19th International Conference on
Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes
3-6 June 2019, Bruges, Belgium

THREE YEARS SIMULATION OF METEOROLOGICAL PARAMETERS AND AIRBORNE POLLUTANTS OVER ITALY FOR EXPOSURE ASSESSMENT OF POPULATION

Camillo Silibello[1], Giuseppe Calori[1], Sandro Finardi[1], Paola Radice[1], Francesco Uboldi[1], Massimo Stafoggia[2], Claudio Gariazzo[3], Giovanni Viegi[4], on behalf of the BEEP collaborative group

[1]ARIANET Srl, Milan, Italy
[2]Department of Epidemiology SSR Lazio/ASL Rome 1, Rome, Italy
[3]INAIL, Department of Occupational and Environmental Medicine, Monte Porzio Catone, Rome, Italy
[4] Institute of Biomedicine and Molecular Immunology (IBIM) National Research Council, Palermo, Italy

Abstract:
A modelling system, based on WRF and FARM chemical-transport model (CTM), has been used to perform three years (2013-2015), high-resolution (5x5 km), simulations over Italy of air pollution and meteorological parameters to be used to estimate exposure for national epidemiological applications. WRF model has been applied over two nested domains, covering Europe and Italy, while FARM has been applied over the Italian domain only. Air quality simulations have been performed using the meteorological fields provided by WRF simulations, chemical boundary conditions from the QualeAria forecast system and emission data from national and European inventories. A statistical analysis comparing modelling results against routine monitoring network data has been carried out to evaluate the simulation performance, indicating an acceptable model performance and confirming the suitability of models' results for the epidemiological investigations foreseen in the BEEP project. To improve the spatial distribution of airborne pollutants maps, the modelled concentration fields have been further integrated with the available observations using the Optimal Interpolation method.

Key words: *Air quality models, spatial analysis, data assimilation, emission inventories, exposure modelling.*

INTRODUCTION

Air pollution, especially particulate matter (PM), is one of the major threats to human health. Recently, the World Health Organization estimated around 4.2 million of premature deaths attributable to air pollution exposure worldwide (WHO, 2018). Similarly, air temperature is projected to increase due to global warming, with adverse consequences to human health and the ecosystem (IPCC, 2018). During the last decades, many epidemiological studies reported consistent health effects of PM and air temperature related to both short-term (i.e. daily peaks) and long-term (i.e. chronical) exposures. These studies have been mainly conducted in major cities, where larger populations live and more evidence on the health effects of air pollution is available. In such places, denser monitoring networks provide a basis on which to describe the spatiotemporal pollutants variability, as can be done through the application of Land Use Regression (LUR) models (Stafoggia *et al.*, 2013). Otherwise in smaller cities, sub-urban and rural areas, where observations are often insufficient, meteorological models and CTMs can effectively provide high spatial (down to about one km) and temporal (hourly) resolution data to investigate the health effects of air quality and meteorology on the population living in these areas. To estimate such effects on the Italian population, the National Institute for Insurance against Accidents at Work (INAIL) has funded the BEEP project (Big Data in Environmental and Occupational Epidemiology) which, among other objectives, planned the reconstruction of three years (2013-2015) of daily maps of air pollutants concentration and air temperature over Italy to support epidemiological investigations. Data fusion methods, that combine information from air quality monitoring networks with other sources such as reanalysis data, satellite data, and data sets obtained from statistical models or CTMs, have been used to improve the air quality assessment and consequently to reduce uncertainty in exposure estimates (Denby *et al.*, 2009; Physick *et al.*, 2007). In this work, we have used the Optimal Interpolation (OI) method to combine modelling results and observations. The modelling system used in this project and the obtained results are described in the next sections.

