**Tethys**

Journal of Mediterranean
Meteorology & Climatology

# A performance evaluation of the CAMx air quality model to forecast ozone and PM10 over the Italian region of Veneto

**A. Dalla-Fontana**[1]**, S. Pillon**[1] **and S. Patti**[1]

[1]Agenzia Regionale per la Prevenzione e Protezione Ambientale del Veneto (ARPAV)

*Correspondence to:* alberto.dallafontana@arpa.veneto.it

## Abstract

*Since air pollution impacts our health and the environment, air quality forecasts represent information of primary importance, especially for an area susceptible to episodes of pollution such as the Po Valley. The Environmental Protection Agency of the Italian region of Veneto (ARPAV) has implemented a system based on the CAMx photochemical model to predict concentrations of particulate matter and ozone. Meteorological input is estimated on the basis of the numerical output from the COSMO limited area model. Emissions are calculated using the INEMAR software and the concentrations at the boundaries of the computational domain are provided by the PREV'AIR modelling system. The system generates, on a daily basis, forecasts of up to three days and produces interactive maps which can be viewed from the agency's website. An evaluation of the forecasts of 1hr-ozone daily maximums and PM10 daily means has been conducted, by means of a statistical comparison with a set of air quality stations. The result is satisfactory since the forecasts comply with the benchmarks established for the model evaluation. The forecast skill respect to the persistence method has also been investigated.*

## 1 Introduction

According to European Air Quality Directive (2008/50/EC), air quality forecasts provide useful information about the potential risks of exceedance of alert thresholds or limit values of pollutants, and can help authorities to take temporary control measures to limit the impact of such events (EEA, 2011). To accomplish this, the most widespread approach is usage of Eulerian-grid models in "forecast mode": coupled with numerical outputs produced by a meteorological prognostic model. Such systems can give, in a relatively short time, a prediction of the future state of atmospheric pollutants over a wide area and high spatial resolution. Although the model quality objectives described in the AQ Directive only apply to the assessment of the current air quality, there is clearly an expectation, when using forecast models, that they have been verified and validated in an appropriate way (Denby et al., 2011). Accordingly, several studies about air quality forecast

verification have emerged in recent years. Honoré et al. (2008) presented a long-term evaluation of the forecasts of PM10 and ozone issued by the French national platform for air quality, PREV'AIR, based on an implementation of the CHIMERE model. Arasa et al. (2010) presented a performance evaluation of the MM5/MNEQA/CMAQ air quality modelling system to forecast ozone concentrations in Catalonia. D'Allura et al. (2018) presented a performance evaluation of QualeAria, a forecast system based on the RAMS meteorological model coupled with the FARM photochemical model, providing forecasts for the whole of Italy. A number of forecast maps can be accessed online: at a continental scale, the Copernicus Atmosphere Monitoring Service (CAMS) delivers an ensemble forecast based on nine numerical air quality models developed in Europe, and in Italy, the previously mentioned QualeAria modelling system, in addition to forecasts, provides the boundary conditions for regional/local forecasting systems implemented by some of the Italian Regional Environmental Agencies.
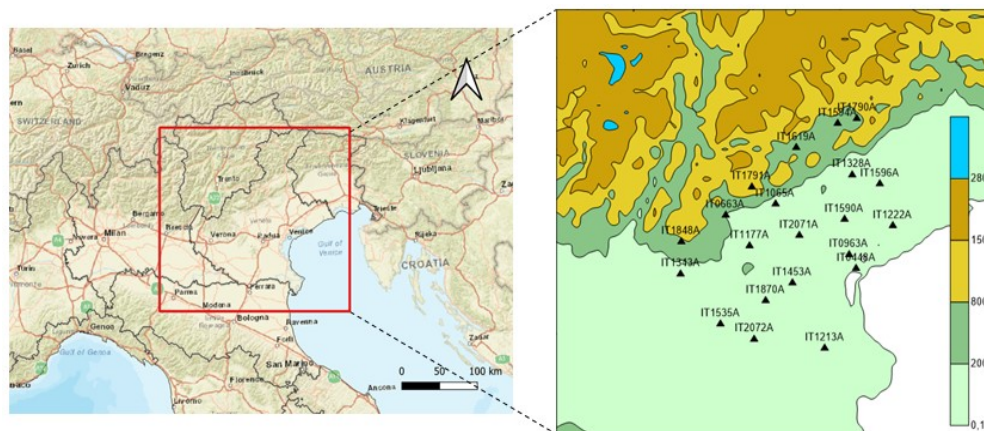
**Figure 1**: Geographic map of Northeast Italy; the domain of the CAMx model is highlighted to the right. Terrain elevation and stations used for model evaluation are shown.

As for the region of Veneto, ARPAV has long since implemented the INEMAR, a "bottom-up" oriented emissions inventory, at a regional scale, and recently developed a meteorological processor that, based on prognostic numerical outputs from the meteorological prognostic model COSMO, can provide the required input data to drive the chemical transport model, CAMx. This led to the implementation of the forecasting system whose evaluation constitutes the objective of this research. The study will focus on forecasts of $O_{3,max}$ (Ozone daily 1-hour peak), issued in summer, and forecasts of PM10 (daily mean), issued in winter. In fact, on a regional scale, these are the pollutants of concern since exceedances of their thresholds foreseen by the AQ directive are not infrequent. For PM10 in particular, although noticeable improvements have been seen in recent years, the maximum allowed number of exceedances of the daily threshold is surpassed every year for most stations in the Po Valley.

The paper is subdivided into four sections. Section 2 illustrates the most relevant inputs and aspects of the modelling system; technical details about steps to translate "raw" emissions and meteorological fields into final input files for the CAMx are provided in a separate Appendix. Section 3 illustrates a verification of the forecasts of $O_{3,max}$ and PM10 issued in a three-month period, in summer and winter respectively. A statistical comparison with measurements is presented: performance metrics and categorical statistics quantifying the capability of the system to forecast exceedances of a given threshold were elaborated. The forecast skill respect to the persistence model was also investigated. Section 4 discusses the results of the evaluation carried out in the previous section pointing to possible interpretations. Section 5 summarizes the relevant conclusions of the research.

## 2   General description of the forecasting system

The air quality forecasting system provides forecasts of up to three days (starting from the current day: D+0,

D+1, D+2). The core of the modelling system is the CAMx photochemical model (Comprehensive Air Quality Model with Extensions). The model, managed and distributed by Ramboll Environ (http://www.camx.com/), is continuously developed and has reached version 7 (version 6.5 is currently implemented in our system). Although the CAMx incorporates two-way grid nesting capability, in order to limit the requirements in terms of computing power and disk space, the model runs on a single grid with a cell size of 4 km, which is a trade off between the necessity to resolve local features retaining a domain large enough to include the study area. The domain grid consists of 64 x 59 horizontal cells and eleven levels, terrain-following, with the height of each layer interface ranging from 20 m up to 6,000 m above ground level. The study area is the north-east of Italy; it is a densely populated area with many industrial and craft activities and large cultivated areas. Fig. 1 shows the simulation domain of the CAMx model, bounded to the west and north by the Alps and to the east by the Adriatic Sea, covering a 256-km x 236-km area and including the whole region of Veneto. Terrain altitudes range from sea level to greater than 3,000 m in the Alps. CAMx inputs are developed using specifically devised processing tools that characterize meteorology, emissions and other environmental conditions (Fig. 2).

