Toward an Early Warning System for Health Issues Related to Particulate Matter Exposure in Brazil: The Feasibility of Using Global PM 2.5 Concentration Forecast Products


To cite this version:

HAL Id: hal-03060437
https://hal.univ-reunion.fr/hal-03060437v2
Submitted on 13 Jan 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution| 4.0 International License
Article

Toward an Early Warning System for Health Issues Related to Particulate Matter Exposure in Brazil: The Feasibility of Using Global PM$_{2.5}$ Concentration Forecast Products

Emmanuel Roux $^{1,2,3,*}$, Eliane Ignotti $^4$, Nelson Bègue $^5$, Hassan Bencherif $^5$, Thibault Catry $^{1,2}$, Nadine Dessay $^{1,2}$, Renata Gracie $^{2,3}$, Helen Gurgel $^{2,6}$, Sandra de Sousa Hacon $^7$, Mônica de A. F. M. Magalhães $^{2,3}$, Antônio Miguel Vieira Monteiro $^8$, Christophe Revillion $^1$, Daniel Antunes Maciel Villela $^9$, Diego Xavier $^{2,3}$ and Christovam Barcellos $^{2,3}$

$^1$ ESPACE-DEV (Institut de Recherche pour le Développement, Université de la Réunion, Université des Antilles, Université de Guyane, Université de Montpellier), 97490 Saint-Denis de La Réunion, France; thibault.catry@ird.fr (T.C.); nadine.dessay@ird.fr (N.D.); christophe.revillion@univ-reunion.fr (C.R.)

$^2$ LMI Sentinel, International Joint Laboratory “Sentinela” (Fiocruz, UnB, IRD), Rio de Janeiro RJ-21040-900, Brazil; renata.gracie@icict.fiocruz.br (R.G.); helengurgel@unb.br (H.G.); monica.magalhaes@icict.fiocruz.br (M.d.A.F.M.M.); diego.ricardo@icict.fiocruz.br (D.X.); christovam.barcellos@fiocruz.br (C.B.)

$^3$ LIS, Laboratório de Informação em Saúde, Instituto de Comunicação e Informação Científica e Tecnológica em Saúde (ICICT), Fundação Oswaldo Cruz (Fiocruz), Climate and Health Observatory, Rio de Janeiro RJ-21040-900, Brazil

$^4$ UNEMAT, Universidade do Estado de Mato Grosso, Programa de Pós-Graduação em Ciências Ambientais, Campus de Cáceres, Cáceres MT-78217-900, Brazil; eliane.ignotti@unemat.br

$^5$ LACy, Laboratoire de l’Atmosphère et des Cyclones (UMR 8105 CNRS, Université de La Réunion, Météo-France), 97490 Saint-Denis de La Réunion, France; nelson.begue@univ-reunion.fr (N.B.); Hassan.Bencherif@univ-reunion.fr (H.B.)

$^6$ LAGAS, Laboratório de Geografia, Ambiente e Saúde, Universidade de Brasilia (UnB), Departamento de Geografia, Campus Universitário Darcy Ribeiro, Brasilia DF-70910-900, Brazil

$^7$ ENSP, Escola Nacional de Saúde Pública, Departamento de Endemias Samuel Pessoa, Fiocruz, Rio de Janeiro RJ-21041-210, Brazil; shacon@ensp.fiocruz.br

$^8$ LiSS, Laboratório de investigação em Sistemas Socioambientais, Instituto Nacional de Pesquisas Espaciais (INPE), São José dos Campos SP-12227-010, Brazil; miguel.monteiro@inpe.br

$^9$ PROCC, Programa de Computação Científica, Fiocruz, Rio de Janeiro RJ-21040-900, Brazil; daniel.villela@fiocruz.br

* Correspondence: emmanuel.roux@ird.fr

Received: 16 October 2020; Accepted: 8 December 2020; Published: 12 December 2020

Abstract: PM$_{2.5}$ severely affects human health. Remotely sensed (RS) data can be used to estimate PM$_{2.5}$ concentrations and population exposure, and therefore to explain acute respiratory disorders. However, available global PM$_{2.5}$ concentration forecast products derived from models assimilating RS data have not yet been exploited to generate early alerts for respiratory problems in Brazil. We investigated the feasibility of building such an early warning system. For this, PM$_{2.5}$ concentrations on a 4-day horizon forecast were provided by the Copernicus Atmosphere Monitoring Service (CAMS) and compared with the number of severe acute respiratory disease (SARD) cases. Confounding effects of the meteorological conditions were considered by selecting the best linear regression models in terms of Akaike Information Criterion (AIC), with meteorological features and their two-way interactions as explanatory variables and PM$_{2.5}$ concentrations and SARD cases, taken separately, as response variables. Pearson and Spearman correlation coefficients were then computed between the residuals of the models for PM$_{2.5}$ concentration.

and SARD cases. The results show a clear tendency to positive correlations between PM$_{2.5}$ and SARD in all regions of Brazil but the South one, with Spearman’s correlation coefficient reaching 0.52 ($p < 0.01$).

