FORECASTING SYSTEM - PARMA

Deliverable D.T2.2.4

Version final
30/09/2020

Leading role: ARPAE

Stefano Zauli Sajani Tel.: +39 059 433626
Email: szauli@arpae.it

Stefano Marchesi Tel.: +39 059 433635
Email: smarchesi@arpae.it

Roberta Amorati Tel.: +39 051 525924
Email: ramorati@arpae.it

Giulia Giovannini Tel.: +39 051 525924
Email: ggiovannini@arpae.it

Michele Stortini Tel.: +39 051 525 916
Email: mstortini@arpae.it

Description of the aims

Forecasting systems (1 per each city) able to announce the SAPEs, to categorize them (per different level of pollution) and to evaluate the realiability and effectiveness of provisional systems (linked to AWAIR-APP D.T2.2.7).

Table of contents

1. Introduction

We define atmospheric model any mathematical procedure giving as a result an estimation of ambient air quality entities (i.e. concentrations, deposition, exceedances). In general, a distinction between process-oriented models and statistical models can be made. Process oriented models are based on the description of physical/chemical processes: starting with emissions, atmospheric advection, dispersion, chemical transformation and deposition air pollutant concentrations are calculated. Statistical models are based on interpolation and extrapolation of measured data and are valuable tools in the diagnosis of present and future air quality without any understanding of cause-effect relationships. Once a model has been developed, the further application of the model will be relatively cheap; however, collecting the necessary input data might be cumbersome. In addition, uncertainties in model results may be large due to the combined effect of shortcomings in model concept and input parameters (e.g. emission data, meteorology). Therefore, model results are a representation of real world only to a limited degree as well as in most models an implicit spatial and temporal average is necessarily introduced which may make inappropriate a direct comparison with measurements at one location at a given moment.

Models describing the dispersion and transport of air pollutants in the atmosphere can be distinguished on many grounds, such as:

- on the spatial scale (Global; Macroscale - characteristic lengths exceeding 1,000 km; Mesoscale - characteristic lengths between 1 and 1000 km; Microscale (characteristic lengths below 1 km);
- on the temporal scale (episodic models, long-term models);
- on the treatment of the transport equations (Eulerian model, Lagrangian models);
- on the treatment of various processes (chemistry, wet and dry deposition).

In the following chapters the forecasting system implemented by Arpae is described in detail. Within the scope of the AWAIR project two main activities were carried out. First, a regular flow of air quality forecasts for the Parma FUA has been set up to allow for optimal operation of the AWAIR app. Second, a specific assessment of the forecasting system performance in the PARMA FUA has been performed and compared with the performances of the orher models operated in the other FUAs.

2. The air quality forecasting system managed by Arpae Emilia-Romagna

Arpae Emilia-Romagna operates a complete set of air quality models depending on the specific purposes. In particular, three main types of modeling systems are applied:

- ADMS-Urban. ADMS-Urban is an advanced three dimensional quasi-Gaussian model designed to study dispersion in the atmosphere of passive, buoyant or slightly dense, continuous releases from single or multiple sources (including point area or line –road- sources). It is mainly applied to air quality assessment studies in urban areas or to study dispersion plumes in area with simple orography.
- LAPMOD. LAPMOD is a Lagrangian model mainly used for detailed analysis in situations of complex terrain in the presence of a limited number of sources.
- NINFA. NINFA is the operational AQ model of the Environmental Agency of the Emilia-Romagna Region (ARPAE). The model suite includes the CTM model CHIMERE, the National NWP model COSMO used by the National Civil Protection Department l and a statistical post-processing module called IBIS.

The first two modeling systems are used for case study applications at local scale and for primary pollutant while NINFA model suite is the modeling system used to deliver daily air quality forecasts and assessment.

CHIMERE (https://www.lmd.polytechnique.fr/chimere/) is the core component of the NINFA modelling system. CHIMERE is an open-access multi-scale Eulerian chemistry transport model which simulates transport, dispersion, chemical transformations and deposition (dry and wet) of air pollutants and aerosols. The concentrations are computed by solving the continuity equation for processes such as emissions, transport, deposition, chemical reactions, and aerosol dynamics. The selection of appropriate parametrization schemes, which characterize the model, was based on previous modeling studies, as detailed in Chimere documentation The model produces daily full 3D pollutants concentrations in analysis and forecasts CHIMERE uses the MELCHIOR gas-phase chemistry mechanism. From the complete mechanism, a reduced mechanism system of equations has been derived according to the concept of chemical operators (the reduced mechanism includes 44 chemical species and about 120 reactions).