Emissions are based on the regional scale inventory INEMAR (Air EMissions INventory, http://www.inemar.eu). The inventory is adopted by almost all the regions of Northern Italy and follows a "bottom-up" approach, beginning with local data and, using this information and proper emission factors, it assesses emissions at a municipal level. The INEMAR software is largely based on the EMEP/EEA methodology for the definition of estimation methods and emission factors. The macro-pollutants in the inventory are $CH_4$ (methane), CO (carbon monoxide), $CO_2$ (carbon dioxide), VOC (volatile organic compounds), $N_2O$ (nitrous oxide), $NH_3$ (ammonia), $NO_x$ (nitrogen oxides), PTS (total suspended dust), PM10 (fine dust with an aerodynamic diameter of less than 10 μm), PM2.5 (fine dust with an aerodynamic diameter of less than 2.5

**Table 1**: Characteristics of stations used for model evaluation. T* is the threshold used to evaluate categorical statistics. Station sites are shown in Fig. 1.

| Station code | Classification | Altitude (m a.s.l.) | Measured pollutants | T* ($O_{3,max}$) | T* (PM10) |
|---|---|---|---|---|---|
| IT1594A | Urban | 378 | PM10,$O_3$ | 129 | 33 |
| IT1619A | Suburban | 356 | PM10,$O_3$ | 143 | 55 |
| IT1790A | Rural | 690 | PM10,$O_3$ | 145 | 18 |
| IT1870A | Rural | 12 | PM10,$O_3$ | 154 | 60 |
| IT1791A | Rural | 1,366 | $O_3$ | 174 | n.a. |
| IT1065A | Urban | 114 | $O_3$ | 158 | n.a. |
| IT0663A | Urban | 190 | PM10,$O_3$ | 163 | 46 |
| IT1848A | Rural | 814 | PM10,$O_3$ | 169 | 23 |
| IT1328A | Urban | 61 | PM10,$O_3$ | 151 | 42 |
| IT1343A | Urban | 88 | PM10,$O_3$ | 162 | 63 |
| IT2071A | Rural | 24 | PM10,$O_3$ | 147 | 69 |
| IT2072A | Rural | 6 | PM10,$O_3$ | 142 | 73 |
| IT1213A | Urban | 4 | PM10,$O_3$ | 135 | 65 |
| IT1596A | Rural | 8 | PM10,$O_3$ | 140 | 60 |
| IT1590A | Urban | 15 | PM10,$O_3$ | 146 | 75 |
| IT1222A | Urban | 0 | $O_3$ | 133 | n.a. |
| IT0963A | Urban | 0 | PM10,$O_3$ | 140 | 65 |
| IT0448A | Urban | 0 | PM10,$O_3$ | 139 | 70 |
| IT1177A | Urban | 36 | PM10,$O_3$ | 166 | 69 |
| IT1535A | Urban | 13 | PM10,$O_3$ | 157 | 69 |
| IT1453A | Urban | 9 | $O_3$ | 147 | n.a. |
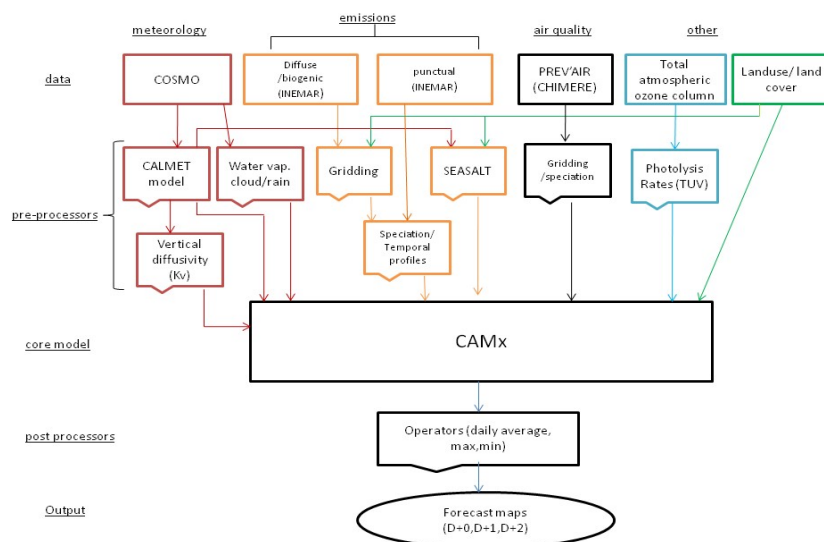


**Figure 2**: Schematic diagram of the modelling system.

μm), $SO_2$ (sulphur dioxide). Currently, the implemented version provides the yearly 2015 emissions and an update to 2017 is in progress. Since emissions required by the CAMx must be in the form of grid and point sources, "Diffuse" and "biogenic" emissions from the INEMAR inventory are gridded and assigned to the first model level whereas for "point" sources (industrial stacks), the level is assigned at each model time step according to its release height. Emissions are then speciated and modulated in time according to a set of emission sector-specific profiles (details in Appendix A).

Meteorological fields are provided by the COSMO-ITA model, the Italian version of the model developed by the Consortium for Small Scale Modelling (http://www.cosmo-model.org). The COSMO is a non-hydrostatic, limited-area atmospheric model, to be used both for operational and for research applications (Baldauf et al., 2011). The current operational resolution of the model is 5 km. The numerical outputs of COSMO-ITA, of up to D + 2, are recovered daily by the national meteorological service. COSMO fields include basic meteorological variables, available on an hourly basis on the pressure (isobaric) levels and surface. The interpolation of the variables to the CAMx grid is performed using the CALMET model; CALMET (Scire

**Table 2**: Definitions of the statistical metrics used for model evaluation.

| Statistics | Definition | Notes |
|---|---|---|
| Correlation coefficient | $r = \frac{\sum_i (M_i - \overline{M})(O_i - \overline{O})}{\sqrt{\sum_i (M_i - \overline{M})^2 (O_i - \overline{O})^2}}$ | $-1 \leq r \leq 1$; $r = 1$ : perfect correlation; $r = 0$ : totally uncorrelated |
| Normalized Mean Bias (%) | $NMB = \frac{\sum_i (M_i - O_i)}{\sum_i O_i} * 100$ | $-\infty < NMB < \infty$; $NMB = 0$ : perfect |
| Normalized Mean Error (%) | $NME = \frac{\sum_i |M_i - O_i|}{\sum_i O_i} * 100$ | $0 \leq NME < \infty$; $NME = 0$ : perfect |
| Probability of detection (POD) | $POD = \frac{H}{H+M}$ | $0 \leq POD \leq 1$; $POD = 1$ : perfect |
| False Alarm Ratio (FAR) | $FAR = \frac{F}{H+F}$ | $0 \leq FAR \leq 1$; $FAR = 0$ : perfect |
| Percent Correct (PC) | $PC = \frac{H+C}{H+M+F+C} * 100$ | $0 \leq PC \leq 100$; $PC = 100$ : perfect |

**Table 3**: Benchmarks established for model evaluation.

| Statistics | $O_{3,max}$ | | PM10 | |
|---|---|---|---|---|
| | Criteria | Goal | Criteria | Goal |
| NMB | $\pm 15\%$ | $\pm 5\%$ | $\pm 30\%$ | $\pm 10\%$ |
| NME | 25% | 15% | 50% | 35% |
| r | 0.5 | 0.75 | 0.4 | 0.7 |

et al., 2000) is a diagnostic meteorological model that reconstructs the 3D wind and temperature fields starting with meteorological data of input, orography and land use. CALMET was implemented in "no observation" mode, using gridded fields provided by the COSMO, rather than meteorological observations. The CALMET processor is needed to interpolate the wind and temperature on the CAMx grid and to determine the 2D fields of the micro meteorological variables (mixing height, Obukov length etc.) required to calculate the vertical diffusivity ($K_v$) using the methodology employed in the CMAQ model (Community Multiscale Air Quality, Byun (1999)). Other variables required by CAMx, quantifying the vapour and aqueous content of the atmosphere, are estimated based on the available variables (details in Appendix A).

Concentrations at the boundaries of the domain are provided by PREV'AIR (http://www2.prevair.org/), the French national platform for air quality, based on a implementation of the CHIMERE model at a European scale (Honoré et al., 2008). Concentrations of gasses (NO, $NO_2$, $O_3$, CO, $SO_2$, $NH_3$, VOCs) and PM species are allocated, by means of conversion tables, to CAMx species. Since we noticed that "desert dust" episodes simulated by CHIMERE often led to particulate peaks that were not actually observed by AQ stations, we resolved to suppress the crustal particulate fraction.

