Abstract: In this paper, an operational forecasting and daily assessment system of air quality is presented. This new system is thought of as a Copernicus-CAMS downstream national service, able to develop and implement a service for air quality forecasting and monitoring in the Italian domain, running every day on the National territory. The system is being developed on behalf of a cooperation between Agenzia Spaziale Italiana (ASI) and Sistema Nazionale Protezione Ambiente (SNPA). SNPA is the network between Istituto Superiore per la Protezione e Ricerca Ambientale (ISPRA) and the Regional Environmental Agencies (ARPAs). The objective of the cooperation is to provide full operation service in terms of continuity, sustainability, and availability of the air quality forecast and evaluation services at the national level. The system forecasts and analyzes air quality throughout Italy, with a focus on Italian regions, for the principal pollutants: Particulate matter with diameter smaller than 10 µm (PM$_{10}$), ozone (O$_3$), and nitrogen dioxide (NO$_2$). It includes a Chemical Transport Model (CTM) nested with the Copernicus Atmosphere Monitoring Service (CAMS) global model and data from the air quality monitoring stations in Italy. The system, under public control and based on open software, is now under testing. To date, it is able to deliver free open data, which is available to environmental agencies and citizens. The data are delivered both as maps and graphs, and as numerical data, useful for providing boundary conditions to local–high resolution-air quality models or for developing customized services. In this work, a downscaling application to a regional nested domain highlights how the new air quality forecasting system gains better results than the Copernicus-CAMS system.

Keywords: air pollution; air pollution forecast modeling; Italy; CAMS; remote sensing

1. Introduction

For a long period, the adverse effects of air pollution on human health and the environment have been demonstrated. In Italy, in the early 1990s, the association between energy consumption (main driver of atmospheric pollution) and emissions into the air was revealed. Consequently, a significant reduction of the latter is sought. As energy consumption increased, emissions into the atmosphere did not increase proportionally. Nonetheless, the high concentrations of air pollutants still occur, and the problem is particularly acute, especially in the cold months during the fall and winter seasons, when air quality issues are most detected [1]. Current health studies demonstrate that pollution plays a significant role in issues related to health due to high urban population densities, for example, at the local scale. Prior to the 20th century, most urban air pollution problems arose from the burning of wood, coal, and other raw materials without any emission controls. Such burning resulted in significant increases
in health issues related to urban pollution [2]. They can be inhaled into the respiratory system and thus cause respiratory lung diseases and even premature death [3].

Facing the challenge of reducing the impact of air pollution on health, the setting up of a prevention system to anticipate pollution events has been, and is, a crucial target. Air quality forecasting is one of the main tools that can help to deal with this objective. Operational systems yielding daily forecasts of air quality over Europe from regional three-dimensional models have been developed in several European countries and are available today [4].

The purpose of the present paper was to present the description and applications of a new Italian national air quality forecasting and monitoring system, recently born and operative, besides the European systems. The system became operational as a result of a close cooperation between Agenzia Spaziale Italiana (ASI) and Sistema Nazionale Protezione Ambiente (SNPA) through the development of a National project. SNPA is a network between Istituto Superiore per la Protezione e Ricerca Ambientale, (ISPRA) and the Regional Environmental Agencies (ARPAs). The National project, named “Piattaforma Tematica del Sentinel Collaborative GS per la Qualità dell’Aria”, involves the definition, design, development, and realization of a component of a thematic platform capable of allowing the delivery of an operational service with daily execution for the assessment, forecast, and analysis of air quality throughout the National Italian domain and to provide boundary to high-resolution air quality models at the regional scale. The new system was based on the implementation of a chemical transport and dispersion model that uses the initial and boundary conditions of the CAMS/Copernicus [5] global model as input data. Hence, the system, under public control and based on open software, is able to deliver free open data, which is available to environmental agencies and citizens.

Due to health problems associated with urban air pollution, many environmental protection agencies have developed capabilities for the continuous monitoring and assessment of air pollution from ground-based stations and for improving sampling techniques. The real-time description of air quality is usually provided by monitoring data. Over the past few decades, air quality monitoring networks have grown up in many regions of Italy. These ground-based measurements are necessary to guide studies on possible ways to reduce the air pollution problems. Yet, ground-based observations represent point measurements and do not have the necessary coverage to characterize the regional distribution of pollutants in the atmosphere.

According to the literature [6], real-time regional monitoring can be improved with spaceborne measurements. The development of satellite remote sensing aerosol products has permitted the exploration of new research techniques for monitoring global air quality [7]. This alternative approach for air quality monitoring provides air quality data where ground-based measurements are not available. Yet, in support of this, models can be used as smart interpolators of the monitoring observations in order to produce near-real time analysis of concentrations using data assimilation techniques. The new system provides pollutants concentrations in near-real time.

