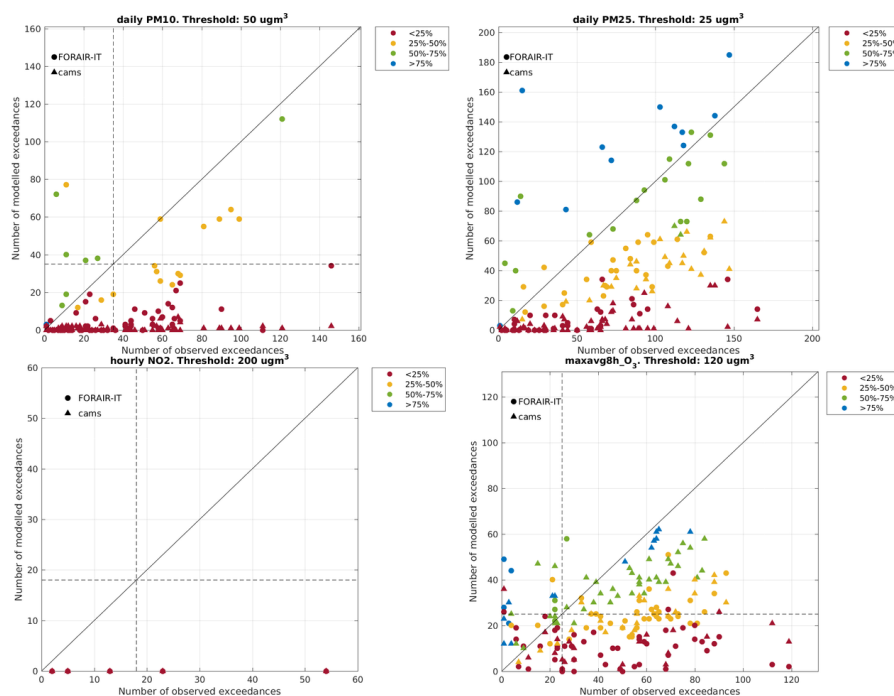
Figure A2. Capability of FORAIR_IT (dots) system in reproducing the exceedances compared with CAMS_50 (triangle) over the Italian domain. Colours refers to Probability of detection (POD).

10 *Author contributions.* MD prepared the meteorological model setup,MA and GB prepared the chemical model set-up. GG and MA performed model simulations. AC AP ID prepared the anthropogenic emission. MA and LV elaborated all the data analysis. MA and LV and prepared the manuscript with all the co-authors contributions with the assistance of LC MM ID

Competing interests. The Authors do declare that no competing interests are present

Acknowledgements. We acknowledge ARIANET s.p.a. especially in the persons of Alessio D'allura, Camillo Silibello, Sandro Finardi and Giuseppe Calori for sharing with us ideas and softwares. We acknowledge the Italian Environmental Ministry for the financial support in the development of FORAIR_IT forecasting system. We acknowlege the Environmental Protection Agency of Emilia Romagna, in the person
5 of Arianna Trentini, for providing the observations on chemical composition of PM2.5. The computing resources and the related technical support used for this work have been provided by CRESCO/ENEAGRID High Performance Computing infrastructure and its staff. CRESCO/ENEAGRID High Performance Computing infrastructure is funded by ENEA, the Italian National Agency for New Technologies, Energy and Sustainable Economic Development and by Italian and European research programmes, see http://www.cresco.enea.it/english for information.

10 References

Adani, M., Mircea, M., D'Isidoro, M., Costa, M.P, Silibello, C., 2015. Heavy Metal Modelling Study over Italy: Effects of Grid Resolution, Lateral Boundary Conditions and Foreign Emissions on Air Concentrations. Water, Air, & Soil Pollution, 226, 46, doi:10.1007/s11270-015-2324-7.

Arianet, 2011. SURFPRO3 User's guide (SURFace-atmosphere interface PROcessor, Version 3). Software manual. Arianet R2011.31.

15 Arianet, 2014. Emission Manager. Modular processing system for model-ready emission input Preparation. Software Manual.