## 3   Statistical evaluation

Forecast verification "involves exploring and summarising the relationship between sets of forecast and observed data and making comparisons between the performance of forecasting systems and that of reference forecasts. Verification is therefore a statistical problem" (Jolliffe and Stephenson, 2012). Accordingly, we present an evaluation based on different statistical metrics, somewhat complementary, in order to provide enough information about the model's performance. Forecasts were investigated in the following periods:

- December 2018 - February 2019 (three months) for evaluation of PM10

- June - August 2019 (three months) for evaluation of $O_{3,max}$

ARPAV is in charge of the management of an extended network of air quality stations over the whole regional territory devoted to constantly measuring the concentration of major atmospheric pollutants. Among these substances, PM10 and ozone are of particular relevance since exceedances of their limit values are observed in many stations. The Po Valley is notoriously subjected to episodes of particulate matter pollution occurring in winter, during periods of dry weather and low wind conditions. On the other hand, in summer, conditions leading to exceedances of ozone information threshold are frequent. Accordingly, most stations are equipped with sensors to measure ozone and PM10. Ozone sensors transmit measurements in real time to a centralised server. As for PM10, some stations implement a beta gauge particulate sampler while others use the classic gravimetric method consisting of weighing the sampling filter after removal from the instrument. Air quality stations are classified based on their location, as "background" (stations in rural or urban areas), "traffic" or "industrial" (stations near sources of pollutants) stations. For the purpose of the present evaluation, a set of "background" stations were selected, since their representativeness is more comparable to the resolution of the model. Table 1 summarizes the typology and altitude of the stations considered for model evaluation.

The dataset amounts to 21 monitoring stations for $O_{3,max}$ (corresponding to approximately 1,932 observation-prediction pairs) and 17 stations for PM10 (1,530 pairs).
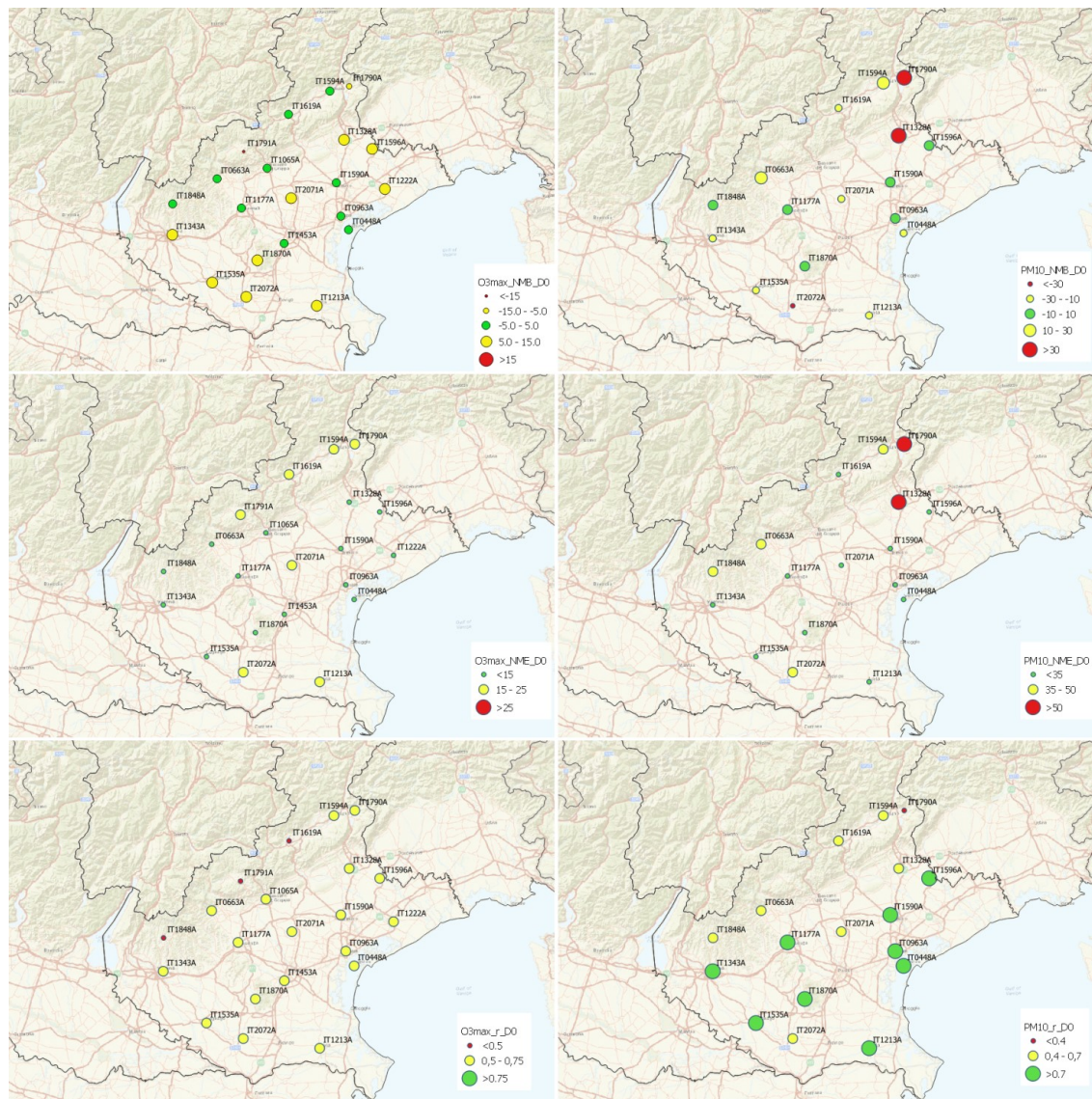
**Figure 3**: "Dot plots" of metrics used for model evaluation. From top to bottom: NMB, NME and $r$. $O_{3,max}$ is on the left and PM10 on the right. Dots are coloured green, yellow or red if they met the "Goal", "Criteria" or neither of the two, respectively. Only the results related to D+0 are shown since no big differences emerged at D+1 and D+2.

There are plenty of indicators that can be used to conduct a model performance evaluation (see EEA (2011) for a comprehensive list). We adopted the method proposed by Emery et al. (2017) which, based on previous research about the performance of photochemical models (Simon et al., 2012), proposed to update established benchmarks for ambient ozone (Doll, 1991) and PM concentrations (Boylan and Russell, 2006) and recommended the evaluation of three well-established statistical metrics: normalized mean bias (NMB), normalized mean error (NME) and correlation coefficient ($r$). Formal definition of these metrics is given in Table 2. For each metric, two benchmarks were developed: a minimum target ("Criteria") and a stricter target ("Goal"), both are reported in Table 3.

In fact, benchmarks had been established for 1-hour ozone (and MDA8 ozone) and PM2.5, while in the present study they are associated to $O_{3,max}$ and PM10. However,

since NMB and NME are normalized from the observed mean, there should be no large variation. As regards the correlation index, it is possible that the variations of $O_{3,max}$ are more difficult to capture than MDA8 and in this case, the benchmark may be considered too strict. Nevertheless, we resolved to refer to these benchmarks, as a minimum, as a guidance, and computed the metrics for each station. Fig. 3 illustrates the results: each site is marked with a dot with a size proportional to the value of the metric. For NMB, smaller dots indicate a bias that is less than the "Goal" and larger dots show a bias that is greater than the "Goal". Dots are coloured green, yellow or red if they met the "Goal", "Criteria" or neither of the two respectively.