In the current study, the Italian air quality forecasting system is presented. Section 2 presents a description of the system, its architecture, input data setup, and outputs. Section 3 is devoted to the presentation of an under-development approach based on estimation of the concentration of particulate matter by remotely sensed data. In Section 4, a system application and a downscaled example at regional scale (Emilia Romagna), an area in the Po Valley, Northern Italy, are presented. It provides a concrete example of how the new Italian system could be used and operatively applied on regional domains by agencies. In fact, the Regional Environmental Emilia Romagna Agency (ARPAE) is now operatively running the new air quality forecasting system and is using it to support daily regional air quality plans. Finally, Section 5 contains concluding remarks.
2. Air Quality Forecasting System

2.1. General Description

This paper aims to present the development and implementation of an operational service for the air quality monitoring and forecasting, running every day over Italian domain. The system was designed to forecast air quality in the domain up to three days ahead, sufficiently in advance about, e.g., pollution episodes in the Italian Peninsula, which could inform the Italian regional agencies of the environment about air quality levels and provide boundary conditions for nested regional domains with higher resolution. The system ingested input data from various origins and used numerical models to produce a forecast of the principal pollutants. Air quality observations were combined with the forecast in order to produce analyses, which were used to evaluate the predicted air quality data, as better explained further in the manuscript. In Figure 1, a schematic idea of the system data flow, from the origin of the continental boundary condition used by the National system until the output from the air quality system, is presented. Further in the text, it is explained how the outputs are distributed and shared with the public. The principal aim was to provide boundary conditions for higher resolution air quality models, which might be used independently by the air quality system described in this manuscript.

Figure 1. In the figure, the forecast suite from the continental CAMS boundary condition is used by the national nested domain to model the distribution of the principal products obtained by the system.
The system, in addition to a visualizer on the web presented in Section 2.5, includes a new and under development verification tool. The tool, called METOPA (Model Evaluation Tool for Observational versus Predicted Air quality data), compares ground-based data with model outputs. METOPA, to date under testing, was written in the programming language R project for Statistical Computing “R”. The methodology applied by METOPA is described in Section 2.6, while some examples of the tool outputs are shown in Sections 4.1 and 4.2 for the regional national domain, respectively.

Another tool, under implementation in the system, is an algorithm for estimating the daily mean particulate matter PM$_{10}$ concentration using satellite AOD retrieval data. The algorithm, presented in Section 3, was applied to obtain pollution maps that allow to check the pollution patterns produced by the Chemical Transport Model (CTM).

2.2. Numerical Air Model: CHIMERE Model Description

CHIMERE is an open-access multi-scale Eulerian chemistry transport model mainly intended to produce hourly forecasts of several aerosol and pollutant gas species concentrations [8,9]. 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 [9]. In the following sections of this paper, an accurate description of the main model setup steps are defined and used in order to build the new daily air quality forecasting system. In the system, the latest version available and downloaded at [10] of CHIMERE (version CHIMERE2017) was used.

The model produces daily full 3D pollutants concentrations in analysis and forecasts. These could be used as boundary conditions for photochemical models at the regional scale, which are used by regional environmental agencies.

2.3. Input Data Setup

2.3.1. Domain Characterization and Geometric Configuration

In the new service for air quality monitoring and forecasting over the Italian domain, it is possible to define three domains of work at the global, continental, and national scale. At the global scale, the pollutant concentrations—approximately 40 km in horizontal resolution and for a limited number of species—are given by the CAMS services. They are processed to be used as boundary conditions for the CTM air quality model at the continental scale. At the continental scale (European domain MEDLn20, with approximately 20 km in horizontal resolution), the model—as input—uses concentrations at the global scale given by the CAMS/Copernicus services, which are post-processed to be used again by the photochemical model CHIMERE. As a result, the model produces pollutant concentrations, which are used as boundary conditions for the model at the national scale. The horizontal resolution of the National domain—ITA7—is approximately 7 km in longitude and latitude, and, as a coordinate system associate, the projection WGS84, long/lat regular grid. CHIMERE simulates the atmosphere below 500 hPa only. Below this altitude, nine levels are taken into account in the model, as described in hybrid sigma-p coordinates. The thickness of the layers increases exponentially with height in the lowest 200 hPa of the atmosphere, and they have a constant thickness in pressure coordinates from 200 hPa to the top of the model. At the national scale (nested in the continental one), the model—as input—uses pollutant concentrations at the continental scale. In Figure 2, the sequence of the three nested domains are mapped. As initial and boundary conditions of the air quality model, the largest domain (dx = 20 km) uses global data (CAMS—where the global concentrations are stored), and the European domain forces the Italian domain (dx = 7 km).
2.3. Input Data Setup

2.3.1. Domain Characterization and Geometric Configuration

In the new service for air quality monitoring and forecasting over the Italian domain, it is possible to define three domains of work at the global, continental, and national scale. At the global scale, the pollutant concentrations—approximately 40 km in horizontal resolution and for a limited number of species—are given by the CAMS services. They are processed to be used as boundary conditions for the CTM air quality model at the continental scale. At the continental scale (European domain MEDLn20, with approximately 20 km in horizontal resolution), the model—as input—uses concentrations at the global scale given by the CAMS/Copernicus services, which are post-processed to be used again by the photochemical model CHIMERE. As a result, the model produces pollutant concentrations, which are used as boundary conditions for the model at the national scale. The horizontal resolution of the National domain—ITA7—is approximately 7 km in longitude and latitude, and, as a coordinate system associate, the projection WGS84, long/lat regular grid.