Bessagnet, B., Colette, A., Meleux, F., Rouïl, L., Ung, A., Favez, O., Cuvelier, C., Thunis, P., Tsyro, S., Stern, R., Manders, A., Kranenburg, R., Aulinger, A., Bieser, J., Mircea, M., Briganti, G., Cappelletti, A., Calori, G., Finardi, S., Silibello, C., Ciarelli, G., Aksoyoglu, S., Prévôt, A., Pay, M.-T., Baldasano, J. M., García Vivanco, M., Garrido, J. L., Palomino, I., Martín, F., Pirovano, G., Roberts, P., Gonzalez, L., White, L., Menut, L., Dupont, J.-C., Carnevale, C., Pederzoli,A., 2014. The EURODELTA III exercise: Model evaluation
20   with observations issued from the 2009 EMEP intensive period and standard measurements in Feb/Mar 2009 Technical Report 1/2014. http://emep.int/publ/reports/2014/MSCW_technical_1_2014.pdf

Bessagnet, B., Pirovano, G., Mircea, M., Cuvelier, C., Aulinger, A., Calori, G., Ciarelli, G., Manders, A., Stern, R., Tsyro, S., García Vivanco, M., Thunis, P., Pay, M.-T., Colette, A., Couvidat, F., Meleux, F., Rouïl, L., Ung, A., Aksoyoglu, S., Baldasano, J. M., Bieser, J., Briganti, G., Cappelletti, A., D'Isidoro, M., Finardi, S., Kranenburg, R., Silibello, C., Carnevale, C., Aas, W., Dupont, J.-C., Fagerli, H.,
25   Gonzalez, L., Menut, L., Prévôt, A. S. H., Roberts, P., and White, L., 2016. Presentation of the EURODELTA III intercomparison exercise – evaluation of the chemistry transport models' performance on criteria pollutants and joint analysis with meteorology, Atmos. Chem. Phys., 16, 12667–12701, https://doi.org/10.5194/acp-16-12667-2016.

Bigi, A. and Ghermandi, G., 2016. Trends and variability of atmospheric PM2.5 and PM10-2.5 concentration in the Po Valley, Italy. Atmos. Chem. Phys., 16, 15777–15788.

30 Binkowski, F. S., 1999. The aerosol portion of Models-3 CMAQ. In Science Algorithms of the EPA Models-3 Community Multiscale Air Quality (CMAQ) Modeling System. Part II: Chapters 9-18. D.W. Byun, and J.K.S. Ching (Eds.). EPA-600/R-99/030.

Binkowski, F.S. and Roselle, S.J., 2003. Models-3 community multiscale air quality (CMAQ) model aerosol component 1. Model description. Journal of Geophysical Research, 108, D6, 4183, doi:10.1029/2001JD001409.

Byun D. W., Pleim J. E., Tang R. T., Bourgeois A., 1999. Meteorology-chemistry interface processor (MCIP) for MODELS-3 community
35   multiscale air quality (CMAQ) modeling system - Chapter 12. EPA/600/R-99/030.

Carson, D.J., 1973. The development of a dry inversion-capped convecticely unstable boundary layer. Quarterly Journal of the Royal Meteorological Society, 99, 450-467, doi:10.1002/qj.49709942105.

Carter, W. P. L., 2000. Documentation of the SAPRC-99 Chemical Mechanism for VOC Reactivity Assessment. Final Report to California Air Resources Board.

Ciancarella, L., Adani, M., Briganti, G., Cappelletti, A., Cremona G., Ciucci A., D'Elia, I., D'Isidoro, M., Piersanti, A., Righini, G., Russo, F., Vitali, L., Zanini, G., 2016. La simulazione nazionale di AMS-MINNI relativa all'anno 2010. ENEA Technical Report, RT/2016/12/ENEA
5   (in Italian).