CHIMERE accounts for five types of emissions: Biogenic emissions, sea-salt emissions, dust emissions, anthropogenic emissions and fire emissions . The biogenic emissions are calculated using the MEGAN model, while the marine aerosol and dust emissions are calculated with different parametrization schemes.

The CHIMERE model accounts for the size distribution of the aerosols using a size-bin approach: the aerosol particles for each of the model species are distributed in N size bins, covering a diameter range from Dmin = 0.01 m to Dmax = 40 m, where the values 2.5 m and 10 m are retained as cut off diameters to allow a meaningful evaluation of PM2.5 and PM10 in the model, with quantities typically available from routine measurements. The main aerosol processes considered in CHIMERE are coagulation, deposition (dry and wet), the absorption of the semi-volatiles, and nucleation of sulfuric acid. The internally mixed particles are assumed to be made of anthropic primary particulate matter (PPM), the composition of which is unspecified: sulfates, nitrates, ammonium, secondary organic species (SOA), and telluric dust or particulate matter resuspended by wind and turbulence.

The model COSMO (http://www.cosmo-model.org) is the official Limited Area atmospheric model, which is used at the Italian National Civil Protection Departmentand developed and maintained in the framework the COnsortium for Small-scale MOdelling. It is a nonhydrostatic limited-area atmospheric It is a nonhydrostatic limited-area atmospheric prediction model based on the primitive thermohydrodynamic equations describing compressible flow in a moist atmosphere with a variety of physical processes taken into account by parameterization schemes. It has been designed for both operational numerical weather prediction (NWP) and various scientific applications on the meso- (2–20 km) and meso- (20–100 km) scale. In the air quality system, the COSMO model operates in two main configurations: One with a 5-km grid spacing and 45 vertical levels over the Mediterranean area (COSMO-5M) and one with a 2.2-km grid spacing and 65 vertical levels over Italy (COSMO-2I). The CHIMERE model was set up with a specific interface able to read the meteorological variables given by the COSMO model and add to variables, such as the planetary boundary layer height.

The time-dependent boundary conditions (with hourly frequency) can be provided either by PREV'AIR (the French national platform of quality air forecast covering Europe), or by the Italian CAMS downstream service (https://doi.org/10.3390/atmos11050447) .

The specific application of the Chimere model in Emilia-Romagna is sketched in figure 1. Data related to pollutant emission are based on:

- the regional inventory of the emission in the atmosphere fulfilled by Arpae using INEMAR software on behalf of Emilia-Romagna Region;
- the national inventory of the pollutant emission published by Ispra for the Italian territory outside Emilia-Romagna;
- MACC Project inventory for the part of the model domain falling outside Italian territory (including the Mediterranean Sea area).

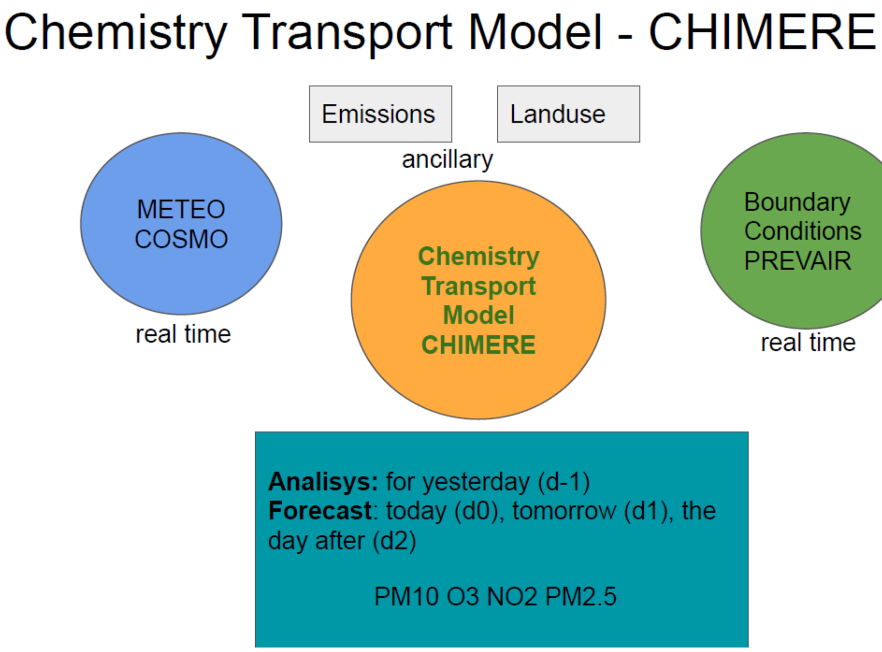

Figure 1. Input and output data for the Chimere modeling system

NINFA is run by Arpae Emilia-Romagna in nested configuration. In particular, the current nesting chain includes a run at 5 km over Northern Italy and a run at 3 km resolution focused on Emilia-Romagna region (figure 2). Daily forecasts are issued everyday on a dedicated area of the Arpae website (https://www.arpae.it/it/temi-ambientali/aria/previsioni/previsioni-di-qualita-dellaria/previsioni-di-qualita-dellaria).