Positive significant correlations were also found in the South region by previously correcting the effects of viral infections on the SARD case dynamics. The possibility of using CAMS global PM$_{2.5}$ concentration forecast products to build an early warning system for pollution-related effects on human health in Brazil was therefore established. Further investigations should be performed to determine alert threshold(s) and possibly build combined risk indicators involving other risk factors for human respiratory diseases. This is of particular interest in Brazil, where the COVID-19 pandemic and biomass burning are occurring concomitantly, to help minimize the effects of PM emissions and implement mitigation actions within populations.

**Keywords:** particulate matter forecasts; severe acute respiratory diseases; Brazil; early warning system; remotely sensed observation assimilation

---

1. Introduction

Solid and liquid particles in the air come from a variety of sources: transport, residential heating, agriculture, biomass burning (including forest fires), etc. Both long- and short-term exposures to particulate matter (PM) of diameter less than 10 µm (PM$_{10}$) and 2.5 µm (PM$_{2.5}$) affect human health.

Based on the converging results of a large body of epidemiological and laboratory studies, PM (PM$_{10}$ and PM$_{2.5}$) from outdoor air pollution was classified by the International Agency for Research on Cancer (IARC) as carcinogenic to humans (IARC Group 1) [1]. Among the PM, PM$_{2.5}$ is considered to be the most dangerous due to its capacity to penetrate deeper into the pulmonary alveoli and the blood. In 2017, long-term exposure to ambient PM$_{2.5}$ was estimated to contribute to 2.9 million deaths (deaths that likely occurred earlier than would be expected in the absence of PM$_{2.5}$ pollution) and to a loss of 83 million disability-adjusted life-years (DALYs, defined as the sum of the years of life lost from early deaths plus the years lived with a disability) worldwide, making PM$_{2.5}$ responsible for 5.2% of all deaths and 3.3% of all DALYs in the world [2]. For this estimation, the study considered ischemic heart disease, cerebrovascular disease (ischemic stroke and hemorrhagic stroke), lung cancer, chronic obstructive pulmonary disease (COPD), and lower-respiratory infections (in particular, pneumonia) and included type 2 diabetes for the first time [3].

In Brazil, 99.8 million people live in metropolitan regions (Regiões Metropolitanas), Integrated economic development regions (Regiões integradas de desenvolvimento econômico), or urban agglomerations with more than 1 million people [4]. This population is potentially exposed to urban pollutants emitted notably by vehicles—vehicular emission is the main contributor to urban air pollution [5]—and industry, and to emissions due to biomass burning, which can be transported to the urban areas [5,6]. Andreão et al. [5] studied four metropolitan areas (MAs) of the Southeast region of Brazil (São Paulo, Rio de Janeiro, Belo Horizonte, and Vitória), corresponding to 102 cities. They estimated the daily PM concentration levels and the number of deaths that would be avoided by satisfying the World Health Organization (WHO) Air Quality Guidelines (AQGs) [7]. Among these 102 cities, 67 presented mean annual PM$_{2.5}$ concentrations above the WHO guideline (10 µg m$^{-3}$).

The total number of all-cause avoidable premature deaths was estimated as 32,000 ± 5300 for the four MAs, representing 6.5% of all-cause deaths. Regarding the consequences of short-term exposure to PM$_{2.5}$, 3436 ± 2340 hospitalizations of children from one to nine years were attributable to daily mean PM$_{2.5}$ concentrations exceeding the WHO AQG (25 µg m$^{-3}$ daily mean). For circulatory system diseases in the elderly, this number reached 9850 ± 3950 for the four MAs.

More than 10 million people are directly exposed to high levels of pollutants resulting from deforestation and agricultural fires over the Brazilian Amazon [8]. During the biomass burning season, high concentrations of PM$_{10}$ (PM$_{10}$ including PM$_{2.5}$) have been measured (ranging from 400 to
600 μg m$^{-3}$) [9], exceeding the WHO AQG by 8 to 12 times [7,8]. Alves et al. [8] demonstrated that exposure to such emissions causes toxic effects at the molecular and cellular levels in human lung cells, especially through the transportation of retene. However, exposure to PM produced by fires in the Amazon also induces acute health disorders: a 10 μg m$^{-3}$ increase in the exposure levels to PM$_{2.5}$ was associated to 2.9% and 2.6% increases in the number of ambulatory receptions of children for respiratory diseases on the sixth and seventh days following the exposure, respectively [10].

In the municipality of Alta Floresta, Mato Grosso State, Brazil, the effects of the short-term exposure to PM$_{2.5}$ on the lung function of 309 school children between 6 to 15 years were studied during a 4-month period [11]. The mean 24-h PM$_{2.5}$ concentration ranged from 6.39 to 99.91 μg m$^{-3}$ (overall mean: 24.35 μg m$^{-3}$). A statistically significant reduction in peak expiratory flow (PEF) measures, ranging from 0.26 L min$^{-1}$ (95% CI: 0.49, 0.04) to 0.38 L min$^{-1}$ (95% CI: 0.71, 0.04), was observed with an increase in the PM$_{2.5}$ concentration of 10 μg m$^{-3}$.