The forecasts can be regarded as categorical when used to predict exceedance of a concentration threshold; in this case a yes-forecast is defined as an event in which an exceedance is predicted and a yes-event as an event for which
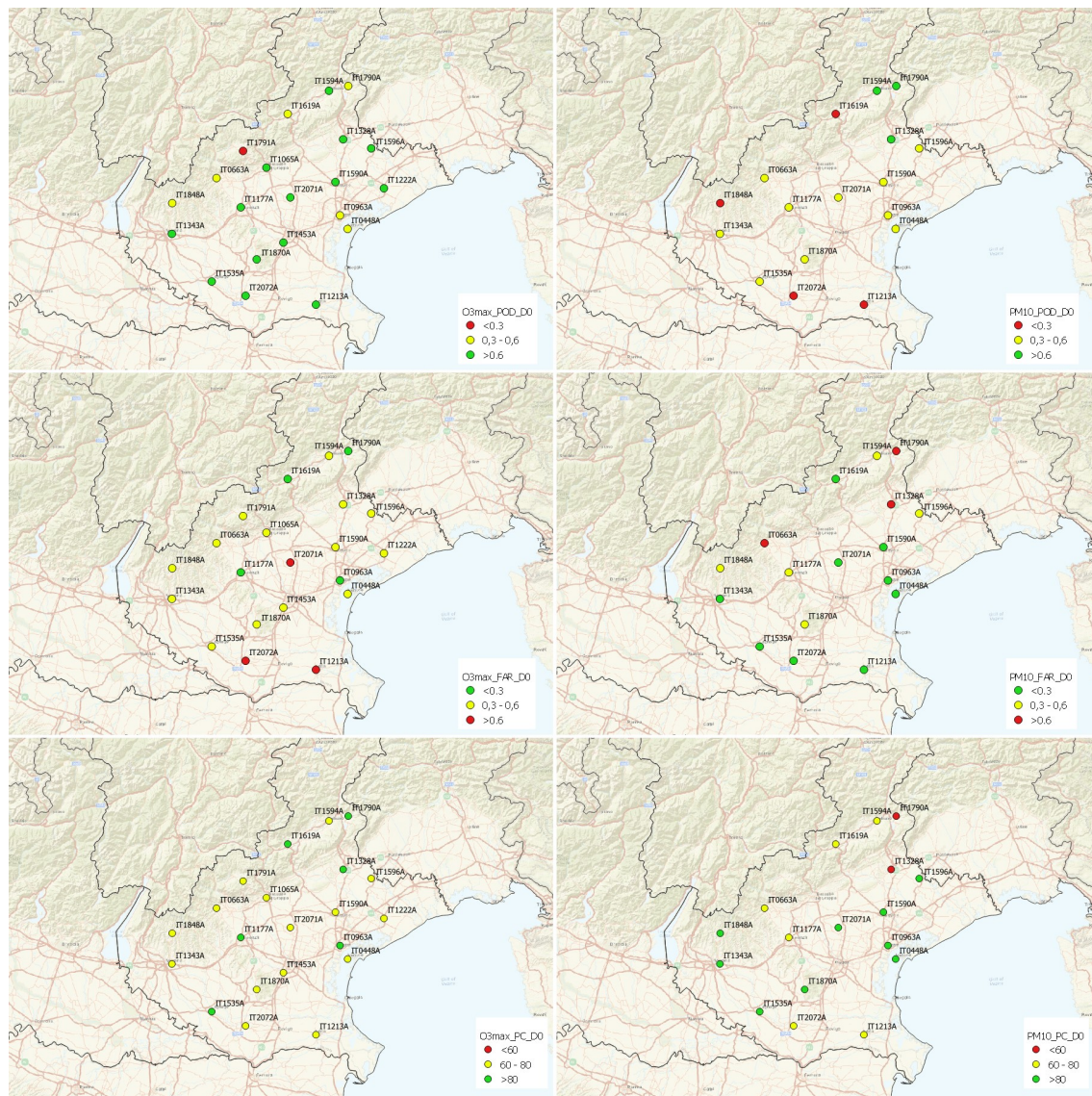
**Figure 4**: "Dot plots" of categorical statistics used for model evaluation. From top to bottom: POD, FAR and PC. $O_{3,max}$ is on the left and PM10 on the right. Only the results related to D+0 are shown since no big differences emerged at D+1 and D+2.

an exceedance is observed. A 2x2 contingency table can be defined that shows the frequency of yes and no forecasts and their corresponding occurrences. The elements of the table are:

- H (hits): number of yes-events correctly predicted

- M (missing alarms): number of yes-events not correctly predicted

- F (false alarms): number of yes-forecasts not actually observed

- C (correct rejections): number of no-events correctly predicted

A large variety of categorical statistics can be computed from the elements of the table. The most frequently used are POD (Probability Of Detection or Hit rate) and FAR (False

Alarms Ratio), as they can be seen as complementary. PC (Percent Correct or Accuracy) is also a useful score since it takes into account all the elements of the contingency table. Definitions of these scores are reported in Table 2; POD is the fraction of the yes-events correctly predicted, FAR is the fraction of yes-forecasts not actually observed, PC is the fraction of total events correctly predicted. Values of categorical statistics may depend on the selection of the threshold. Since exceedance of limits/targets foreseen by the European Directive for $O_{3,max}$ (180 $\mu g/m^3$) and PM10 (50 $\mu g/m^3$) is a relatively rare event for some stations, a threshold based on the 75th percentile of the observed concentrations was selected (Pay et al., 2014). Thresholds computed on this basis are reported for each station and pollutant in Table 1. Fig. 4 summarizes the results obtained for POD, FAR and PC. For each statistic, a colour scale was devised with tentative steps.
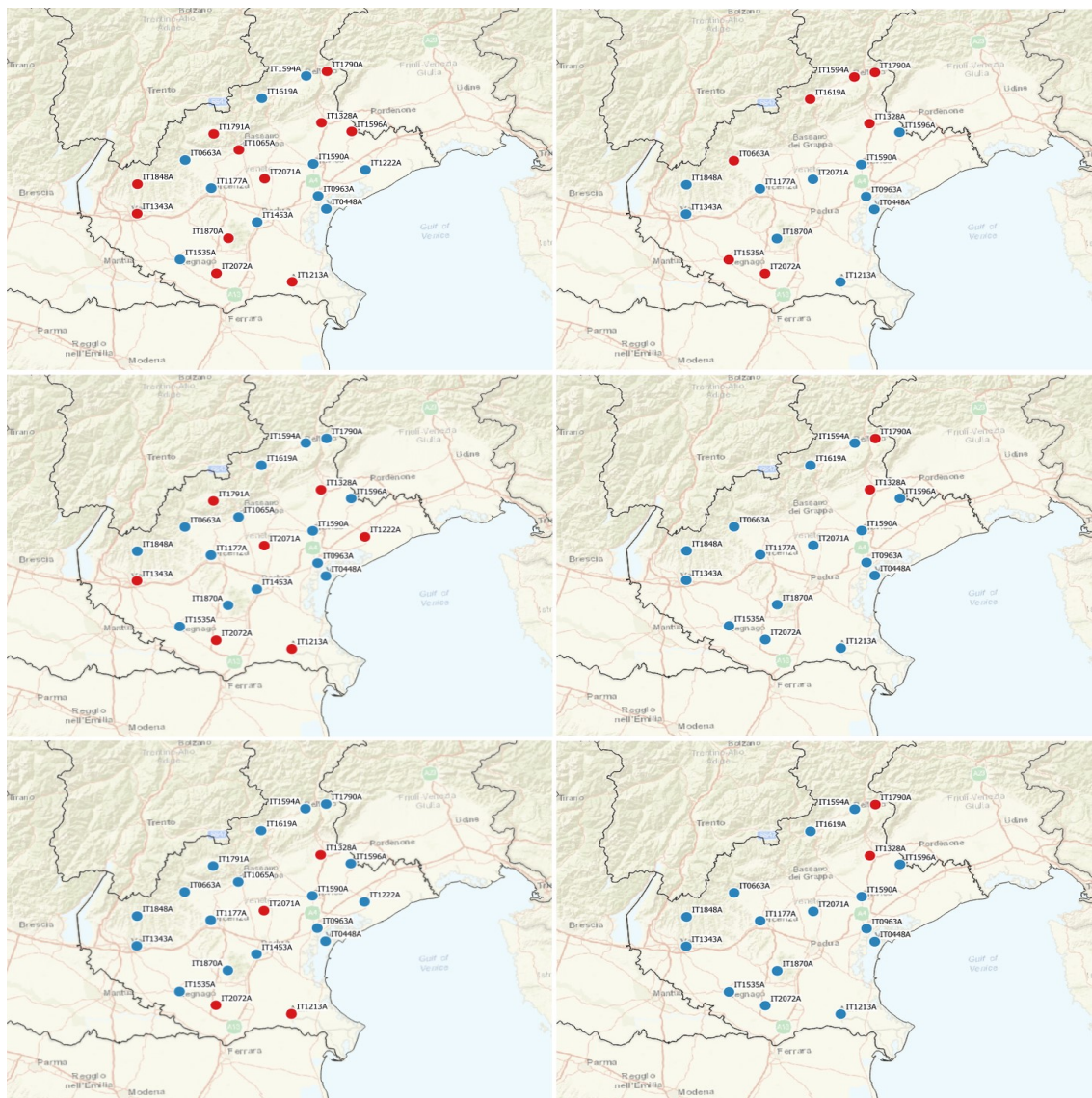
**Figure 5**: "Dot plots" of SI index. For $O_{3,max}$ (left) and PM10 (right). Blue indicates a positive SI (forecast better), red indicates a negative SI (persistence better). Forecast lead time increases from top (D+0) to bottom (D+2).