CHIMERE simulates the atmosphere below 500 hPa only. Below this altitude, nine levels are taken into account in the model, as described in hybrid sigma-p coordinates. The thickness of the layers increases exponentially with height in the lowest 200 hPa of the atmosphere, and they have a constant thickness in pressure coordinates from 200 hPa to the top of the model. At the national scale (nested in the continental one), the model—as input—uses pollutant concentrations at the continental scale. In Figure 2, the sequence of the three nested domains are mapped. As initial and boundary conditions of the air quality model, the largest domain (dx = 20 km) uses global data (CAMS–where the global concentrations are stored), and the European domain forces the Italian domain (dx = 7 km).

Figure 2. In the figure, the sequence of the three nested domains are displayed. The European domain (MEDLn20 with an extension of $-24.81, 27.05, 33.50, 54.99$ as $x_{\text{min}}, y_{\text{min}}, x_{\text{max}}, y_{\text{max}}$, respectively) is nested on the global domain, while the National domain (ITA7 with an extension of $4.36, 35.2, 19.93, 48.88$ as $x_{\text{min}}, y_{\text{min}}, x_{\text{max}}, y_{\text{max}}$, respectively) is nested on the European one.

2.3.2. Meteorological Data

The model COSMO is the official Limited Area atmospheric model, which is used at the Italian National Civil Protection Department (DPCM) and developed and maintained in the framework of the COnsortium for Small-scale MOdelling [11,12]. 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 [10]. It has been designed for both operational numerical weather prediction (NWP) and various scientific applications on the meso-β (2–20 km) and meso-γ (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.
2.3.3. Emissions Data

CHIMERE accounts for five types of emissions: Biogenic emissions, sea-salt emissions, anthropogenic emissions, fire emissions, and dust emissions. The biogenic emissions are calculated using the MEGAN model [13], while the marine aerosol production is based on the scheme proposed by [14]. The anthropogenic emissions are incorporated into the CHIMERE model domain, at the European scale, using the TNO–MACC-III (2011) [15] emissions inventory at the surface. At the Italian scale, the ISPRA inventory of 2015 is considered, while, at the regional scale, the finer and latest available inventory is used. Dust emissions are calculated using the parametrization scheme proposed by [16].

2.3.4. Surface Land Use

The CHIMERE chemistry transport model can use different Land Cover Surface global databases as input during the preprocessing phase: The USGS and the GlobCover—two global land cover maps intensively used all around the world and recommended into forecast modeling procedures, and the Corine land cover (CLC 2012), which is the European land cover. In this study, in particular, a merging of the CORINE 250m 2012–3liv with the European GlobCover 2009 Landuse on the Italian domain—ITA7—was used. The use of the European land use was forced by the need to cover extra European area, not covered by Corine, but physically present in ITA7 domain model used. Finally, as required by the model setup [9], the land use categories were aggregated as the CHIMERE model uses its own land use categories (nine categories), mainly to speed up the calculations. The selection of appropriate parametrization schemes, which characterize the model, was based on previous modeling studies, as detailed in the CHIMERE documentation [9].

2.3.5. Gas Phase Chemistry

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 [9].

2.3.6. Aerosol Chemistry

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 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 $D_{\text{min}} = 0.01 \, \mu m$ to $D_{\text{max}} = 40 \, \mu m$, where the values $2.5 \, \mu m$ and $10 \, \mu m$ are retained as cutoff diameters to allow a meaningful evaluation of PM$_{2.5}$ and PM$_{10}$ in the model, with quantities typically available from routine measurements [9]. The main aerosol processes considered in CHIMERE are coagulation, deposition (dry and wet), the absorption of the semi-volatiles, and nucleation of sulfuric acid [17].

2.4. Ground-Based Measurements

Figure 3 shows the spatial distribution of the monitoring air quality sites throughout Italy. The ground-based measurements sites are spread out over all Italian administrative regions, where they almost uniformly cover most of the Italian Peninsula. In the framework of the new system presented in this paper, the ISPRA-NRT database was used, without a discrimination on the type of the air quality station (e.g., traffic, rural, background) for a total of 688 sites.
2.4. Ground-Based Measurements

Figure 3 shows the spatial distribution of the monitoring air quality sites throughout Italy. The ground-based measurements sites are spread out over all Italian administrative regions, where they almost uniformly cover most of the Italian Peninsula. In the framework of the new system presented in this paper, the ISPRA-NRT database was used, without a discrimination on the type of the air quality station (e.g., traffic, rural, background) for a total of 688 sites.

Figure 3. The figure shows the geographic study domain. The blue colored dots mark locations of the ARPA ground-based stations, distributed on the Italian peninsula.

2.5. The Visualizer on Web

As mentioned in the section above, in the framework of the air quality downstream service, a unique database was created where the near-real-time observed concentration data from the ground-based air quality stations of each Italian administrative region were gathered and stored. In the same logic, a visualizer on the web was created. On the visualizer, the model data map (for the PM\textsubscript{10}, NO\textsubscript{2}, and O\textsubscript{3} atmospheric pollutant—Figure 4a) and the observed data (overlapped—Figure 4b) are displayed. The data were not only visualized, but the visualizer was created with the goal to give the public the opportunity to download the data and use them. In particular, the observed data of the air quality stations at the ground were directly downloaded from the web site of the visualizer as numeric data, archived as a csv and xlsx file, while on the visualizer the model was displayed just as a mean (or maxima–based on the pollutant) daily map. The sharing of the model data happens from the open-source platform CKAN, a Comprehensive Knowledge Archive Network. Here, it is possible to download not only the 2D daily maps—analysis and forecasts of the main pollutant concentrations (i.e., PM\textsubscript{10}, NO\textsubscript{2}, and O\textsubscript{3}), but also full hourly 3D fields in analysis and forecast. Therefore, the aim of this data structure was to create a platform where the model data of the new air quality system are stored daily and are available daily for downloading and use by each Regional Environmental Agencies as boundary conditions of their own air quality models.
Figure 4. The figure shows what is displayed on the visualizer, (a) for example, the daily mean 2D model map of PM10 concentration at the ground level. The colored map has the extension of the Italian model domain ITA7, which is overlapped on the geographical map (Google map). On the right, the referred color palette. (b) As above, with adding overlapped colored dots which represent the observed data measured at the ground station.
2.6. Air Quality Model System Validation Approach