In May 2020, the Brazilian space agency (Instituto Nacional de Pesquisa Espacial (INPE)) warned of a forest fire season in 2020 in Brazil (which occurs between August and November, with interannual variations between June and December and with a peak in September [12]) that could overcome that of 2019 in terms of burned surface due to high deforestation rates, dry climate conditions, and the weakening of the control of illegal activities, especially by the Brazilian institute of the environment and renewable natural resources (Instituto Brasileiro do meio ambiente e dos recursos naturais renováveis, IBAMA) [13]. In such context, several Brazilian scientific institutions alerted to the impact of the concomitant biomass burning and the COVID-19 epidemic on population health and health service overcrowding [13,14]. Such a potential disastrous situation is also reported in the United States [15].

Particular matter acts on viral infections in different ways: it is associated with the aggravation or acute events of existing pathologies, especially in children and the elderly; it increases patient susceptibility and immune response to virus infection; and it may enhance exposure, acting as a virus carrier [16,17]. Specific relationships between the exposure to atmospheric PM and COVID-19 have already been noted [18–22]. Andrée et al. [21] considered various confounding factors: population density, health preconditions, and spatial distribution of cases to account for the spread of the disease. They concluded that PM$_{2.5}$ is a strong predictor of the number of confirmed COVID-19 cases in the Netherlands, and that an increase in PM$_{2.5}$ concentration from 10 to 12 μg m$^{-3}$ resulted in a 100% increase in the COVID-19 case number during the study period. By adjusting for 20 county-level covariates in the United States, Wu et al. [22] established that an increase of 1 μg m$^{-3}$ in the 17-year long-term average PM$_{2.5}$ was significantly associated with a 11% (95% CI, 6 to 17%) increase in the COVID-19 mortality.

The above-mentioned effects of PM on health, as well as the current situation with the COVID-19 pandemic and biomass burning in Brazil, necessitate the creation of an early alert system for the Brazilian population about exposure to PM, to implement urgent actions to reduce emissions, and to mitigate the effects of pollution on health.

The WHO defined AQGs for PM$_{2.5}$ and PM$_{10}$ concentrations that “reflect the concentrations at which increased mortality responses due to PM air pollution are expected based on current scientific findings” [7]: daily average exposure of 25 and 50 μg m$^{-3}$ and annual mean exposure of 10 and 20 μg m$^{-3}$ for PM$_{2.5}$ and PM$_{10}$, respectively. The WHO AQG for long-term PM$_{2.5}$ concentration exposure (10 μg m$^{-3}$) was chosen here as it corresponds to the lower end of the range over which significant effects on mortality due to cardiopulmonary diseases and lung cancer (considering 95% confidence interval) were observed, making use of American Cancer Society’s data [7,23]. The AQG for daily PM$_{2.5}$ concentration exposure (25 μg m$^{-3}$) is related to the annual mean value of 10 μg m$^{-3}$, considering the observed relationship between the 99th percentile of the distribution of the daily mean levels and the annual mean [7]. Eventually, AQGs for PM$_{10}$ were defined by considering a PM$_{2.5}$/PM$_{10}$ ratio of 0.5, reflecting both the typical and lower values observed for urban areas in developing and developed countries, respectively [7].
However, according to the WHO, given that no evidence exists of a safe level of exposure to PM\textsuperscript{2.5} and that significant interindividual variability can be observed in responses to PM exposure, “the standard-setting process needs to aim at achieving the lowest concentrations possible in the context of local constraints, capabilities and public health priorities” \cite{7}. Thus, countries define their own standards, sometimes supposed to evolve toward increasingly stringent regulations, according to the concentration level observed in their territory and their measurement and control capacities. In Brazil, resolution CONAMA 491/2018, which states the air quality standards, was approved by the National Environmental Council (Conselho Nacional de Meio Ambiente, CONAMA) in 2018. The resolution plans to adopt the WHO air quality reference values in four stages. For PM\textsubscript{2.5}, the intermediate reference values for the daily average concentration will thus be 60 (currently in application), 50, and 37 \(\mu g m^{-3}\), before reaching the final reference value of 25 \(\mu g m^{-3}\). For the annual average concentration, these values are 20, 17, 15, and 10 \(\mu g m^{-3}\) \cite{25}. However, no timetable has been defined for the adoption of the next reference values \cite{25,26}. In addition, countries can define alert thresholds. They relate to high concentration levels over short periods of time and are normally associated with actions to reduce the main emissions to protect the health of the population (restriction of the movement of certain vehicles and/or of certain industrial activities, etc.). In Brazil, levels of attention, alert, and emergency are defined as follows for PM\textsubscript{2.5}: 125, 210, and 250 \(\mu g m^{-3}\), respectively. The technical evaluation of the CONAMA 491/2018 resolution by the Brazilian Health and Sustainability Institute (Instituto Saúde e Sustentabilidade, ISS) reported that these values are very high (among the highest worldwide) and will rarely be reached to force immediate measures to protect the population \cite{25}.