**Table 4**: Categorical statistics obtained from the "multi-site" elaboration at D+0 forecast time.

| Statistics | $O_{3,max}$ | PM10 |
|:---:|:---:|:---:|
| POD | 61% | 45% |
| FAR | 49% | 34% |
| PC | 76% | 80% |

A global evaluation of model performance was also performed building a "multi-site" time series based on all the prediction-observation pairs. This led to the statistics reported in Table 4. Categorical statistics were evaluated assuming a threshold given by the 75th percentile of observed values.

Finally, a comparison with an alternative low-skill forecasting method was carried out. This is relevant information, since it can be argued that a prediction with no skill is not useful and can be replaced by a trivial prediction.

A commonly used baseline method, in particular for short range forecasts, is persistence: this method consists of forecasting the future values as equal to the last available measurement. Accordingly, predictions at D+0, D+1 and D+2 are set equal to the concentration measured the previous day. A time series of persistence forecasts can be obtained by shifting the measurements by one, two and three days and a performance statistics is computed. This value represents the reference to be compared to the actual forecasts. NME and $NME_{ref}$ can be combined in a Skill Index formulation:

$$SI = 1 - NME/NME_{ref} \qquad (1)$$

A value of 1 is the maximum value for SI and indicates a perfect forecast ($NME = 0$); a value of 0 indicates a model forecast equivalent to the persistence method ($NME = NME_{ref}$), a negative value implies that the model is, in terms of NME, worse than persistence ($NME >$
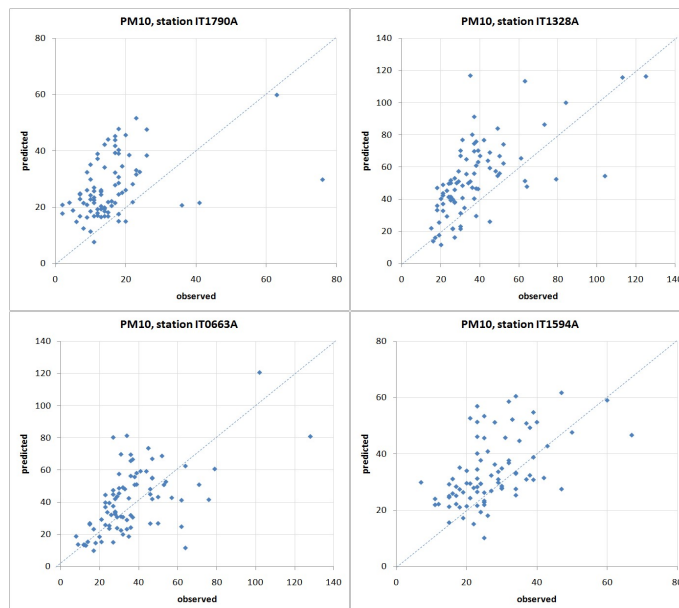
**Figure 6**: Scatter plots of prediction-observation pairs of PM10 from air quality stations: IT1790A and IT1328A (top), IT0663A and IT1594A (bottom).
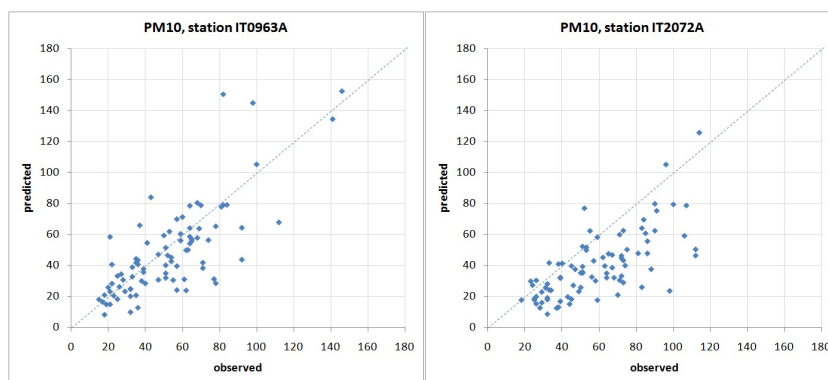


**Figure 7**: Scatter plots of prediction-observation pairs of PM10 from air quality stations: IT0963A (left) and IT2072A (right).

$NME_{ref}$). Fig. 5 shows SI computed for each station and forecast time, both for $O_{3,max}$ and PM10. A red dot means negative SI (persistence is better), a blue dot means positive SI (forecast is better).

## 4 Discussion

As can be seen in Fig. 3, for most stations NMB, NME and *r* meet, as a minimum, the "Criteria" target, both for $O_{3,max}$ and PM10. In many cases, they meet the "Goal". However for PM10, a strong overestimation emerged for stations IT1790A and IT1328A, for which $NMB > 30\%$ and $NME > 50\%$. Looking at the scatter plots of predicted-observed values of these two stations (Fig. 6, top) a clear tendency to overestimation can be noticed in both cases. These stations also show a relevant FAR ($> 60\%$) and low PC ($< 60\%$) (Fig. 4). Moreover, although for most stations the skill was positive, since the persistence method gives worse results, for these two stations, it is negative up to the third day of forecast (Fig. 5). As both are located in complex

terrain, IT1328A at the foot of the hills and IT1790A on an elevated small basin in the initial ranges of the Alps (Fig. 1), this deviation could possibly be traced back to local wind circulations not well described by the meteorological model. In particular, this could be the case of nocturnal katabatic winds, driven by local cooling and occurring in the absence of synoptic scale winds. In fact, the same tendency can also be seen, to a lesser extent, for stations IT0663A and IT1594A (Fig. 6, bottom) which are also close to mountain slopes. This defect could lead to an enhanced stagnation of air masses in those areas, increasing concentrations of particulate matter.

On the other hand, for many stations on the plain, a good performance for PM10 was found (Fig. 7, left), especially regarding correlation that meets the "Goal" target, in most cases. An exception is station IT2072A (Fig. 7, right), in the south, that is marked by a negative NMB ($< -30\%$), low POD and quite a low correlation. Also the neighbouring stations show a negative bias and a low POD. Indeed, although this part of the plain is relatively free from
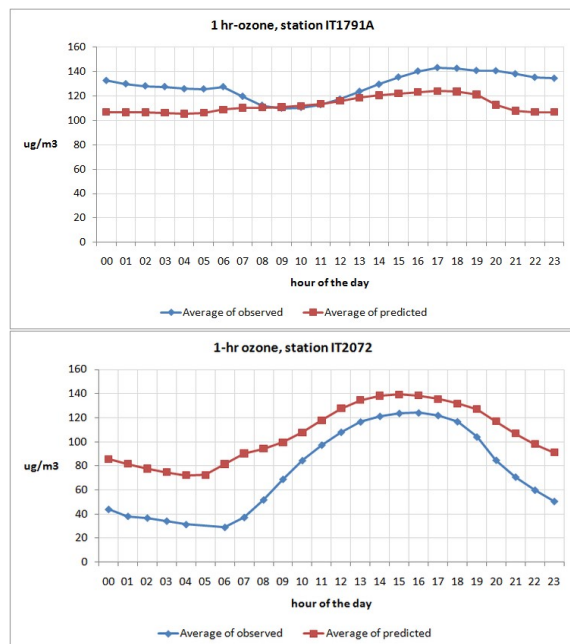
**Figure 8**: Mean daily evolutions of predicted and measured 1hr-ozone. Top: IT1791A station (altitude = 1300 m ), bottom: IT2072A station (altitude= 6 m).

pollutant sources and less inhabited with respect to the north, measured values of PM10 are higher than expected. This is something that is not well understood and that requires further investigation.

As for $O_{3,max}$, NMB, and NME as well, they meet the "Criteria", as a minimum, for most stations and are mostly positive. Correlation meets the "Criteria" target, but never the "Goal", and is poor ($< 0.5$) for three stations located on the mountain ranges (IT1848A, IT1791A and IT1619A). For the IT1791 station, the one at the highest altitude ($> 1,300$ m), a relevant negative NMB resulted ($< -15\%$) and a very low POD ($10\%$) indicating that most exceedances are missed. On the contrary, ozone tends to be overpredicted in some part of the plain, where a relevant FAR ($> 60\%$) was found in three stations (IT2071A, IT2072A and IT1213A). For these sites, skill respect to persistence is also negative up to D+2. To explain this different behaviour, it is worth inspecting the hourly data of observed and predicted ozone. Fig. 8 shows the predicted daily evolution (D+0) of 1-hr ozone for station IT1791A, located on a high plateau in the western pre-alpine ranges. Alongside, for comparison, the same diagram is shown, for station IT2072A, almost at sea level in flat terrain.