Collin, G., Pithon, P., Plu, M., Petiot, V., Assar, N., Arteaga-Pichard, V. et al. 2018. Regional Production, Quarterly report on the daily analyses and forecasts activities, and verification of the ENSEMBLE performances at the surface and above surface. Tech. Rep. CAMS50_2015SC3_D50.4.1.2.ENSEMBLE_201808_Daily_Analyses_Report_v1, CAMS50_2015SC3_D50.4.2.2.ENSEMBLE_201808_Daily_Forecasts_Report_v1,

10   CAMS50_2015SC3_D50.5.1.2.ENSEMBLE_201808_NRT_Verification_Report_v1,
      CAMS84_2015SC3_D84.5.1.10_D84.6.1.5_2018MAM_v1. URL https://atmosphere.copernicus.eu/sites/default/files/2018-09/
      CAMS50_2015SC3_D50.4.1.2_D50.4.2.2_D50.5.1.2.ENSEMBLE_201807_Production_Report_v1_VF_.pdf

      Cotton, W. R., Pielke Sr., R. A., Walko, R. L., Liston, G. E., Tremback, C. J., Jiang, H., McAnelly, R. L., Harrington, J. Y., Nicholls, M. E.,
      Carrio, G. G., McFadden, J. P., 2003. RAMS 2001: Current status and future directions. Meteorology and Atmospheric Physics, 82, 5-29,
15   doi:10.1007/s00703-001-0584-9.

      D'Allura, A., Costa, M.P., Silibello, C., 2018. QualeAria: European and national scale air quality forecast system performance evaluation.
      International Journal of Environment and Pollution (IJEP), 64, 1/2/3, doi:10.1504/IJEP.2018.099152.

      European Commission (EC), 2004. Directive 2004/107/EC of the European Parliament and of the Council of 15 December 2004 relating to
      arsenic, cadmium, mercury, nickel and polycyclic aromatic hydrocarbons in ambient air.
20   Official Journal European Union En. Series, L23/3

      European Commission (EC), 2008. Directive 2008/50/EC of the European Parliament and of the Council of 21 May 2008 on ambient air
      quality and cleaner air for Europe (The Framework Directive). Official Journal European Union En. Series, L152/51.

      European Environmental Agency (EEA), 2019. Air quality in Europe — 2019 report. EEA Report | No 20/2019, doi:10.2800/822355

      EMEP,2003. Simpson, D., Fagerli, H., Jonson, J.E., Tsyro, S., Wind, P., Tuovinen, JP., 2003. Transboundary Acidification, Eutrophication
25   and Ground Level Ozone in Europe. PART I. Unified EMEP Model Description. EMEP Report 1/2003. https://www.emep.int/publ/reports/
      2003/emep_report_1_part1_2003.pdf

      Fagerli, H., Tsyro, S., Jonson, JE., Nyiri, A., Gauss, M., Simpson, D., Wind, P., Benedictow, A., et al., 2019. Transboundary particulate
      matter, photo-oxidants, acidifying and eutrophying components. EMEP Status Report 1/2019. URL https://emep.int/publ/reports/2019/
      EMEP_Status_Report_1_2019.pdf

30   Fountoukis, C. and Nenes, A., 2007. ISORROPIA II: A Computationally Efficient Aerosol Thermodynamic Equilibrium Model for K+,
      Ca2+, Mg2+, NH4+, Na+, SO42-, NO3-, Cl-, H2O Aerosols. Atmospheric Chemistry and Physics, 7(17), 4639-4659, doi:10.5194/acp-7-
      4639-2007

      Gariazzo, C., Silibello, C., Finardi, S., Radice, P., Piersanti, A., Calori, G., Cecinato, A., Perrino, C., Nussio, F., Cagnoli, M., Pelliccioni,
      A., Gobbi, G. P., Di Filippo, P., 2007. A gas/aerosol air pollutants study over the urban area of Rome using a comprehensive chemical
35   transport model. Atmospheric Environment, 41, 7286-7303, doi:10.1016/j.atmosenv.2007.05.018.

      Guenther, A., Karl, T., Harley, P., Wiedinmyer, C., Palmer, P. I., and Geron, C., 2006. Estimates of global terrestrial isoprene emissions using
      MEGAN (Model of Emissions of Gases and Aerosols from Nature). Atmospheric Chemistry and Physics, 6, 3181-3210, doi:10.5194/acp-
      6-3181-2006.