Monitoring PM in Brazil would require a sufficiently dense and distributed network of PM measurements over the whole territory. However, in 2017, “only 24 Brazilian cities (0.43\% of all cities) monitored fine particles with 50 monitoring stations, all located in the southeastern region” \cite{5}. An inventory of the capabilities of air quality monitoring performed in 2018 \cite{25} showed that air pollution was monitored in only 7 of 27 Brazilian states (26\%). Among the 319 measurement stations in operation in the country, 298 (93.4\%) were present in the Southeast region, 13 (4.1\%) in the South, and 4 (1.3\%) in the Center West and Northeast regions \cite{25}. Only 186 (58.3\%) and 65 (20.4\%) stations were able to measure PM\textsubscript{10} and PM\textsubscript{2.5} concentrations, respectively. No station appeared to be present in the North region (Amazon), which is the region among those most affected by forest fires over the last few years \cite{27}.

Satellite-derived PM\textsubscript{2.5} concentration and composition estimates were proven to be related to respiratory disorders (children’s lung function) \cite{28}. This demonstrates the possibility of overcoming the limits due to the discontinuity of the information provided by ground stations, both in space and time, using remotely sensed products. However, in South America, this data source “is still not much employed in scientific research, and for air quality decision-makers, probably because of access, processing, validation, and interpreting difficulties” \cite{29}. Andreão et al. \cite{5} estimated daily PM concentrations at a 5 km spatial resolution using the weather research and forecasting model coupled with chemistry (WRF-Chem) and remotely sensed data including land use/cover and fire detection. Wu et al. \cite{22} estimated PM\textsubscript{2.5} concentrations across the United States on a \(0.01 \times 0.01\)-degree grid, using a model taking into account remotely sensed aerosol optical depth (AOD), model-based information, and ground-based observations. However, to the best of our knowledge, already available and easily accessible global PM concentration forecast products have not been associated with respiratory disorders to develop an early warning system. PM concentration forecasts provided on regional or global scales by models are of particular interest. Several products already exist, like those provided by the United States National Aeronautics and Space Administration (NASA) Global Monitoring and Assimilation Service (GMAO), the Copernicus Atmosphere Monitoring Service (CAMS) in Europe, or the INPE in Brazil.

In this context, we aimed to evaluate the feasibility of using available global PM\textsubscript{2.5} concentration forecast data to construct an early warning system for PM\textsubscript{2.5} exposure and its health consequences in Brazil.
2. Materials and Methods

2.1. Data

2.1.1. Cases of Severe Acute Respiratory Diseases (SARD)

Retrospective data on hospitalizations due to acute respiratory disorders were provided by the InfoGripe surveillance system developed by the Group of Analytical Methods in Epidemiological Surveillance (MAVE, Scientific computation program of the Oswaldo Cruz Foundation, PROCC/Fiocruz, School of Applied Mathematics of the Getulio Vargas Foundation, EMAp/FGV) and the Influenza Working Group of the Health Surveillance Secretariat of the Ministry of Health (GT-Influenza, SVS, MS) [30] (see Appendix A for all data sources and main features). The InfoGripe system aims to monitor and generate alerts regarding hospitalizations using the Brazilian Notifiable Diseases Information System (Sistema de Informação de Agravos de Notificação, SINAN). It strictly includes severe acute respiratory syndromes (SARS), i.e., viral respiratory diseases caused by coronaviruses, as well as other viral infection types such as seasonal influenza. It also gathers cases satisfying a less restrictive definition: all hospitalization or death cases due to respiratory disorders, without necessarily including the presence of fever or other symptoms related to a viral infection. This broadest definition of the cases in the InfoGripe system was considered in this study and is referred to as severe acute respiratory diseases (SARD) hereafter.

The InfoGripe system provides data by epidemiological week, Brazilian state, age groups, and sex. The total number of positive and negative tests for viral infections are also provided, as well as test results for the main viruses (influenza A, influenza B, SARS-CoV-2, respiratory syncytial virus (RSV), parainfluenza 1, parainfluenza 2, parainfluenza 3, and adenovirus).

Data for the 2015–2019 period were downloaded from the InfoGripe online repository [31]. As the COVID-19 pandemic is considerably affecting the Brazilian health system as a whole and SARD surveillance in particular [14], cases for 2020 were excluded from the study to avoid biases in the analysis.

2.1.2. PM$_{2.5}$ Concentrations

Data on PM$_{2.5}$ concentrations were all provided by the Copernicus Atmosphere Monitoring Service (CAMS). CAMS forecasting is based on the Integrated Forecasting System (IFS) developed by the European Centre for Medium-Range Weather Forecasts (ECMWF). Predictions of atmosphere composition (including PM$_{2.5}$) are derived from an ensemble modeling approach, which assimilates satellite observations. In particular, the Moderate Resolution Imaging Radiospectrometer (MODIS) instrument aboard the Terra and Aqua satellites (NASA) and the Polar Multi-sensor Aerosol product (PMAp) provided by the MetOp satellites (European Organisation for the Exploitation of Meteorological Satellites, EUMETSAT, and European Space Agency (ESA)) are used to estimate the AOD (https://atmosphere.copernicus.eu/satellite-observations) [32]. MODIS active fire products are also assimilated in the IFS via the Global Fire Assimilation System (GFAS), which estimates emissions from biomass burning [32]. In this study, three different products were used: CAMS reanalysis data, archived CAMS near real-time (NRT) forecast data, and current CAMS NRT forecast data. As detailed below, each product was associated with a specific question.