Both stations are classified as "rural" sites, that is, relatively far from anthropogenic sources. Ozone is fairly constant at the mountain site whereas, for the station in the plain, it exhibits a pronounced deviation between day and night, with a steady rise from dawn until afternoon and a drop in the evening. In the first case, there is an evident underprediction while in the second case, an overprediction occurs, especially in the night hours. The behaviour of the model could be attributed to different factors: on the mountains, ozone is linked to the transport of air masses

from the free troposphere, so it could be the case that this contribution is not well predicted by the model. On the other hand, ozone in the flat terrain site is mainly determined by the photochemistry of $NO_x$ and CO, emitted in urban areas, and VOCs, so the overprediction could be linked to inaccuracies in the estimate of emissions of such precursors. Inaccuracies in predicted temperature and wind fields can also play an important role; verification studies on the COSMO showed a tendency to overestimate nocturnal temperatures at low altitude sites (Tesini and Cacciamani, 2008); this "warming" can reduce the nocturnal scavenging of ozone by nitrogen monoxide. Simulated deposition processes are also affected by uncertainties; the CAMx implementation of the Zhang model for dry deposition (Zhang et al., 2003) has shown a tendency to generate lower ozone deposition rates respect to other methods, leading to higher ozone predictions.

## 5 Conclusions

This paper describes the model performance evaluation of a system based on the CAMx photochemical model, used to forecast ozone and particulate matter over the region of Veneto, in the north-eastern Italy. The COSMO meteorological prognostic model, coupled with the CALMET diagnostic model in "no observation" mode, provides the main meteorological input. Supplementary inputs describing atmospheric water vapour and aqueous contents are estimated through a specific processor based on available variables. Emissions in the form of grid and point sources derive from regional scale INEMAR inventories and boundary conditions are provided by the CHIMERE model at a European scale. The investigation, consisting of a statistical comparison with measurements, highlights the periods of major concerns for these pollutants; ozone daily 1h-peak ($O_{3,max}$) predictions were evaluated in summer, from June to August 2019, and PM10 (daily mean) in winter, from December 2018 to February 2019. Altogether, the evaluation is satisfactory since the NMB, NME and $r$ metrics selected for model evaluation, for most of the air quality stations, meet, as a minimum, the minimum quality target ("Criteria") established for PM10 and $O_{3,max}$, and in many cases also a stricter "Goal" target. The correlation coefficient of PM10 is the main strength of the forecasting system. However, some local issues emerged. The forecasts of $O_{3,max}$ on elevated terrain showed a poor correlation and a relevant underestimation in the site at higher elevation, whereas a tendency to overestimation was observed for some stations on the plain. Skill respect to the persistence method increases with forecast times although, for a minority of stations, remains negative until the third day of the forecast. Inaccuracies in precursor estimates, in deposition processes and in description of ozone transport through the troposphere could be the reasons for these deviations. For PM10, a strong overestimation was obtained in two sites where skill respect to persistence remains negative up to the third day of forecast. This could be

traced back to limitations in the meteorological model concerning the prediction of local wind circulations, driven by mountain slopes. This leads to an excess of stability of air masses in those areas, increasing concentrations of PM10. Conversely, in some areas in the plain a tendency to underestimation was shown, possibly linked to inaccuracies in emissions estimates. Indeed, part of the deviation can be traced back to the fact that the forecasts are based on the 2015 emissions inventory. In the near future, an update of the emissions inventory is planned to align emissions and replace the INEMAR biogenic module with an implementation of the MEGAN (Model of Emissions of Gases and Aerosols from Nature) model (Guenther et al., 2012) based on current meteorological data and high-resolution land cover inventory.

# References

Arasa, R., Soler, M., Ortega, S., Olid, M., and Merino, M., 2010: *A performance evaluation of MM5/MNEQA/CMAQ air quality modelling system to forecast ozone concentrations in Catalonia*, Tethys, **7**, 11–23, doi:10.3369/tethys.2010.7.02.

Baldauf, M., Seifert, A., Förstner, J., Majewski, D., Raschendorfer, M., and Reinhardt, T., 2011: *Operational convective-scale numerical weather prediction with the COSMO model: Description and sensitivities*, Mon. Weather Rev., **139**, 3887–3905, doi:10.1175/mwr-d-10-05013.1.

Belis, C., Pernigotti, D., Pirovano, G., Favez, O., et al.: *Results of the First European Source Apportionment Intercomparison for Receptor and Chemical Transport Models*, Tech. rep., JRC, iSBN 978-92-79-86573-2, 2018.

Boylan, J. W. and Russell, A. G., 2006: *PM and light extinction model performance metrics, goals, and criteria for three-dimensional air quality models*, Atmos. Environ., **40**, 4946–4959, doi:10.1016/j.atmosenv.2005.09.087.

Byun, D., 1999: *Science algorithms of the EPA Models-3 community multiscale air quality (CMAQ) modeling system*, EPA/600/R-99/030.

Chang, J., Brost, R., Isaksen, I., Madronich, S., Middleton, P., Stockwell, W., and Walcek, C., 1987: *A three-dimensional Eulerian acid deposition model: Physical concepts and formulation*, J. Geophys. Res., **92**, 14 681–14 700, doi:10.1029/jd092id12p14681.

D'Allura, A., Costa, M. P., and Silibello, C., 2018: *QualeAria: European and national scale air quality forecast system performance evaluation*, Int. J. Environ. Pollut., **64**, 110–124, doi:10.1504/ijep.2018.099152.

Denby, B., Douros, I., and Fragkou, L., 2011: *Modelling of Nitrogen Dioxide (NO2) for air quality assessment and planning relevant to the European Air Quality Directive*, in: Forum for Air Quality Modelling in Europe (FAIRMODE) Guidance Document.

Doll, D.: *Guideline for regulatory application of the Urban Airshed Model*, Tech. rep., Environmental Protection Agency, Research Triangle Park, NC, 1991.

EEA: *The application of models under the European Union's Air Quality Directive: A Technical Reference Guide*, Tech. rep., European Environment Agency, doi:10.2800/80600, 2011.

Emery, C., Liu, Z., Russell, A. G., Odman, M. T., Yarwood, G., and Kumar, N., 2017: *Recommendations on statistics and benchmarks to assess photochemical model performance*, J. Air Waste Manage., **67**, 582–598, doi:10.1080/10962247.2016.1265027.

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

Hertel, O., Berkowicz, R., Christensen, J., and Hov, Ø., 1993: *Test of two numerical schemes for use in atmospheric transport-chemistry models*, Atmos. Environ., **27**, 2591–2611, doi:10.1016/0960-1686(93)90032-t.

Hess, M., Koepke, P., and Schult, I., 1998: *Optical properties of aerosols and clouds: The software package OPAC*, Bull. Amer. Meteorol. Soc., **79**, 831–844, doi:10.1175/1520-0477(1998)079¡0831:opoaac¿2.0.co;2.

Honoré, C., Rouil, L., Vautard, R., Beekmann, M., Bessagnet, B., Dufour, A., Elichegaray, C., Flaud, J.-M., Malherbe, L., Meleux, F., et al., 2008: *Predictability of European air quality: Assessment of 3 years of operational forecasts and analyses by the PREV'AIR system*, J. Geophys. Res., **113**, doi:10.1029/2007JD008761.

Jolliffe, I. T. and Stephenson, D. B., 2012: Forecast verification: a practitioner's guide in atmospheric science, John Wiley & Sons.

Karl, M., Guenther, A., Köble, R., Leip, A., and Seufert, G., 2009: *A new European plant-specific emission inventory of biogenic volatile organic compounds for use in atmospheric transport models*, Biogeosciences, **6**, 1059–1087, doi:10.5194/bg-6-1059-2009.

Marshall, J., Langille, R., and Palmer, W. M. K., 1947: *Measurement of rainfall by radar*, Journal of Meteorology, **4**, 186–192, doi:10.1175/1520-0469(1947)004¡0186:MORBR¿2.0.CO;2.