      Ha, S., Liu, Z., Sun, W., Lee, Y., Chang, L., 2019. Improving air quality forecasting with the assimilation of GOCI AOD retrievals during the
      KORUS-AQ period. Atmospheric Chemistry and Physics, https://doi.org/10.5194/acp-2019-648, in review, 2019.

      Honoré, C., Rouïl, 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., Poisson, N., 2008. Predictability of European air quality: Assessment of 3 years of operational
5   forecasts and analyses by the PREV'AIR system. Journal of Geophysical Research: Atmospheres, 113, D4, doi:10.1029/2007JD008761.

      ISPRA,   2013.   http://www.sinanet.isprambiente.it/it/sia-ispra/inventaria/disaggregazione-dellinventario-nazionale-2010/
      aggiornamento-con-dati-della-submission-2013/view downloaded in December 2013 and not more available.

      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., Eben, K., 2012. A review of operational, regional-scale, chemical weather forecasting models in Europe. Atmospheric Chemistry and Physics, 12, 1-87, doi:10.5194/acp-12-1-2012.

Kuo, H. L., 1974. Further Studies of the Parameterization of the Influence of Cumulus Convection on Large-Scale Flow. Journal of the Atmospheric Sciences, 31(5), 1232-1240, doi:10.1175/1520-0469(1974)031<1232:FSOTPO>2.0.CO;2.

Lange, R., 1989. Transferrability of a three-dimensional air quality model between two different sites in complex terrain. Journal of Applied Meteorology, 78 (7), 665-679, doi:10.1175/1520-0450(1989)028<0665:TOATDA>2.0.CO;2.

MACIII project, 2016. Final Report Monitoring Atmospheric Composition and Climate 3. https://atmosphere.copernicus.eu/sites/default/files/repository/MACCIII_FinalReport.pdf

Marécal, V., Peuch, V.-H., Andersson, C., Andersson, S., Arteta, J., Beekmann, M., Benedictow, A., Bergström, R., Bessagnet, B., Cansado, A., Chéroux, F., Colette, A., Coman, A., Curier, R. L., Denier van der Gon, H. A. C., Drouin, A., Elbern, H., Emili, E., Engelen, R. J., Eskes, H. J., Foret, G., Friese, E., Gauss, M., Giannaros, C., Guth, J., Joly, M., Jaumouillé, E., Josse, B., Kadygrov, N., Kaiser, J. W., Krajsek, K., Kuenen, J., Kumar, U., Liora, N., Lopez, E., Malherbe, L., Martinez, I., Melas, D., Meleux, F., Menut, L., Moinat, P., Morales, T., Parmentier, J., Piacentini, A., Plu, M., Poupkou, A., Queguiner, S., Robertson, L., Rouïl, L., Schaap, M., Segers, A., Sofiev, M., Tarasson, L., Thomas, M., Timmermans, R., Valdebenito, Á., van Velthoven, P., van Versendaal, R., Vira, J., and Ung, A.,2015. A regional air quality forecasting system over Europe: the MACC-II daily ensemble production. Geoscientific Model Development, 8(9), 2777–2813, doi:10.5194/gmd-8-2777-2015.

Maul, P.R., 1980. Atmospheric transport of sulphur compound pollutants. Central Electricity Generating Bureau MID/SSD/80/0026/R.

Mircea, M., Ciancarella, L., Briganti, G., Calori, G., Cappelletti, A., Cionni, I., Costa, M., Cremona, G., D'Isidoro, M., Finardi, S., Pace, G., Piersanti, A., Righini, G., Silibello, C., Vitali, L., and Zanini, G., 2014. Assessment of the AMS-MINNI system capabilities to predict air quality over Italy for the calendar year 2005. Atmos. Environ. 84: 178–188.

Mircea, M., Grigoras, G., D'Isidoro, M., Righini, G., Adani, M., Briganti, G., Cremona, G., Ciancarella, L., Cappelletti, A., Calori, G., Cionni, I., Finardi, S., Larsen, B., Pace, G., Perrino, C., Piersanti, A., Silibello, C., Vitali, L., Zanini, G., 2016. Impact of grid resolution on aerosol predictions: a case study over Italy. Aersol Air Qual. Res. 16, 1253-1267.