The CAMS reanalysis data, provided on an $\sim$80 km spatial resolution regular grid, were used to describe and discuss the surface PM$_{2.5}$ concentration distributions from 2015 to 2019 in Brazil. The CAMS reanalysis system retrospectively applies the last improvements in the models and observations to provide reference and homogeneous data on atmosphere composition over a long past period [33], whereas NRT forecasts are impacted at each IFS upgrade implementation (Figure 1). Reanalysis data were accessed using the Copernicus Atmosphere Data Store.

For the retrospective analysis intended to demonstrate the feasibility of an early warning system, by comparing PM$_{2.5}$ concentration forecasts with SARD cases, the archived surface forecasts of PM$_{2.5}$ concentrations were considered (archived CAMS NRT). For each day, the CAMS NRT dataset consists
of a 0.4 degree (∼40 km) spatial resolution forecast on a regular grid every 3 h from 00:00 UTC on the current day to 24:00 UTC on the fifth day after. Time steps are coded 0, 3, 6, ..., 120 in the data repository. To study the feasibility of an early warning system development, the maximum prediction horizon of the daily mean PM concentrations in Brazil was considered. In other words, the PM$_{2.5}$ concentration values considered for day $d$ were those provided by CAMS NRT on the fourth previous day ($d - 4$) at the horizon $d + 4$. In practice, the average of the forecasts from time steps 75 to 96 of the prediction performed four days before was computed (these time steps correspond to the 00:00 UTC–3 to 24:00 UTC–3 interval) to obtain the daily mean PM$_{2.5}$ concentrations in accordance with the main Brazilian time zone (UTC–3). Daily values were then averaged by epidemiological week for comparison with the weekly SARD data.

The successive IFS upgrades significantly impacted the CAMS NRT forecast results and had to be considered to define the study period. This was particularly visible after the implementation of the 43R1 and 46R1 IFS versions on 24 January (4th epidemiological week), 2017, and on 9 July (28th epidemiological week), 2019, respectively (Figure 1). These two dates correspond to significant drops in PM$_{2.5}$ concentration forecast amplitudes in Brazil, especially for the Center West, Southeast, and South regions. During the period between the 5th epidemiological week of 2017 (29 January 2017) and the 27th epidemiological week of 2019 (6 July 2019) (represented by the grey background in Figure 1), CAMS NRT time-series predictions were stationary for all regions: an augmented Dickey–Fuller stationarity test resulted in $p$-values $\leq 0.01$ for all the regions but the Southeast one ($p$-value $= 0.013$). During this period, the dynamics of the PM$_{2.5}$ concentrations provided by CAMS NRT and CAMS reanalysis appeared similar. Given that, and despite the CAMS NRT predictions seeming to overestimate the PM$_{2.5}$ concentrations in comparison with the CAMS reanalysis and the newest version of the IFS, this period (from 29 January 2017 to 6 July 2019, i.e., from the 5th epidemiological week of 2017 to the 27th epidemiological week of 2019) was considered in the following to investigate the feasibility of an early warning system.

From the perspective of early warning system development, the forecasts (4-day horizon) of the last IFS cycle (46R1) were considered for one entire year from 10 July 2019 to 9 July 2020. This allowed us to simulate and discuss the effects of the different warning thresholds determined by the CONAMA 491/2018 resolution, considering the IFS version currently in use.

Eventually, an operational early warning system would be based on the PM$_{2.5}$ concentration forecasts for the current and subsequent days. CAMS provides access to such forecasts on a regular ∼40 km spatial resolution grid and for every hour from 03:00 UTM on the present day to 24:00 UTM on the fifth following day. Datasets are accessible through a File Transfer Protocol (FTP) service. From such data, daily mean PM$_{2.5}$ concentrations can be calculated for the four next days from the present one, considering the main Brazilian time zone (UTC–3). This dataset was not strictly included in the analysis but was used to provide an example of indicators upon which an early warning system could be based (Appendix B).
Figure 1. Weekly mean PM$_{2.5}$ concentrations as a function of the Brazilian regions from 2015 to 2019, provided by the near real time (NRT) forecast archive (4-day forecast horizon) of the Copernicus Atmosphere Monitoring Service (CAMS) (top) and by the CAMS reanalysis database (bottom). Dates of the implementation and the names of successive Integrated Forecasting System (IFS) versions are represented by the vertical red dashed lines. The 41R1 cycle was maintained on 21 June 2016 but with a horizontal resolution change, from $\sim$80 to $\sim$40 km. The grey background corresponds to the period considered in this study for investigating the feasibility of an early warning system.

2.1.3. Meteorological Data

Particulate matter concentrations are partly related to meteorological conditions. Notably, precipitation tends to leach suspended particles from the atmosphere; temperature impacts both the emissions of PM and chemical processes in the atmosphere. CAMS includes a meteorological forecasting model to predict the atmosphere composition. On the other hand, extremely low or high air humidity and/or temperature can directly and indirectly induce respiratory problems. For instance, low air humidity can induce respiratory tract irritations and thus increase susceptibility to viral
infections; low temperatures tend to increase virus lifetime. As such, meteorological conditions need be considered as possible confounding factors when studying the relationships between PM$_{2.5}$ concentrations and SARD cases.