Methodology

As mentioned by the authors of [18], the model output and in-situ measurements are not straightforwardly comparable. The model's outputs release a 3D air quality evaluation of pollutant concentrations on a grid. The concentrations can be thought of as volume averages over each grid cell and time averages over the integration time step. Whereas the in-situ observations are point measurements, they are local data and are influenced by other processes. Nevertheless, a verification approach to evaluate the model performance versus the observations data was built.

To estimate the degree of agreement between the concentrations of the principal pollutant in-situ versus CTM model simulations over each of the air quality stations in the domain under investigation, classical statistical scores were used for the evaluation. They were cumulated over the entire period of analysis considered in this paper: BIAS (in µg/m³), root mean square error (RMSE—in µg/m³), correlation (r-unitless), and mean absolute error (MAE—in µg/m³). Bias indicates, on average, if the predictions are under or overpredicted. In this study, a negative value indicates an underestimation, while a positive one indicates an overestimation. RMSE gives information about the skill in predicting the overall magnitude of the observations. Correlation is a measure of whether predicted and observations change together. The closer the value is to unity, the greater the match predicted concentrations versus observed. The absolute error is the absolute value of the difference between the forecasted value and the actual value. Therefore, MAE tells how big of an error we can expect from the forecast on average. The equations of the statistical parameters, following [19], are given in Appendix A.

3. Remotely Sensed Data from Satellite Sentinel-3

Due to health problems associated with urban air pollution, as mentioned before, many environmental protection agencies have been developed capabilities for the continuous monitoring and assessment of air pollution from ground-based stations and for improving sampling techniques. Yet, ground-based observations represent point measurements and do not have the necessary coverage to characterize the region distribution of aerosols in the atmosphere. The development of satellite remote sensing aerosol products has permitted the exploration of new research techniques for monitoring global air quality [20]. Over the Italian domain, a remotely sensed approach to monitoring air quality has been applied in the past and is available in the literature, especially over Northern Italy [6]. The potential for using space-based sensors for air quality monitoring was demonstrated using Aerosol Optical Depth (AOD) data in combination with the PM ground-based stations, as the literature suggests [20,21]. In this section of the paper, a preliminary under-development approach is presented, with the goal to estimate the daily mean particular matter PM10 concentration in µg/m³ at the ground using satellite AOD retrieval data.

For remotely sensed data, the Sentinel-3 SYN Level-2 product was used [22]. The SYN Level-2 processing aims to combine information from the OLCI and SLSTR instruments to provide improved data for land surface analysis. In particular, the T550 Aerosol Optical Thickness (AOT), at a wavelength of 550 nm and a resolution of 300 m, which is part of the Level-2 SYN product (SY_2_SYN), was included in this product as a measurement data file. The AOD Non Time Critical product was generated by the Level-2 synergy processor, which was the result of synergy between the OLCI OL_1_EFR and SLSTR SL_1_RBT products. Figure 5 shows an example day of AOD retrieval from the SY_2_SYN product on the left, which was compared to the MODIS MOD04_3K product on the right. The AOD granule covered most of the Italian domain. Due to the high pollution levels, especially in the Northern Italy, the use of a product with a high resolution may better achieve details than the standard MODIS product, with, for example, 3 km of spatial resolution. Again, the high spatial resolution product AOD retrieval often provides data where others have reduced data coverage. It is true that one of the fundamental limitations of satellite data is the unavailability of air pollution observations both when clouds obstruct the satellite sensors’ field of view and over domains with high reflectivity surfaces. To bypass this
problem, the presented approach to derive particulate matter at the ground from the satellite AOD retrieval has adopted a gap-filling procedure, named the Data INterpolating Empirical Orthogonal Functions (DINEOF) algorithm, to the granule [23]. Moreover, with a bilinear interpolation on the grid model domain, it is possible to obtain the AOD retrieval data on the domain under investigation.

\[ PM_{10}^{\text{sat}} = A \times \left( \frac{\text{AOD}_{\text{sat}}}{\text{hpbl}} \right) + B \]  

(1)

where the coefficients of the linear relationship (A and B) were obtained on the Po Valley domain and for two years of data training. The \( PM_{10}^{\text{sat}} \) identifies the daily average satellite-based particulate matter \( PM_{10} \), in \( \mu g/m^3 \), at the ground.