Miskell, G., Pattinson, W., Weissert, L., Williams, D., 2019. Forecasting short-term peak concentrations from a network of air quality instruments measuring PM2.5 using boosted gradient machine models. Journal of Environmental Management, 242, 56-64

Nabielek, K., Hamers, D., Evers, D., 2016. Cities in Europe Facts and figures on cities and urban areas. PBL Netherlands Environmental Assessment Agency The Hague,ISBN 978-94-91506-97-0.

Neal, L.S., Agnew, P., Moseley, S., Ordónez, C., Savaaage, N.H., Tilbee, M., 2014. Application of a statistical post-processing technique to a gridded, operational, air quality forecast. Atmos. Environ. 98: 385-393

Nenes A., Pandis S.N., Pilinis C., 1998. ISORROPIA: A new thermodynamic equilibrium model for multiphase multicomponent inorganic aerosols. Aquatic Geochemistry, 4(1), 123-152, doi:10.1023/A:1009604003981.

Nieuwstadt, F. T. M., 1981. The steady-state height and resistance laws of the nocturnal boundary layer: Theory compared with Cabauw observations. Boundary Layer Meteorology, 20, 3-17, doi:10.1007/BF00119920.

Peng, Z., Lei, L., Liu, Z., Sun, J., Ding, A., Ban, J., Chen, D., Kou, X., and Chu, K., 2018. The impact of multi-species surface chemical observation assimilation on air quality forecasts in China. Atmospheric Chemistry and Physics, 18, 17387–17404, https://doi.org/10.5194/acp-18-17387-2018.

Ponti, G., Palombi, F., Abate, D., Ambrosino, F., Aprea, G., Bastianelli, T., Beone, F., Bertini, R., Bracco, G., Caporicci, M., Calosso, B., Chinnici, M., Colavincenzo, A., Cucurullo, A., Dangelo, P., De Rosa, M., De Michele, P., Funel, A., Furini, G., Giammattei, D.,

10 Giusepponi, S., Guadagni, R., Guarnieri, G., Italiano, A., Magagnino, S., Mariano, A., Mencuccini, G., Mercuri, C., Migliori, S., Ornelli, P., Pecoraro, S., Perozziello, A., Pierattini, S., Podda, S., Poggi, F., Quintiliani, A., Rocchi, A., Scio, C., Simoni, F., Vita, A.,2014. The role of medium size facilities in the HPC ecosystem: the case of the new CRESCO4 cluster integrated in the ENEAGRID infrastructure", Proceedings of the 2014 International Conference on High Performance Computing and Simulation, HPCS 2014, art. no. 6903807, 1030-1033.

15 Pope III., C.A., Coleman, N., Pond, Z.A., Burnett, R.T., 2019. Fine particulate air pollution and human mortality: 25+ years of cohort studies. Environmental Research, in press, doi: https://doi.org/10.1016/j.envres.2019.108924

Schell, B., Ackermann, I. J., Hass, H., Binkowski, F. S., Abel, A., 2001. Modeling the formation of secondary organic aerosol within a comprehensive air quality modeling system. Journal of Geophysical Research, 106, D22, 28275-28293, doi:10.1029/2001JD000384.

Seinfeld, J.H. and Pandis, S.N., 1998. Atmospheric chemistry and physics: from air pollution to climate change. Ed. Wiley.

20 Silibello, C., Calori, G., Brusasca, G., Giudici, A., Angelino, E., Fossati, G., Peroni, E., Buganza, E., 2008. Modelling of PM10 concentrations over Milano urban area using two aerosol modules. Environmental Modelling & Software, 23, 333-343, doi:10.1016/j.envsoft.2007.04.002.

Silibello, C., Calori, G., Costa, M.P., Dirodi, M.G., Mircea, M., Radice, P., Vitali, L., Zanini, G., 2012. Benzo[a]pyrene modelling over Italy: comparison with experimental data and source apportionment. Atmospheric Pollution Research, 3, 399-407, doi:10.5094/APR.2012.046.