Precipitation, temperature, and relative humidity were considered in this study. Measurements recorded by the automatic meteorological stations of the National meteorological institute of Brazil (INMET) were downloaded from the Institute’s website. They included minimum and maximum hourly temperature and relative humidity, as well as hourly rainfall. The daily minimum (maximum) was computed for the minimum (maximum) temperature and relative humidity, as well as the daily accumulated rainfall. Then, mean weekly values were computed for each meteorological variable for inclusion in the analysis with the SARD and PM$_{2.5}$ data.

2.2. Data Analysis

2.2.1. Population Exposure to Environmental Factors

The InfoGripe system provides SARD cases by state (the variable is referred here to as SARD), considering the notification health center (and not the patients’ residence) as the case location. As hospitals do not necessarily admit patients who reside in the same state, the putative exposure areas of the patients were defined by considering the recruiting zones of the hospitals (and not the state boundaries). The Brazilian Institute of Geography and Statistics (IBGE) defines the regions of influence of Brazilian cities (REGIC) by considering the patient movements to seek hospitals providing high complexity healthcare, i.e., healthcare that requires a set of procedures that involve high technology and high cost, as is the case for the events gathered by the InfoGripe database. In this study, REGIC were considered to delimit the areas within which the environmental factors possibly impacted SARD case occurrence. Such areas were defined by aggregating the REGIC of the cities that belong to the same state, providing areas that are referred to as statewide REGIC (SREGIC) hereafter.

Weekly mean concentrations of PM$_{2.5}$ were averaged by SREGIC (variable referred to as PM) (Appendices C and D). Meteorological variable means within the SREGIC were also computed. However, not all the ~600 automatic meteorological stations of the INMET network were functioning at the same time; even when a station was in operation, missing data produced gaps in the time series. Selecting only stations with complete time-series would have led to too-restrictive data selection and to SREGIC without data. Consequently, a compromise was achieved by selecting, for each SREGIC, the maximum number of stations that simultaneously provided data for at least 80% of the study period (i.e., 29 January 2017 to 6 July 2019). This allowed: (1) maximizing the number of selected stations (to provide an accurate estimation of the mean meteorological conditions within the SREGIC) and (2) minimizing the amount of missing values (up to 20%) for the meteorological variable time series (Appendices E–G).

2.2.2. Correlation between PM$_{2.5}$ Concentrations and SARD Cases

To consider the confounding effects of the meteorological conditions, the part of the variance (information) of the PM and SARD variables that was explained by meteorological conditions was previously removed from the two variables. The PM and SARD variables were considered separately, and for each SREGIC, the best multivariate linear regression model, in terms of the minimum Akaike information criterion (AIC), was selected by considering all combinations of 1 to 6 explanatory variables. The set of explanatory variables was composed of the 5 meteorological variables and their 10 possible two-way interactions (which corresponds to $C_{10}^{1} C_{10}^{2} = 9948$ models tested for each SREGIC). Then, for each SREGIC, the Pearson and Spearman correlation coefficients, to account for linear (Pearson correlation) and nonlinear (rank-based association detected by Spearman correlation) relationships, were computed between the residuals of the best models found for the PM and SARD variables, referred to as $uPM$ and $uSARD$, respectively. The procedure is schematically presented Figure 2.
Figure 2. Method used to consider the confounding effects of the meteorological variables in the computation of the correlation coefficients between PM and SARD. PREC: precipitation; TN: minimum temperature; TX: maximum temperature; HN: minimum relative humidity; HX: maximum relative humidity; SARD: Severe Acute Respiratory Disease case number; PM: particulate matter ($PM_{2.5}$) concentration.

3. Results

3.1. Characterization of the Exposure to Environmental Factors

The regions of exposure (SREGIC) differed from the state limits, as shown in Figure 3.

We selected 176, 202, and 229 meteorological stations to compute the mean weekly precipitation, relative humidity, and temperature in the SREGIC, respectively (Figure 4).

Distribution of these stations is highly heterogeneous, with 1 to 8 stations in the SREGIC of the North region, 2 to 17 in the Northeast, 6 to 10 in the Center West, 6 to 30 in the Southeast, and 10 to 20 in the South. Only one station was considered in Amapá and Roraima (both in the North region), whereas 30 were selected in Minas Gerais and 16 in the relatively small SREGIC of Rio de Janeiro.
Figure 3. (a) Brazil states and regions; (b) regions of Influence of cities (REGIC, source: Brazilian Institute of Geography and Statistics (IBGE)) aggregated per state (statewide REGIC (SREGIC)).

Figure 4. Distribution of the selected meteorological stations as a function of the meteorological variables.
3.2. Spatial Distribution of SARD Cases

The South region was the one the most affected by SARD during the period 2015–2019, with relatively high annual incidence rates (Figure 5).