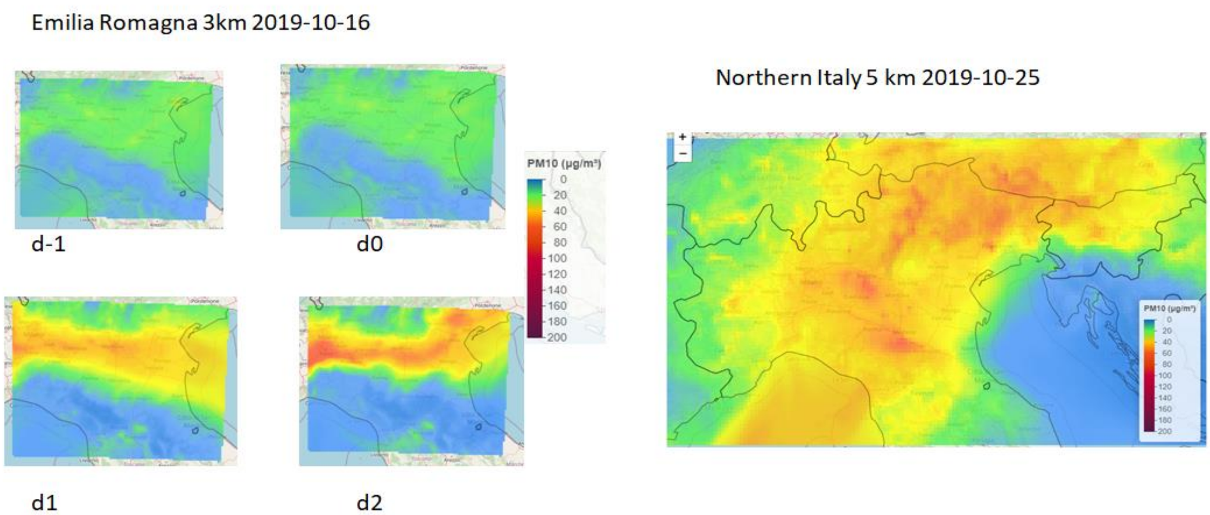

Figure 2. Nested operation application of Chimere model

Concentration fields are post-processed by a statistical Bayesian model called IBIS (Inferenza Bayesiana Inquinamento Simulato). IBIS model uses time series of measured data to correct model output and produce probabilistic predictions for the same day and the next two days (figure 3).