NCAR: *The Tropospheric Visible and Ultraviolet (TUV) Radiation Model*, https://www2.acom.ucar.edu/modeling/tropospheric-ultraviolet-and-visible-tuv-radiation-model, 2011.

Nenes, A., Pandis, S. N., and Pilinis, C., 1998: *ISORROPIA: A new thermodynamic equilibrium model for multiphase multicomponent inorganic aerosols*, Aquat. Geochem., **4**, 123–152.

Pay, M. T., Martínez, F., Guevara, M., and Baldasano, J. M., 2014: *Air quality forecasts on a kilometer-scale grid over complex Spanish terrains*, Geosci. Model Dev., **7**, 1979–1999, doi:10.5194/gmd-7-1979-2014.

Ramboll Environment and Health: *User's guide to the Comprehensive Air Quality Model with Extensions (CAMx)*, 2018.

Scire, J. S., Robe, F. R., Fernau, M. E., and Yamartino, R. J.: *A user's guide for the CALMET Meteorological Model*, Tech. rep., Earth Tech, 2000.

Seinfeld, J. H. and Pandis, S. N., 2016: Atmospheric chemistry and physics: from air pollution to climate change, John Wiley & Sons.

Simon, H., Baker, K. R., and Phillips, S., 2012: *Compilation and interpretation of photochemical model performance statistics published between 2006 and 2012*, Atmos. Environ., **61**, 124–139, doi:10.1016/j.atmosenv.2012.07.012.

Stephens, G., 1978: *Radiation profiles in extended Water Clouds. II: Parameterization Schemes*, J. Atmos. Sci., **35**, doi:10.1175/1520-0469(1978)035¡2123:rpiewc¿2.0.co;2.

Strader, R., Lurmann, F., and Pandis, S. N., 1999: *Evaluation of secondary organic aerosol formation in winter*, Atmos. Environ., **33**, 4849–4863, doi:10.1016/s1352-2310(99)00310-6.

Stull, R., 2000: Meteorology for scientists and engineers, Brooks/Cole.

Tesini, M. and Cacciamani, C.: *High resolution verification of COSMO-I7 2m Temperature over Emilia Romagna region*, Tech. rep., 2008.

Wood, R. and Field, P. R., 2000: *Relationships between total water, condensed water, and cloud fraction in stratiform clouds examined using aircraft data*, J. Atmos. Sci., **57**, 1888–1905, doi:10.1175/1520-0469(2000)057¡1888:rbtwcw¿2.0.co;2.

Yarwood, G., Rao, S., Yocke, M., and Whitten, G., 2005: *Updates to the carbon bond chemical mechanism: CB05*, **8**, 13, final report to the US EPA, RT-0400675.

Zhang, L., Gong, S., Padro, J., and Barrie, L., 2001: *A size-segregated particle dry deposition scheme for an atmospheric aerosol module*, Atmos. Environ., **35**, 549–560, doi:10.1016/s1352-2310(00)00326-5.

Zhang, L., Brook, J., and Vet, R., 2003: *A revised parameterization for gaseous dry deposition in air-quality models*, Atmos. Chem. Phys., **3**, 2067–2082, doi:10.5194/acp-3-2067-2003.

## Appendix A

Technical details are provided below about the principal steps to translate "raw" input data in the proper format and about schemes adopted in the CAMx implementation. Pre-processing of emissions and meteorological fields are depicted. For a full description of CAMx structure and mechanisms please refer to the CAMx User's Guide (Ramboll Environment and Health, 2018).

### Pre-processing of emissions

The INEMAR software includes a series of modules for estimating emissions from industrial and residential sectors, traffic, ports, airports and other off-road mobile machinery, as well as agriculture and livestock. Emissions are grouped according to SNAP97 categories. A "Diffuse emissions" module allows estimating anthropogenic emissions that cannot be localized and a "Punctual emissions" module is used to estimate emissions by industrial stacks of relevant importance. A specific "biogenic emissions" module is used to estimate biogenic VOC emissions: the algorithm estimates the emissions of isoprene, monoterpenes and other VOCs on the basis of the methodology and data proposed by (Karl et al., 2009). The sea salt emissions, expressed in terms of sodium and chloride, are estimated by SEASALT, a processor distributed by Ramboll that generates aerosol emissions of sodium and chloride using meteorological and land use files. Emissions are projected on the model grid, by means of an algorithm based on land use elements in each model grid cell (each land use element is a 250 m x 250 m square). For a given pollutant and emission sector, emission at municipality level is split into contributions according to land use-specific factors. Then the emission

of each land use element is computed as the ratio between that contribution and the number of land use elements. Finally, the emission of a model grid cell is given by the sum of emissions of all land use elements included in the cell. Gridded emissions, and point sources as well, are then elaborated by means of a processor that makes the chemical speciation and disaggregation in terms of $NO_x$ (partitioned as 95% NO and 5% $NO_2$, independently from the emissions sector), VOCs and PM. Speciation tables and time profiles, available as monthly, daily and hourly factors, derive from the experience gained in inter-comparison exercises within the European community (Belis et al., 2018).

### The CAMx model formulation

The CAMx simulates the emission, dispersion, chemical reaction and removal of pollutants by marching the Eulerian continuity equation forward in time for each chemical species. The model time step is dynamically determined during the simulation; time steps typically range from 5-15 minutes for grid cell sizes of 10-50 km, to a minute or less for small cell sizes of 1-2 km. Chemistry is treated by simultaneously solving a set of reaction equations defined by specific chemical mechanisms. The following table summarizes the most relevant chemical processes implemented in the CAMx model. For each process, reference to the physical model and numerical method is given.

### Pre-processing of meteorological fields

Beside wind and temperature and vertical diffusivity, the CAMx model requires 3D water vapour content (WVAP) and pressure (P) as inputs. Moreover, cloud/rain water content and cloud optical depth, even if optional, are actually recommended for photochemistry, aqueous chemistry and deposition calculations. A processor has been devised to estimate these fields based on the following outputs from the COSMO model: wind speed and direction, temperature, geopotential and relative humidity on constant pressure levels (1,000 hPa, 950 hPa, 925 hPa, 850 hPa, 700 hPa, 500 hPa, 400 hPa, 300 hPa), wind speed and direction at 10 m, temperature and dew-point temperature at 2 m, sea-level pressure, total precipitation at surface and cloud cover fraction (low, middle, high). The implemented algorithms are outlined thereafter, equations referred to in the text are listed below. WVAP is computed from temperature T and dew-point temperature $T_d$: vapour pressure $e$ is calculated first as $e = e_s(T_d)$, where the function of temperature $e_s(T)$ is given as Equation 2 (if relative humidity is available, instead of dew point temperature, vapour pressure is given as Equation 3). Water vapour content is then given as Equation 4 in mass units and as Equation 5 in the volumetric units required by the CAMx. Based on a very simplified scheme, low cloud cover (LCC) was assigned to 850 hpa level (about 1,500 m above sea level), middle clouds (MCC) to 700 hpa (about 3,000 m) and high clouds (HCC) to 500 hpa (about

**Table A1**: Summary of the processes implemented in CAMx

| Process | Physical Models | Numerical Methods |
|---|---|---|
| Vertical diffusion | K-theory 1st order closure | Implicit backward-Euler (time) centred (space) solver |
| Gas-phase chemistry | Carbon Bond 2005 (Yarwood et al., 2005) <br> Advanced photolysis model | EBI[a] (Hertel et al., 1993) <br> TUV[b] (NCAR, 2011) |
| Aerosol chemistry | Coarse/Fine (CF) <br> Aqueous inorganic chemistry <br> Inorganic thermodynamic/partitioning <br> Organic thermodynamic/partitioning | Static 2-mode <br> RADM[c] AQ (Chang et al., 1987) <br> ISORROPIA (Nenes et al., 1998) <br> SOAP[d] (Strader et al., 1999) |
| Dry Deposition | Resistance model for gases (Zhang et al., 2003) <br> and aerosols (Zhang et al., 2001) | Deposition velocity as surface boundary condition in vertical diffusion solver |
| Wet deposition | Scavenging model for gases <br> and aerosols (Seinfeld and Pandis, 2016) | Exponential decay as a function of scavenging coefficient |