MODELLING SYSTEM DESCRIPTION

The adopted modelling system is based on FARM (Silibello *et al.*, 2014) and on *meteorology, emission* and *boundary-condition* modules. The meteorology module is made up by WRF prognostic non-hydrostatic model and an interface module that calculates further information required by FARM (i.e. gas-phase species deposition velocities, horizontal and vertical diffusivities, natural emissions of aerosols – sea salts and soil dust driven by surface wind - and trace species from vegetation) as function of meteorological parameters (i.e. wind speed, solar radiation, temperature) and geographic environment characteristics (i.e. soil type and land coverage). ECMWF ERA5 reanalyses have been used to drive WRF simulations that have been performed over two nested domains, covering Europe and Italy at 25 and 5 km resolution respectively. To improve the meteorological fields over the target domain, the observation nudging data assimilation scheme implemented in WRF has been applied using METAR, ship and buoy observations from NCEP/MADIS archives (an example of the spatial distribution of such observations for the year 2015 is given in Figure 1). The emission module generates hourly emissions at every cell of the Italian domain, by disaggregating ISPRA Italian national inventory and TNO inventory for the surrounding countries. Different temporal activity profiles (daily, weekly and yearly time modulations), and gridded spatial proxies have been used to disaggregate in time and space the emission inventories data. Activity-related speciation profiles have been used to split NMVOC and PM inventories emissions according to the chemical mechanism (SAPRC99, Carter, 2000) and the aerosol module (aero3, Binkowski and Roselle, 2003) implemented in FARM. Boundary conditions for the Italian grid have been supplied by QualeAria system (www.qualearia.it), that provides air quality forecasts over the Italian peninsula downscaling synoptic scale weather and chemical forecasts from US National Center for Environmental Prediction (NCEP) and Copernicus Atmosphere Monitoring Service (CAMS).

Figure 1. Monitoring sites used by WRF observation nudging scheme

SIMULATION RESULTS

The modelling system has been used to compute daily maps of PM_{10}, $PM_{2.5}$, NO_2 and O_3 levels over Italy along three years (2013-2015). The capability of the system to capture the spatiotemporal distribution of airborne pollutants has been assessed through a comparison with monitoring network data and the use of following statistical indicators (Table 1): the fraction of predictions within a factor of two of observations (FAC2) and the Index of Agreement (IA, Willmott, 1981). The analysis of the table evidences 'acceptable' model performance values for the considered pollutants: FAC2 > 60% and IA > 0.5 are generally considered to be a good result. Figure 2 shows scatterplots of average computed concentrations versus observed ones for the year 2015 for background monitoring sites located in northern, central and southern Italy. Traffic stations are not included in the analysis due to the horizontal resolution of the predicted fields that does not permit to capture the local phenomena induced by nearby road traffic emissions. As shown in Figure 2, the stations are not uniformly distributed across the country, with a larger number of monitoring sites in the northern portion of the territory. The analysis of scatterplots evidences better results in the northern portion of the country, with calculated values generally within a factor of two of observations. A good agreement has been obtained for PM_{10}, $PM_{2.5}$ and O_3, while NO_2 scatterplots evidences a more significant underestimation of observed levels at monitoring stations located in central and southern Italy. This result could be linked to higher uncertainties in emission data for these regions, but also to a general prevalence of a combination of smaller populated areas and a hilly/mountainous terrain, not fully resolved at the adopted model resolution.

Table 1. Model evaluation of PM_{10}, NO_2 and ozone predictions [μg m^{-3}] over Italy; FAC2: percentage of predictions within a factor of two of observations; IA: Index of Agreement ($IA \in [0,1]$, with 1 indicating the best agreement).

Pollutant	year	No. of sites	Mean Obs.	Mean Pred.	IA	FAC2
PM_{10}	2013	449	26.1	16.5	0.7	57.5
	2014	466	24.3	15.4	0.6	58.6
	2015	449	27.1	16.4	0.7	52.5
$PM_{2.5}$	2013	191	17.8	14.6	0.8	69.1
	2014	217	15.9	13.5	0.8	70.1
	2015	228	18.71	14.6	0.8	65.4
NO_2	2013	518	26.9	15.5	0.7	51.8
	2014	526	24.6	15.6	0.7	56.2
	2015	537	25.8	15.4	0.7	51.5
O_3	2013	289	54.3	56.4	0.8	86.8
	2014	300	52.4	56.6	0.9	87.0
	2015	283	56.9	64.9	0.9	84.9

To improve the concentration fields provided by the modelling system, we have then combined them with monitoring data, applying the OI algorithm. This method gives the best estimate of the chemical state of the atmosphere provided that suitable values are used for the scale parameters defining the influence of the observation along the horizontal and vertical directions (L_h and L_z). We have used the approach described in Silibello *et al.* (2014) to identify the values of these parameters to be used for the pollutants considered in this work. As for L_h we have considered values of 20, 30, 40, 50 km and 300, 500, 700 and 1000 m for L_z. For each combination of them, we performed the analysis and calculated the RMSE at each monitoring station, excluding its observations from the OI calculations ("leave-one-out cross validation"). Once identified the combination of L_h and L_z values (Table 2) that minimizes RMSE, we have then applied the OI method to the daily computed concentration fields for the three years.