Acre was the only state of the North region that exhibited a high incidence for the whole period, whereas the incidence tended to increase for the Amazonas state. In the Northeast region, Pernambuco was the only state with a high incidence in 2015, but from 2016, almost all states of the region exhibited relatively high incidences. The distributions of the incidence rates appeared similar in 2018 and 2019, except for the Amazonas state, which experienced an important increase in the SARD case number between 2018 and 2019 from 227 to 1911 cases.

A diagonal constituted of the states of Mato Grosso, Pará, Maranhão, and Amapá seemed less affected by SARD.

![Figure 5](image)

**Figure 5.** Severe Acute Respiratory Disease (SARD) annual incidence rate (for 100,000 inhabitants) from 2015 to 2019. For incidence rate computation, 2019 population estimates were used (source: Brazilian Institute of Geography and Statistics (IBGE)).

3.3. Spatial PM$_{2.5}$ Distribution

Mapping PM$_{2.5}$ concentrations provided by the CAMS reanalysis dataset shows that more than half of the country was exposed to annual mean concentrations of PM$_{2.5}$ that exceeded the WHO AQG (10 µg m$^{-3}$) (Figure 6, left column). High PM$_{2.5}$ concentrations were predicted all year long and for all the considered years in a strip located in the South of the Amazonas state at the border with Acre. In this region, concentrations significantly exceeded 50 µg m$^{-3}$ in some places (five times the WHO AQG). A region associated with high concentrations was also observed in a west–east-oriented strip crossing the Center West and Southeast regions. The states of Amapá, Pará (to the north), Maranhão, Rondônia, Mato Grosso (central), and Bahia (west part) seemed to be the least exposed to high annual mean concentrations, whereas very large states like Amazonas, Pará, Mato Grosso, and Bahia were exposed to heterogeneous levels of PM$_{2.5}$ concentrations in space.

By studying the maximum daily concentrations observed over the year (middle column in Figure 6), it appears that a main part of the country was punctually exposed to extremely high PM$_{2.5}$
concentrations, especially in comparison with the daily mean concentration reference value defined by the WHO (25 µg m\(^{-3}\)). The results showed that the southern region experienced very high maximum concentrations, despite mean annual concentration values being relatively low.

Figure 6. Annual mean (left column) and maximum (middle column) of the daily PM\(_{2.5}\) concentrations and number of days (as percentage of the year) for which the daily mean exceeded the World Health Organization (WHO) Air Quality Guideline (AQG) of 25 µg m\(^{-3}\)/24 h (right column).
3.4. PM$_{2.5}$ Concentration and SARD Case Correlations

Particulate matter concentrations and SARD cases corrected for the confounding effects of the meteorological conditions, i.e., $u_{PM}$ and $u_{SARD}$, respectively, were positively correlated in all SREGICs of the Center West, Northeast, and North regions, with the exception of the SREGIC of Rio Grande do Norte, which exhibited a nonsignificant negative correlation (Figures 7 and 8; to visualize corrected variables for each SREGIC, see Appendices H and I). Significant ($p$-value < 0.05) positive correlations were obtained for the SREGICs of Amapá, Alagoas, Bahia, Maranhão, Pernambuco, and Mato Grosso, with correlation coefficients varying from 0.24 (Mato Grosso) to 0.35 (Maranhão) considering the maximum of the Pearson and Spearman coefficients (Figures 7 and 8). Correlations were improved by first smoothing the $u_{PM}$ and $u_{SARD}$ variables with a three-week centered moving average. Half of the SREGICs belonging to the previously mentioned regions were associated with a significant correlation coefficient (reaching 0.52 for Maranhão, $p$-value < 0.01), the others being associated with positive but nonsignificant positive correlations, except for the SREGIC of Rio Grande do Norte, which was still associated with a nonsignificant negative correlation. The majority of the SREGICs that presented significant correlations were situated in a west–east-oriented strip from Mato Grosso to the south of the Northeast region, which coincides with exposures to relatively high PM$_{2.5}$ concentrations (Figure 6) and with high SARD incidences (Figure 5).

Negative correlations were found for the SREGICs of the Southeast and South regions, with significant values for Espírito Santo (considering the previously smoothed residuals) and for Rio Grande do Sul and Santa Catarina, the two southernmost SREGICs of the country (Figures 7 and 8).

In the SREGICs of Rio Grande do Sul and Santa Catarina, PM$_{2.5}$ concentrations and SARD cases presented a strong seasonality and a clearly opposite dynamic (Figure 9). However, a clear association was observed between the PM$_{2.5}$ concentrations and the negative test proportion in these two SREGICs (Figure 9).

![Figure 7. Pearson and Spearman correlation coefficients between weekly corrected time series of SARD cases and mean PM$_{2.5}$ concentrations: uSARD and uPM. Green and blue colors correspond to positive and negative correlations, respectively, and red corresponds to the test significance level.](image-url)
Figure 8. Cartographic representation of the correlations between (a) $u_{PM}$ and $u_{SARD}$ and (b) smoothed $u_{PM}$ and $u_{SARD}$ using a 3-week centered moving average.