The mentioned meteorological variable was obtained from the system presented in this paper, where the height of the boundary layer was available daily and extracted from the model output. In Figure 6, the comparison map (left) of the daily average satellite-based \( PM_{10} \) in \( \mu g/m^3 \) at the ground as result of the applied procedure in Equation (1) is shown and compared with both (in the center) the daily average \( PM_{10} \) in \( \mu g/m^3 \) at the ground as output of the model chain, and the daily average satellite-based \( PM_{10} \) in \( \mu g/m^3 \) at the ground using the MOD04_3K as satellite AOD retrieval, for the same day of analysis. The comparison results highlight quite good agreement in the structure of particulate matter, especially over the Po Valley domain, where higher level concentrations were detected. The \( PM_{10} \) pattern over the Po Valley domain was mainly reproduced by the result obtained using the Sentinel product compared to the MODIS product, mainly due to the high resolution of the Sentinel product.

To date, the procedure has been tested in Northern Italy, considering the EMR3 model domain. The idea for the future is to extend the procedure to a larger domain.

**Figure 5.** In the figure, SY_2_SYN (left) AOD retrieval is shown for one day of October 2019, as an example. The higher-resolution data reveals the spatial variability of AOD, captured by the finer product by Sentinel-3, especially near the urban areas. The MOD04_3km (right) AOD retrieval for the same day is compared as example of a coarser product.
8 October 2019

PM10 (μg/m³) PM10 (μg/m³) PM10 (24 h)

Figure 6. In the figure, (left) the daily average satellite-based PM₁₀ map in μg/m³ is presented at the ground obtained from the satellite-based procedure presented in the paper using both satellite retrieval and meteorological variables from the model chain. This result is compared with (center) the daily average PM₁₀ map in μg/m³ at the ground obtained as aerosol concentration variable from the model chain. (right) The results, obtained from the same satellite-based procedure, is shown using the MODIS MOD04_3K satellite retrieval product.

4. System Application and Downscaling

4.1. An Example of a Regional Model Domain System: The EMR3

A regional domain—EMR3—is presented here as a concrete example of a downscaling application at the regional scale. The EMR3 regional domain (Figure 7) covers an area around the Emilia Romagna administrative Italian region, in the Po Valley, Northern Italy. It has an extension range in longitude from 8.7732 to 13.1355 and in latitude from 43.5043 to 45.8691, with a horizontal resolution of about 3 km and, as associated coordinate system, the projection WGS84/UTM zone 32N. Vertically, the regional domain EMR3 has 15 levels, from the ground until 500 hPa, against the 9 vertical levels of the national domain. By doing this, there is greater consistency between the horizontal and vertical spatial resolution. For this, the outputs from the ITA7 domain were used as boundary conditions. In particular, this regional model domain is operative daily and used by the Regional Environmental Agency, ARPAE (Air Quality monitoring forecast unit), in Bologna. At a local scale, urban pollution plays a significant role in issues related to health due to high urban population densities, mostly over the Po Valley domain, which is covered, for the most part, by the EMR3 model domain. The Po Valley, in particular, in the northern part of Italy, is the area with the most severe air pollution problems in the country and Europe, as it is the largest industrial, trading, and agriculture area with a high population density [24]. In relation to this, the model system output is used daily by the agency and it is proving crucial—in many cases—to provide support (and a major clarification of air quality daily episodes) in forecast mode for the next 48 h (complete simulation for yesterday, today, tomorrow and the day after tomorrow at 10 m). In Figure 8, a daily average map of PM₁₀ mass concentration, as example, is reported, where the colored basemap represents the daily average concentration obtained by the model. Overlapped dots represent the ground-based stations distributed on the Emilia Romagna region, which follow the same color scale range on the right of the figure.

The emission inventory used on the regional domain EMR3 was obtained by merging regional and national inventories. The national inventory was provided by ISPRA, the Italian environmental agency, while the regional inventory was provided by ARPAE on behalf of the Emilia Romagna region and derived from INEMAR, a tool developed in the framework of a collaboration among many Italian regions [25].
Figure 7. In the figure, the regional domain EMR3 is shown, with an extension of 8.7732, 43.5043, 13.1355, 45.8691 as \(x_{\text{min}}, y_{\text{min}}, x_{\text{max}}, y_{\text{max}}\), respectively, in the coordinate system UTM-32.

Figure 8. In the figure, as basemap, the model daily average PM\(_{10}\) in \(\mu g/m^3\) is shown, where the overlapped colored dots represent the geographical location of the ground-based air quality stations of the ARPAE network on the Emilia Romagna region. The color of the dot is associated with the daily mean concentration on the site, with the color range scale on the legend on the bottom left of the map. An overlap of color identifies an agreement between the model and the observation range values.

Example of Model Evaluation Tool on Regional Domain (EMR3)

The verification tool was applied on ground-based measurements and model output concentrations at a first layer for the principal atmospheric pollutants PM\(_{10}\), NO\(_2\), and O\(_3\) for the same period of analysis. METOPA was developed to aid in its evaluation of the air quality models within the operational suite. The tool is used to pair observations and gridded model output in space and time. The evaluation of the model simulations includes the comparison of the model-predicted values against the available observed values.

The observed concentrations were compared to the model first-layer concentrations in analysis and co-located in space during the period of analysis for each monitoring station considered over the regional
domain. The spatial co-location of the model cells with the pollutant ground-based stations was accomplished using the nearest neighborhood approach, considering the model file as a Raster in the R world and extracting the values of a Raster object for the cells in which the set of ground-based locations fall.