Table 2. Values of L_h [km] and L_z [m] for PM_{10}, $PM_{2.5}$, NO_2 ed O_3.

PM_{10}		$PM_{2.5}$		NO_2		O_3	
L_h	L_z	L_h	L_z	L_h	L_z	L_h	L_z
20	1000	30	1000	40	1000	50	1000

Figure 3 reports, for the year 2015, FARM and OI yearly averaged concentration fields except for ozone, for which we have considered summer averages (June-August). Significant differences between FARM and OI maps were evidenced for NO_2, in central and southern Italy coherently with previous scatterplot analysis. As for PM_{10} and $PM_{2.5}$ the introduction of observations leads to higher levels along the

Apennine ridge (central Italy), while for ozone higher concentration were estimated in the North, along to eastern Alps.

CONCLUSION AND FUTURE WORK

In this work we presented the results obtained by the application of meteorological and chemical-transport models over Italy for three years considering a high spatial resolution (5 km). A good agreement between predicted and observed concentrations has been obtained for PM_{10}, $PM_{2.5}$ and O_3. As for NO_2, a significant underestimation of observed levels was evidenced at central and southern Italy monitoring stations that could be ascribable to uncertainties in NO_X emission data for these areas. To improve the air quality assessment and to reduce uncertainty in foreseen exposure estimates consequently, the Optimal Interpolation method has been used to integrate the estimated concentration fields with the available observation data. The "data fused" maps evidence, with respect to CTM results, a general increase of yearly NO_2 levels in central and southern Italy, higher PM_{10} and $PM_{2.5}$ levels along the Apennine ridge and higher ozone concentration along eastern Alps. Ongoing work concerns the application of machine learning methods, specifically the Random Forest (similarly to the approach described in Stafoggia *et al.*, 2019), to produce NO_2 and O_3 higher resolution maps (1 km) over Italy.

ACKNOWLEDGMENTS

This work paper has been funded by the National Institute for Insurance against Accidents at Work (INAIL) within the project "BEEP" (http://www.progettobeep.it, project code B72F17000180005, Bando INAIL Ricerche in Collaborazione BRiC 2016-2018).

REFERENCES

Binkowski, F.S, Roselle, S.J., 2003: Models-3 community multiscale air quality (CMAQ) model aerosol component 1. Model description. *Journal of Geophysical Research*, 10 8 (D6), 4183.

Carter, W.P.L., 2000: Documentation of the SAPRC-99 Chemical Mechanism for VOC Reactivity Assessment. Final Report to California Air Resources Board, Contract 92-329 and 95-308, SAPRC, University of California, Riverside, CA.

Denby, B., Garcia, V., Holland, D.M, Hogrege, C., 2009: Integration of Air Quality Modeling and Monitoring Data for Enhanced Health Exposure Assessment. *Environ. Manag.,* 46-49.

IPCC, 2018: Global Warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty [V. Masson-Delmotte, P. Zhai, H. O. Pörtner, D. Roberts, J. Skea, P.R. Shukla, A. Pirani, W. Moufouma-Okia, C. Péan, R. Pidcock, S. Connors, J. B. R. Matthews, Y. Chen, X. Zhou, M. I. Gomis, E. Lonnoy, T. Maycock, M. Tignor, T. Waterfield (eds.)].

Physick, W.L., Cope, M.E., Lee, S., Hurley, P.J., 2007: An approach for estimating exposure to ambient concentrations. *Journal of Exposure Science and Environmental Epidemiology*, **17**, 76–83.

Silibello, C., Bolignano, A., Sozzi, R., Gariazzo, C., 2014: Application of a chemical transport model and optimized data assimilation methods to improve air quality assessment. *Air Qual. Atmos. Health*, 7, 3, 283-296.

Stafoggia, M., Schwartz, J., and others, 2016: Estimation of daily PM_{10} concentrations in Italy (2006–2012) using finely resolved satellite data, land use variables and meteorology. *Environ. Int.*, **99**, 234–244.

Stafoggia, M., Bellander, T., and others, 2019: Estimation of daily PM_{10} and $PM_{2.5}$ concentrations in Italy, 2013–2015, using a spatiotemporal land-use random-forest model. *Environ. Int.*, **124**, 170–179.

Willmott, C.J., 1981: On the validation of models. *Physical Geography*, 2,184-194.