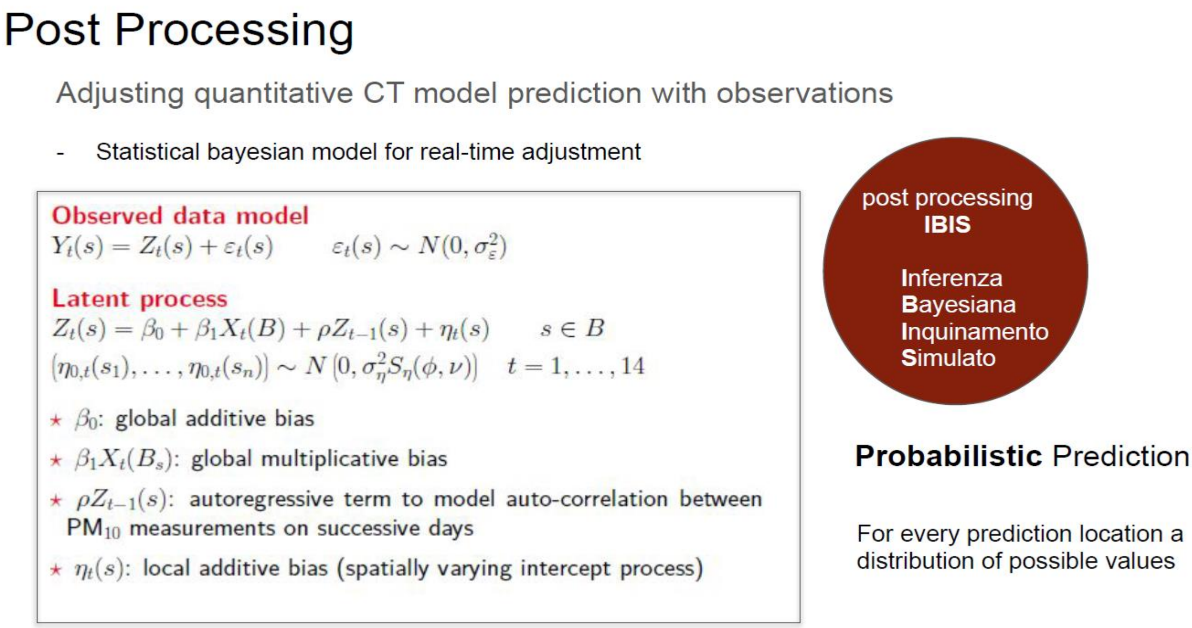

Figure 3. Summary mathematical description of the post-processing model IBIS

While air quality forecast maps and forcasts for the locations where of the fixed site monitoring are provided by IBIS model, air quality maps of the previous day are produced by a geostatistical post-processing that estimates urban and rural background concentrations with a 1km horizontal resolution. A kriging geostatistical technique is used to merge model concentration fields with in situ observations for PM_{10}, $PM_{2.5}$, NO_2 and O_3.

Table 1. Kriging procedure for O_3, NO_2, PM_{10}, $PM_{2.5}$

Pollutant	**Proxy 1**	**Proxy 2**
O_3	Height above sea level	NOX emissions
NO_2	Height above sea level	NOX emissions
PM_{10}	Height above sea level	PM_{10} emissions
$PM_{2.5}$	PM_{10} emissions	-

Since the horizontal resolution is 1 km, only data from monitoring stations which can be representative of an area with a radius of at least 1 km are included in the post-processing procedure. Traffic or industrial stations are not considered since they represent only local situations.

3. The Arpae air quality index

The overall status of air quality is synthesized by Arpae in a simplified way using an air quality index (AQI).

The pollutants usually included in the definition of air quality indices at international level are those associated with short-term health effects, such as carbon monoxide (CO), nitrogen dioxide (NO_2), ozone (O_3), sulfur dioxide (SO_2), particulate matter (total suspended particles, PM_{10} or $PM_{2.5}$). The underlying idea is that indices should be used for everyday information to citizens in order to increase people awareness, and enable actions to reduce exposure and health effects.

Regarding the Arpae AQI, only PM_{10}, NO_2 and O_3 are included in the definition. CO and SO_2 concentrations were excluded from the index computation due to the dramatic concentration drop observed during the last decades and the current levels often below their limit of detection. $PM_{2.5}$ was not included as well, since no daily limit values are available.

The computation of the Arpae AQI consists of two main steps: a) the construction of a dimensionless scale for each pollutant (sub-index); and b) the construction of a unique, synthetic index starting from the afore-mentioned sub-indices. The sub-indices for each pollutant are computed dividing the measured or forecasted pollutant concentration by the corresponding EU limit value. The value is then multiplied by 100. The following Table 2 shows the limit values for each pollutant.

Pollutant	Reference indicator	Value
PM_{10}	Daily mean value	50 µg/m^3
O_3	Daily maximum value of the 8-hour moving average	120 µg/m^3
NO_2	Daily maximum value	200 µg/m^3

Table 2. EU limit values for PM_{10}, O_3, NO_2

The next step is the aggregation of the three sub-indices. According to the approach adopted for the construction of most indices at the International level, Arpae sets the value of the synthetic value of the air quality index as the value of worst sub-index. Therefore, an air quality index above 100 indicates that at least one of the pollutant concentration is above the law thresholds.

Values of the AQI are grouped in 5 classes with an associated color and qualitative adjective scale (Table 3). Figure 4 shows an example of AQI map produced by the Arpae Emilia-Romagna system.

Index value	Color scale	Air quality (IT)	Air quality (EN)
<50		"Buona"	Good
50-99		"Accettabile"	Fair
100-149		"Mediocre"	Moderate
150-199		"Scadente"	Poor
>200		"Pessima"	Very poor