For the PM$_{10}$, the evaluation observation versus the model concentration was made on daily mean concentrations; for the NO$_2$, it was made on the daily maxima computed on the hourly concentrations; while, for the O$_3$, the daily maxima were computed on a moving average calculated on the previous 8 h of the day. At the beginning, a threshold on the observed concentrations was applied: Only those air quality monitoring stations for which at least 75% of the daily measurements were available were considered. Due to the fall/winter period under investigation, in this paper, we limited our analysis to October 2019. The PM$_{10}$ has a wide concentration range and variability larger in the fall/winter period than in the summer. New updates will follow in future works, where the verification period will be extended. Table 1 displays the error statistics obtained for the particulate matter, PM$_{10}$, in µg/m$^3$, the daily mean predicted by the new system in analysis during the month of October 2019, and for other two main pollutants, NO$_2$ and O$_3$, calculated, respectively, as the daily maximum and the daily maxima computed on the moving average calculated on the previous 8 h of the day, in µg/m$^3$. The pollutant concentrations were compared with the CAMS ensemble for the same period. The skill scores reported in the table highlight that, with mostly the same correlation value for all three principal pollutants, the bias value for PM$_{10}$ was less than CAMS, so it better underestimated the daily mean concentration. On the other hand, the NO$_2$ result overestimated the daily mean concentration, due to a higher bias value, compared to CAMS. Large bias occurred for the O$_3$ pollutant, where both models overestimated the daily concentration. The “typical weekly” trend for the PM$_{10}$ is reported in Figure 9, while, for NO$_2$ and O$_3$, respectively, the colored trends represent the “typical day”, with a 95% confidence interval in the mean. Even if the BIAS was higher for EMR3 than CAMS, for the NO$_2$ pollutant, the trend was better represented by EMR3, except for a few hours in the night. The same occurred for the O$_3$ pollutant, where, even if there was a good capture of the trend during the daytime, there was an overestimation during the nighttime. This may be due to meteorological issues, which should be analyzed in more detail (i.e., nighttime boundary layer definition).

![Figure 9](image_url)

*Figure 9.* In the figure, for PM$_{10}$, the typical weekly trend is shown, and for NO$_2$ and O$_3$, the typical daily trend is shown. The trends were obtained considering all the available data in µg/m$^3$, observed and estimated from the model for the period of October 2019 for all the ground-based air quality concentration sites over the Emilia Romagna region (EM3).
Table 1. Skill scores for the maximum on moving average 8-h $O_3$, the daily maximum $NO_2$, and daily mean $PM_{10}$, in $\mu g/m^3$, for the period of October 2019, using EMR3 as regional model domain. The skill scores are compared with values obtain from the comparison with the ensemble CAMS air quality model for the same period of analysis.

| Period: October 2019 | PM$_{10}$ | NO$_2$ | O$_3$ |
|---------------------|-----------|--------|-------|
|                     | EMR3 | CAMS | EMR3 | CAMS | EMR3 | CAMS |
| Bias (µg/m$^3$)     | −4.85 | −8.17 | 4.3  | −6.15 | 16.91 | 10.95 |
| RMSE (µg/m$^3$)     | 11.42 | 12.96 | 16.48| 11.95 | 28.01 | 19.33 |
| r                   | 0.72  | 0.76  | 0.51 | 0.56  | 0.63  | 0.78  |
| MAE (µg/m$^3$)      | 8.72  | 9.35  | 10.98| 7.61  | 22.4  | 15.4  |
| FAR                 | 0.44  | 0.43  | 0    | 0     | 0     | 0     |
| POD                 | 0.22  | 0.06  | 0    | 0     | 0     | 0     |

4.2. Example of Model Evaluation Tool on the Italian Domain (ITA7)

METOPA was built to be applied on any domain, with any extension and number of ground-based stations. The tool was also designed to deliver several types of output. An example of the model performance evaluation, for the period of October 2019 on the Italian domain by comparing the time pattern of modeled and measured $PM_{10}$ daily mean, is shown in Figure 10a. The figure shows that model performance depends on the period considered. In Figure 10b (colored trends), the “typical day” is presented, with a 95% confidence interval for $NO_2$ and $O_3$, respectively.

Table 2 displays error statistics obtained for all three pollutants, computed using the same approach as for the regional scale analysis.

A consistent data gathering procedure from the heterogeneous set of ground-based stations, which feature both different hardware and time averaging strategies, is being implemented and will be subject to future investigation. Still, the reported results suggest that a nationwide air quality monitoring system can be achieved with the current methodology.

Table 2. Skill scores for the maximum on moving average 8-h $O_3$, the daily maximum $NO_2$, and daily mean $PM_{10}$, in $\mu g/m^3$, for the period of October 2019, using ITA7 as the National Italian model domain. The skill scores are compared with values obtain from the comparison with the ensemble CAMS air quality model for the same period of analysis.

| Period: October 2019 | PM$_{10}$ | NO$_2$ | O$_3$ |
|---------------------|-----------|--------|-------|
|                     | ITA7 | CAMS | ITA7 | CAMS | ITA7 | CAMS |
| Bias (µg/m$^3$)     | −5.58 | −6.24 | −17.42| −25.97 | 17.60 | 3.43 |
| RMSE (µg/m$^3$)     | 11.90 | 10.71 | 26.97 | 32.81 | 34.55 | 26.30 |
| r                   | 0.55  | 0.69  | 0.40  | 0.49  | 0.33  | 0.48  |
| MAE (µg/m$^3$)      | 8.86  | 7.71  | 20.17 | 26.06 | 24.02 | 13.69 |
| FAR                 | 0.45  | 0.31  | 0     | 0     | 0     | 0     |
| POD                 | 0.12  | 0.12  | 0     | 0     | 0     | 0     |