WHO (World Health Organization), 2018: available at https://www.who.int/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health, Accessed date: 19 March 2019.

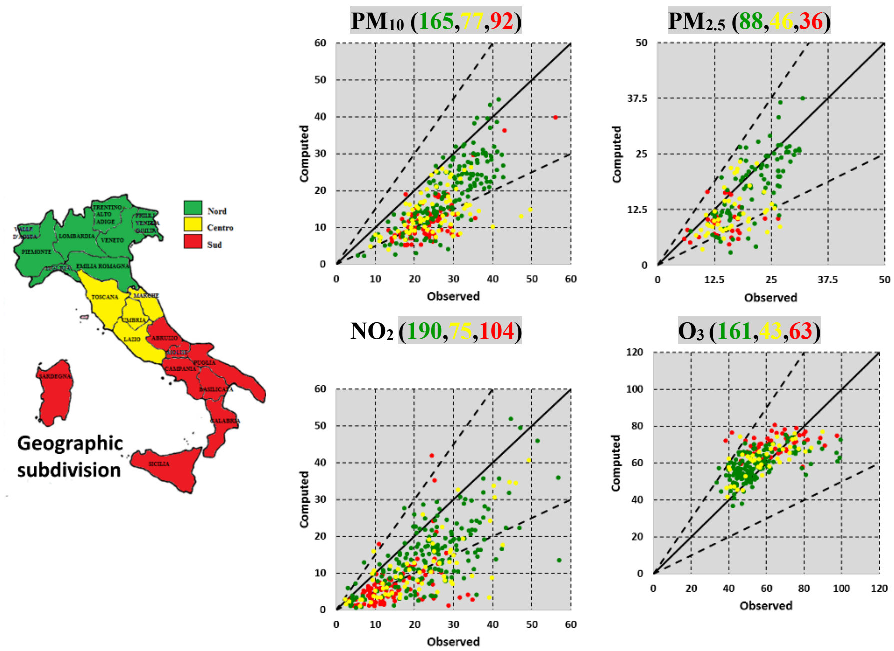

Figure 2. Yearly averaged (2015) PM_{10}, NO_2 and O_3 computed vs observed concentration scatterplots over Northern (green), Central (yellow) and Southern (red) Italy (between parenthesis the number of stations)

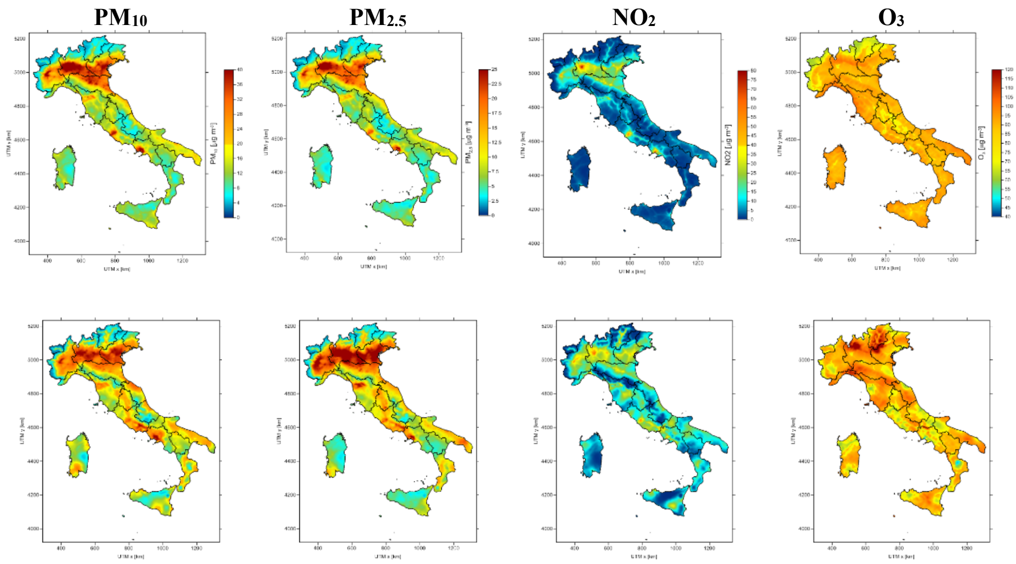

Figure 3. Yearly (PM_{10}, $PM_{2.5}$ and NO_2) and summer (JJA, O_3) levels over Italy, for the year 2015, computed by FARM (top) and OI algorithm (bottom)