Table 3. Intervals of air quality index in Emilia-Romagna

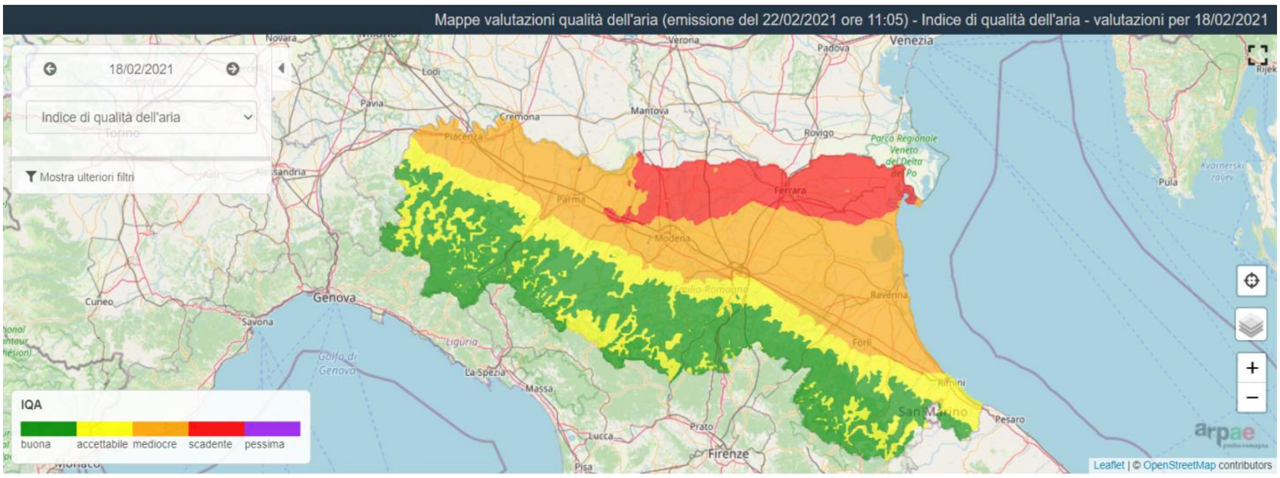

Figure 4. Example of map produced by the post-processing

4. Performance of the forecasting system in the Parma FUA

AWAIR partnership agreed in assessing the performance of the forecasting systems used in the three FUAs using a common set of indicators. It is worth to highlight that in the AWAIR APP the status of air pollution will be communicated by means of the Arpae AQI. However, in Emilia-Romagna region AQI calculation during winter is mainly related to the PM_{10} sub-index and all exceedances of limit values are due to PM_{10} concentrations. According to these assumptions we calculated performance indicators considering directly PM_{10} concentrations instead of AQI values.

Moving back to the specific set of indicators, in addition to Pearson correlation coefficient calculated between measured and forecasted concentration of PM_{10}, partners agreed on two selected sets of statistical indices. The first group includes Mean Error (ME, which corresponds to the bias), Mean Absolute Error (MAE), and Root Mean Square Error (RMSE) calculated with the following formulas (where predicted and observed daily values are indicated as P_i and O_i, respectively):

$$ME = \frac{1}{N}\sum_{i}^{N}(P_i - O_i)$$

$$MAE = \frac{1}{N}\sum_{i}^{N}|P_i - O_i|$$

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

The second group focuses on the performance of the forecasting system in predicting PM_{10} exceedances at different thresholds (50, 75 and 100 $\mu g/m^3$) and includes indices based on a contingency tables. The selected indices are Probability Of Detection (POD), False Alarm Ratio (FAR) and Threat Score (TS). Here below the definitions:

$$POD = \frac{A}{A + C}$$

$$FAR = \frac{B}{A + B}$$

$$TS = \frac{A}{A + B + C}$$

where "A", "B", "C" are the entries of the 2x2 contingency table and stand for the number of "hits" (events observed and predicted), "false alarms" (events predicted, but not observed) and "misses" (events observed, but not predicted) at each selected threshold.

The performance assessment was based on the winter period 2019-2020 because on September 2019 the forecasting system underwent an important upgrade. Unfortunately, no days with PM_{10} concentrations above 100 $\mu g/m^3$ and only 5 days with concentration above 75 $\mu g/m^3$ were observed. Therefore, the selected indices based on contingency tables were calculated taking into account only the threshold 50 $\mu g/m^3$. The calculated statistical indices are reported in table 4.

Table 4. Overview of the results of the performance assessment results of the forecasting systems in the Parma FUA

		d0	d1	d2
	MAE	9.5	12.2	13.4
	ME	-1.5	-0.5	0.27
	RMSE	12.7	15.6	17.5
	COR	0.77	0.66	0.61
>50 µg/m³	POD	0.71	0.64	0.75
>50 µg/m³	TS	0.51	0.45	0.5
>50 µg/m³	FAR	0.35	0.4	0.4

The main conclusions we can draw are the following:

- The model seems reasonably able to predict PM_{10} concentrations and exceedances of the daily limit value;
- The model shows a slight decrease of performance from d0 (forecasts for the same day) to d2 (forecasts for the day after tomorrow)

It is important to highlight that the performance assessment was specific for the Parma FUA and therefore it should not be considered as a general performance assessment of the model.