The model performance evaluation will be assessed regularly by applying the tool on a routine basis during the air quality managing activities. The outputs give information about the confidence level of the model results during the day-by-day operations. They will be used to improve the system.
5. Conclusions

For the past few months, the new Italian air quality forecasting and monitoring system has worked daily. It produces full three-dimensional fields of pollutant concentrations in analysis and forecasts up to three days ahead over the Italian domain, which will be used as boundary conditions by local regional models. The new setup process was based on the three-dimensional deterministic chemistry transport model CHIMERE at the national scale. This paper presents a new service which is now operative in Italy.

Based on this study, the following conclusions were drawn:

1. The presented system is run daily. Results are used by the Regional Italian Environmental Agencies to forecast peak pollution episodes and manage short-term emergency plans and to support long-term air quality improvement plans.
2. The new system delivers everyday hourly boundary conditions to high-resolution air quality models at the regional scale. The model outputs and air quality data from ground monitoring stations which can be easily downloaded and visualized by a web service. They are available for free.
3. Every day, in addition to the gas and PM concentration, the new service provides the meteorological input data for the chemical transport models.
4. An algorithm for estimating the daily mean particulate matter concentration using satellite AOD retrieval data was implemented in the system.

5. A verification tool called METOPA, which works on a routine basis, was included in the system. The tool is used to assess the confidence level of the model outputs and give information for the improvement of the modelling system.

6. Future works will extend the verification analysis to cover at least a complete year. It will be possible to extend the model evaluation approach to the entire national domain with more robust, consistent, and reliable results.

Author Contributions: Data curation, M.S.; Formal analysis, B.A.; Methodology, M.S. and B.A.; Project administration, M.D.; Software, M.S. and B.A.; Supervision, M.D.; Writing—original draft, B.A.; Writing—review & editing, M.S. and B.A. All authors have read and agree to the published version of the manuscript.

Funding: This research was funded by the Project “Piattaforma Tematica del Sentinel Collaborative GS per la Qualità dell’Aria”. The agreement was signed between ASI—Agenzia Spaziale Italiana—and ISPRA—“Istituto Superiore per la Protezione e Ricerca Ambientale”. CUP: F82F17000000005.

Acknowledgments: The authors are thankful for the Regional Agency for Prevention Environment and Energy of Emilia-Romagna, Regional Agency for Environmental Prevention and Protection of Veneto, Regional Agency for Environmental Protection of Piemonte, Regional Agency for Environmental Protection of Lazio, Environmental Protection Agency of Friuli Venezia Giulia, Regional Agency for Environmental Protection of Campania, Regional Agency for Environmental Protection of Umbria, as partners of the project.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Model Statistics

By ref. [19], it is possible to define the following main model statistics used in this manuscript:

- The Root Mean Square Error (RMSE), combines the spread of individual errors and is strongly dominated by the largest values due to the squaring operation,

\[
RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (\text{mod}_i - \text{obs}_i)^2} \tag{A1}
\]

- The BIAS represents the deviation between two datasets;

\[
\text{BIAS} = \frac{1}{n} \sum_{i=1}^{n} (\text{mod}_i - \text{obs}_i) \tag{A2}
\]

The ideal value of this parameter is zero, while negative and positive values indicate respectively underestimation and overestimation.

- The Correlation coefficient (r),

\[
r = \frac{\sum_{i=1}^{n} (\text{obs}_i - \overline{\text{obs}})(\text{mod}_i - \overline{\text{mod}})}{\sqrt{\sum_{i=1}^{n} (\text{obs}_i - \overline{\text{obs}})^2 \times \sum_{i=1}^{n} (\text{mod}_i - \overline{\text{mod}})^2}} \tag{A3}
\]

where \( \overline{\text{obs}} \) and \( \overline{\text{mod}} \) represent the average values of the time series for both observations and model simulations, respectively. The correlation coefficient indicates how efficient is the model to catch up the temporal fluctuations with respect to the observations.

- The Mean Absolute Error (MAE),

\[
\text{MAE} = \left( \frac{1}{n} \sum_{i=1}^{n} |\text{mod}_i - \text{obs}_i| \right) \tag{A4}
\]
is the measure of errors between paired observations and model simulations expressing the same phenomenon.

- The False Alarm Ratio (FAR) represents the number of false alarms over the total number of warnings or alarms
- Probability of Detection (POD) represents the number of detections over the total number of observations.