[a] EBI: Euler-Backward Iterative (solver)

[b] TUV: Tropospheric Ultraviolet and Visible (radiative transfer model)

[c] RADM: Regional Acid Deposition Model

[d] SOAP: Secondary Organic Aerosol Partitioning (model)

**Table A2**: Summary of parameters used to estimate LWC and COD at each level of the COSMOI5 grid. Heights of isobaric levels refer to standard atmosphere.

| Level | Altitude (m) | Cloud fraction (0-1) | $\Delta z^{b}$ (m) | $r_e$ ($\mu m$) |
|---|---|---|---|---|
| Surface level | h[a]+10 | 0.01 | 10 | 5 |
| 1,000 hPa | 100 | 0.01 | 90 | 5 |
| 950 hPa | 500 | 0.01 | 400 | 5 |
| 925 hPa | 750 | 0.01 | 250 | 5 |
| 850 hPa | 1,500 | LCC | 750 | 5 |
| 700 hPa | 3,000 | MCC | 1,500 | 10 |
| 500 hPa | 5,500 | HCC | 2,500 | 50 |
| 400 hPa | 7,000 | 0.01 | 1,500 | 50 |
| 300 hPa | 9,000 | 0.01 | 2,000 | 50 |

[a] h is the terrain height

[b] $\Delta z$ is the thickness of the cloud

5,500 m). This approximation allowed a value of cloud fraction to be associated to each level of the COSMO grid (for levels other than 850, 700 and 500 hPa, a minimum cloud fraction equal to 0.01 was set). Cloud water content was estimated according to an experimental relationship between cloud fraction and condensed water (Wood and Field, 2000): assuming cloud water is entirely in liquid form, liquid water content (LWC) can be computed as Equation 7 as a function of cloud fraction and temperature. Cloud Optical Depth (COD) was estimated according to Stephens (1978): liquid water path (LWP) is calculated first from LWC as Equation 9 in which cloud thickness $\Delta z$ for a level is given, from the bottom to the top, as the difference between the height of the level and the height of the previous one. Then COD can be computed as Equation 10. Integration is performed along the vertical to obtain, for each level, the total COD. Adopted values of $r_e$ are based on the dataset used in the software package OPAC (Hess et al., 1998), in which microphysical properties of some water and ice cloud models were reported. According to these models, $r_e$ is higher for ice clouds so a level-specific value increasing with height was set, from 5 $\mu m$ at 850 hPa to 50 $\mu m$ at 500 hPa, with 10 $\mu m$ at 700 hPa. Table A2 summarizes the factors used to estimate LWC and COD in the COSMO grid.

Rain water content ($M$) was estimated from precipitation rate by means of Equation 11 (Marshall et al., 1947). $M$ was set equal to the surface value for levels up to 1,000 m above the ground, assumed as cloud base, and set to zero above. Pressure at the surface level was calculated from sea-level pressure, by the hypsometric equation (Equation 12), based on temperature and altitude. Finally, all the variables are interpolated on the CAMx grid: first vertical interpolation is performed by means of a linear function and then horizontal interpolation using a simple "nearest neighbour" rule. Fig. A1 shows an example of COD, $M$ ($g/m^3$) and WVAP (ppm), at surface level, resulting from the implementation of the algorithms.

A list of the equations follows below. Equations 2 to 6 and Equation 12 are from (Stull, 2000). Otherwise the reference is given.
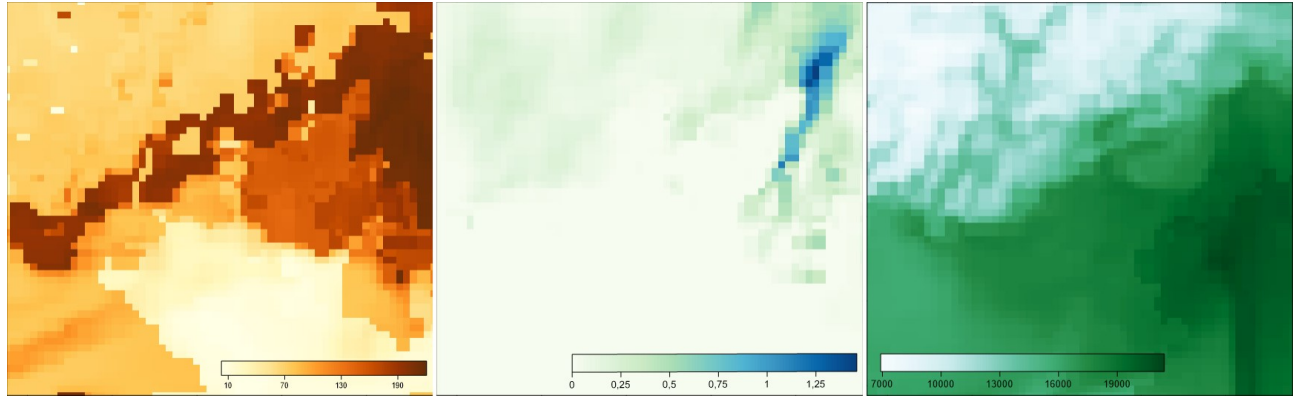
**Figure A1**: Examples of COD, M ($g/m^3$) and WVAP (ppm), from left to right, estimated according to parameterisations described in the text (forecast time +3h, June 5 2020 at 03 UTC).

$$e_s = e_0 e^{b\frac{T-T_1}{T-T_2}} \quad [Pa] \tag{2}$$

Where $e_0 = 611\ Pa$, $b = 17.2694$, $T_1 = 273.16\ K$, $T_2 = 35.86\ K$.

Vapour pressure as a function of relative humidity and temperature

$$e = e_s \frac{RH}{100} \quad [Pa] \tag{3}$$

where $e_s$ is the saturation vapour pressure and RH is the relative humidity.

Water vapour mixing ratio in mass units (WVAP)

$$WVAP = 622 \frac{e}{P-e} \quad [g/kg] \tag{4}$$

where $P$ is the actual pressure and $e$ is vapour pressure.

Water vapour in volumetric units

$$WVAP_{vol} = WVAP \frac{29}{18} 1000 \quad [ppm] \tag{5}$$

where $29\ g/mol$ indicates the molar mass of dry air and $18\ g/mol$, the molar mass of water.

Specific Humidity $q$:

$$q = 622 \frac{e}{P} \quad [g/kg] \tag{6}$$

where $e$ is the vapour pressure and $P$ is the actual pressure.

Cloud water content (LWC)

$$LWC = -ln(1-C) \frac{q_s}{K} \quad [g/kg] \tag{7}$$

where $C$ is the cloud fraction and $K = 75$ is a constant. Obtained by the inversion of equation FWII (Wood and Field, 2000) and assuming condensed water is liquid. Saturation specific humidity qs is given from Equation 7 with $e = e_s(T)$. To avoid undetermined values, a maximum value of 0.99 for $C$ was set.

Volume occupied by 1 $kg$ of dry air (from the ideal gas law)

$$VOL = \frac{RT}{29P} 1000 \quad [m^3] \tag{8}$$

where $R = 8.31\ J/K/mol)$ is the ideal gas constant and $29\ g/mol$ is the molar mass of dry air. It is used to convert units of LWC from $g/kg$ to $g/m^3$.

Liquid water path (LWP)

$$LWP = LWC\Delta z \quad [g/m^2] \tag{9}$$

where $\Delta z$ is the thickness of cloud (Stephens, 1978).

Cloud optical thickness (COD)

$$COD = \frac{3}{2} \frac{LWP}{r_e} \tag{10}$$

where $r_e$ is the effective radius expressed in micrometres and LWP is given by Equation 10 (Stephens, 1978).

Rain water content (M)

$$M = 0.08 RR^{0.83} \quad [g/m^3] \tag{11}$$

where $RR$ is the rain rate in $mm/h$ (Marshall et al., 1947).

Hypsometric equation for the calculation of surface level pressure:

$$P = P_{sea} e^{-\frac{h+10}{A*<T_v>}} \quad [Pa] \tag{12}$$

where $P_{sea}$ is the sea-level pressure and $h$ is the height of terrain. $A = 29.3\ m/K$ is a constant and $<T_v>$ is the layer mean virtual temperature (set equal to temperature $<T>$ and computed assuming a lapse rate of $0.65\ K/100m$).