25 Smagorinsky, J., 1963. General circulation experiments with the primitive equations: 1. The basic experiment. Monthly Weather Review, 91(3), 99-164, doi:10.1175/1520-0493(1963)091<0099:GCEWTP>2.3.CO;2.

Tarrason, L., Hamer, P., Meleux, F., and Rouil, L., 2018. Interim Annual Assessment Report. European air quality in 2017, Tech. Rep. CAMS71_2018SC3_D71.1.1.10_IAAR2017_final. URL https://policy.atmosphere.copernicus.eu/reports/CAMS71_D71.1.1.10_201807_IAAR2017_final.pdf

30 Vautard, R., Bessagnet, B., Chin, M., Menut, L., 2005. On the contribution of natural Aeolian sources to particulate matter concentrations in Europe: Testing hypotheses with a modeling approach. Atmospheric Environment, 39(18), 3291 - 3303, doi:10.1016/j.atmosenv.2005.01.051.

Venkatram, A.,1980. Estimating the Monin-Obukhov lenght in the stable boundary layer for dispersion calculations. Boundary Layer Meteorology, 19, 481-485, doi:10.1007/BF00122347.

35 Vitali, L., Adani, M., Briganti, G., Cappelletti, A., Ciancarella, Cremona G., D'Elia, I., D'Isidoro, M., Guarnieri, G., Piersanti, A., Righini, G., Russo, F., Villani, M.G., Zanini, G., 2019. AMS-MINNI NATIONAL AIR QUALITY SIMULATION ON ITALY FOR THE CALENDAR YEAR 2015. Annual air quality simulation of MINNI Atmospheric Modelling System: results for the calendar year 2015 and comparison with observed dataENEA Technical Report, RT/2019/15/ENEA.

Uno, I., Osada, K., Yumimoto, K., Wang, Z., Itahashi, S., Pan, X., Hara, Y., Kanaya, Y., Yamamoto, S., and Fairlie, T. D., 2017. Seasonal variation of fine- and coarse-mode nitrates and related aerosols over East Asia: synergetic observations and chemical transport model analysis. Atmospheric Chemistry and Physics, 17, 14181–14197, https://doi.org/10.5194/acp-17-14181-2017.

World Health Assembly WHO, 2019. Healthy environments for healthier populations: Why do they matter, and what can we do?
5   Geneva: World Health Organization; https://www.who.int/publications-detail/healthy-environments-for-healthier-populations-why-do-they-matter-and-what-can-we-do

World Health Assembly, WHO, 2015. Health and the environment: addressing the health impact of air pollution: draft resolution proposed by the delegations of Albania, Chile, Colombia, France, Germany, Monaco, Norway, Panama, Sweden, Switzerland, Ukraine, United States of America, Uruguay and Zambia. World Health Organization. https://apps.who.int/iris/handle/10665/253206

10   Žabkar, R., Honzak, L., Skok, G., Forkel, R., Rakovec, J., Ceglar, A., Žagar, N., 2015. Evaluation of the high resolution WRF-Chem
       (v3.4.1) air quality forecast and its comparison with statistical ozone predictions. Geoscientific Model Development, 8(7), 2119–2137,
680   doi:10.5194/gmd-8-2119-2015.

       Zhang, K. M., Knipping, E. M., Wexler, A. S., Bhave, P.V., Tonnesen, G. S., 2005. Size distribution of sea-salt emissions as a function of
       relative humidity. Atmospheric Environment, 39, 3373-3379, doi:10.1016/j.atmosenv.2005.02.032.

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

685   Zhang, Y., Bocquet, M., Mallet, V., Seigneur, C., Baklanov, A., 2012b. Real-time air quality forecasting, part II: State of the science, current
       research, and future prospects. Atmospheric Environment, 60, 656-676.

       Zilitinkevich, S. S., 1972. On the Determination of the Height of the Ekman Boundary Layer. Boundary Layer Meteorology, 3, 141-145,
       doi:10.1007/BF02033914.