References
1. Guerreiro, C.; de Leeuw, F.; Viana, M.; Colette, A. Air Quality in Europe: 2018 Report; European Environment Agency: København, Denmark, 2018; ISBN 978-92-9213-990-2.
2. Jacobson, M.Z. Air Pollution and Global Warming: History, Science, and Solutions; Cambridge University Press: Cambridge, UK, 2012.
3. Pope, C.A., III; Burnett, R.T.; Thurston, G.D.; Thun, M.J.; Calle, E.E.; Krewski, D.; Godleski, J.J. Cardiovascular mortality and long-term exposure to particulate air pollution: Epidemiological evidence of general pathophysiological pathways of disease. *Circulation* 2004, 109, 71–77. [CrossRef] [PubMed]
4. Honoré, C.; Rouil, L.; Vautard, R.; Beekmann, M.; Bessagnet, B.; Dufour, A.; Menut, L. Predictability of European air quality: Assessment of 3 years of operational forecasts and analyses by the PREV’AIR system. *J. Geophys. Res. Atmos.* 2008, 113, D04301. [CrossRef]
5. Copernicus Official Web Site. Available online: https://www.copernicus.eu/en/services/atmosphere (accessed on 27 February 2020).
6. Arvani, B.P. Seasonal monitoring and estimation of regional aerosol distribution over Po valley, northern Italy, using a high-resolution MAIAC product. *Atmos. Environ.* 2016, 141, 106–121. [CrossRef]
7. Gupta, P.C. Satellite remote sensing of particulate matter and air quality assessment over global cities. *Atmos. Environ.* 2006, 40, 5880–5892. [CrossRef]
8. Mailler, S.; Menut, L.; Khvorostyanov, D.; Valari, M.; Couvidat, F.; Siour, G.; Colette, A. CHIMERE-2017: From Urban to Hemispheric Chemistry-Transport Modeling. *Geosci. Model Dev.* 2017, 10, 2397–2423. [CrossRef]
9. CHIMERE Documentation. Documentation of the Chemistry-Transport Model CHIMERE. 2017. Available online: https://www.lmd.polytechnique.fr/chimere/ (accessed on 27 February 2020).
10. CHIMERE A Multi-Scale Chemistry-Transport Model for Atmospheric Composition Analysis and Forecast. Available online: https://www.lmd.polytechnique.fr/chimere/ (accessed on 27 February 2020).
11. COSMO. Source Consortium for Small-scale Modeling. 2020. Available online: http://www.cosmo-model.org (accessed on 27 February 2020).
12. Doms, G.; Baldauf, M. A Description of the Non-hydrostatic Regional COSMO-Model. Part I: Dynamics and Numerics. 2015. Available online: User guide documentation at www.cosmo-model.org (accessed on 27 February 2020).
13. Guenther, A.K. Estimates of global terrestrial isoprene emissions using MEGAN (Model of Emissions of Gases and Aerosols from Nature). *Atmos. Chem. Phys.* 2006, 6, 3181–3210. [CrossRef]
14. Monahan, E.C. The Ocean as a Source of Atmospheric Particles. In *The Role of Air-Sea Exchange in Geochemical Cycling*; Kluwer Academic Publishers: Dordrecht, Holland, 1986; pp. 129–163.
15. MACC-III Monitoring Atmospheric Composition and Climate 3. Available online: https://cordis.europa.eu/project/id/633080 (accessed on 27 February 2020).
16. Kok, J.F.; Mahowald, N.M.; Fratini, G.; Gillies, J.A.; Ishizuka, M.; Leys, J.F.; Zobeck, T.M. An improved dust emission model-Part 1: Model description and comparison against measurements. *Atmos. Chem. Phys.* 2014, 14, 13023–13041. [CrossRef]
17. Bessagnet, B.; Hodzic, A.; Vautard, R.; Beekmann, M.; Cheinet, S.; Honoré, C.; Lioussse, C.; Rouil, L. Aerosol modeling with CHIMERE: Preliminary evaluation at the continental scale. *Atmos. Environ.* 2004, 38, 2803–2817. [CrossRef]
18. Riccio, A.; Barone, G.; Chianese, E.; Giunta, G. A hierarchical Bayesian approach to the spatio-temporal modeling of air quality data. *Atmos. Environ.* 2006, 40, 554–566. [CrossRef]
19. Wilks, D.S. *Statistical Methods in the Atmospheric Sciences*; Academic Press: Cambridge, MA, USA, 2011.
20. Gupta, P.; Christopher, S.A. Seven Year Particulate Matter Air Quality Assessment from Surface and Satellite Measurements. *Atmos. Chem. Phys. Discuss.* 2008, 8, 327–365. [CrossRef]
21. Emili, E.; Lyapustin, A.; Wang, Y.; Popp, C.; Korkin, S.; Zebisch, M.; Wunderle, S.; Petitta, M. High spatial resolution aerosol retrieval with MAIAC: Application to mountain regions. *J. Geophys. Res. Atmos.* 2011, 116, D23211. [CrossRef]

22. Donlon, C.; Berruti, B.; Buongiorno, A.; Ferreira, M.H.; Féménias, P.; Frerick, J.; Nieke, J. The global monitoring for environment and security (GMES) sentinel-3 mission. *Remote Sens. Environ.* 2012, 120, 37–57. [CrossRef]

23. Ping, B.; Su, F.; Meng, Y. An improved DINEOF algorithm for filling missing values in spatio-temporal sea surface temperature data. *PLoS ONE* 2016, 11, e0155928. [CrossRef] [PubMed]

24. Bigi, A.; Ghermandi, G. Long-term trend and variability of atmospheric PM10 concentration in the Po Valley. *Atmos. Chem. Phys.* 2014, 14, 4895–4907. [CrossRef]

25. INEMAR Official Site. Available online: http://www.inemar.eu (accessed on 27 February 2020).