Figure 9. PM$_{2.5}$ concentration forecasts (archived CAMS NRT) provided 4 days before the current date and averaged per epidemiological week (black line) and weekly case number of SARD as a function of the exam status: cases with laboratory test positive to viral infection in red; cases with negative viral infection laboratory test in orange; cases awaiting test results in yellow; cases without laboratory information in blue; cases without test in light blue; others in grey. Data are presented for Rio Grande do Sul and Santa Catarina from 29 January 2017 to 6 July 2019, i.e., from the 5th epidemiological week of 2017 to the 27th epidemiological week of 2019. For a representation of the results for all SREGICs, see Appendices C and D.
This observation led us to consider the negative test rate in place of the total number of cases in the analysis. This negative test rate is defined as follows:

\[
NR = \frac{\text{Neg}}{\text{Neg} + \text{Pos}}
\]  

(1)

where \(\text{Neg}\) and \(\text{Pos}\) are the number of negative and positive tests for viral infection by epidemiological week, respectively. By applying the same procedure to remove the information attributable to meteorological conditions, \(uNR\) and the correlation coefficients between \(uNR\) and \(uPM\) were computed. The results showed a tendency to positive correlations for the SREGICs of the South and Southeast regions (Figure 10). Correlations were very significant for three SREGICs: Rio de Janeiro, Rio Grande do Sul, and Santa Catarina.

<table>
<thead>
<tr>
<th>Regions</th>
<th>State</th>
<th>(uPM) vs. (uNR)</th>
<th>Pearson</th>
<th>Spearman</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>r</td>
<td>p-value</td>
<td></td>
</tr>
<tr>
<td>Southeast</td>
<td>Espírito Santo</td>
<td>0.14</td>
<td>0.128</td>
<td>0.10</td>
</tr>
<tr>
<td></td>
<td>Minas Gerais</td>
<td>0.07</td>
<td>0.447</td>
<td>0.08</td>
</tr>
<tr>
<td></td>
<td>Rio de Janeiro</td>
<td>0.04</td>
<td>0.664</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>São Paulo</td>
<td>-0.02</td>
<td>0.817</td>
<td>0.01</td>
</tr>
<tr>
<td>South</td>
<td>Paraná</td>
<td>-0.01</td>
<td>0.952</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>Rio Grande do Sul</td>
<td>0.31</td>
<td>0.000</td>
<td>0.27</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Smoothed (uPM) vs. Smoothed (uNR)</th>
<th>Pearson</th>
<th>Spearman</th>
</tr>
</thead>
<tbody>
<tr>
<td>r</td>
<td>p-value</td>
<td></td>
</tr>
<tr>
<td>0.07</td>
<td>0.451</td>
<td>0.06</td>
</tr>
<tr>
<td>-0.09</td>
<td>0.346</td>
<td>-0.12</td>
</tr>
<tr>
<td>0.18</td>
<td>0.042</td>
<td>0.18</td>
</tr>
<tr>
<td>-0.03</td>
<td>0.752</td>
<td>0.01</td>
</tr>
<tr>
<td>-0.08</td>
<td>0.400</td>
<td>-0.01</td>
</tr>
<tr>
<td>0.31</td>
<td>0.000</td>
<td>0.30</td>
</tr>
<tr>
<td>0.43</td>
<td>0.000</td>
<td>0.26</td>
</tr>
<tr>
<td>0.00</td>
<td>0.001</td>
<td>0.004</td>
</tr>
</tbody>
</table>

**Figure 10.** Pearson and Spearman correlation coefficients between the corrected negative test rate and mean PM\(_{2.5}\) concentrations: \(uNR\) and \(uPM\). Green and blue correspond to positive and negative correlations, respectively, and red corresponds to the test significance level.

3.5. An Early Warning System

By considering the archived CAMS NRT forecasts on a four-day horizon from 10 September 2019 to 9 September 2020, i.e., for an entire year associated with the latest IFS version (46R1), we found that the daily mean PM\(_{2.5}\) concentrations per municipality exceeded the WHO AQG of 25 \(\mu g m^{-3}\) at least one day over the entire year for 2071 municipalities (37%; Figure 11). Two municipalities (Eirunepé and Itamarati), situated in the south of the Amazonas state, were exposed to concentrations exceeding 25 \(\mu g m^{-3}\) for more than 120 days (~4 months) in the year. The eight municipalities exposed to concentrations above 25 \(\mu g m^{-3}\) for more than 60 days (~2 months) were all situated in the Amazonas state, except for the municipality of Serra do Ramalho in the state of Bahia.

A total of 309, 450 and 905 municipalities had mean PM\(_{2.5}\) concentrations exceeding, at least one day during the considered period, the intermediate reference values (established by the CONAMA 491/2018 resolution) of 60, 50, and 37 \(\mu g m^{-3}\), respectively. A low number of municipalities exceeded, at least one day during the considered period, the alert thresholds established by the CONAMA 491/2018 resolution: 40, 8, and 2 municipalities for the attention, alert, and emergency levels, respectively (Figure 11).
Figure 11. Number of days from 10 July 2019 to 9 July 2020 (one year) that the municipalities of Brazil exceeded the different alert and intermediate reference values (IRVs) established by the CONAMA 491/2018 resolution for daily mean PM$_{2.5}$ concentration. Numbers in square brackets correspond to the numbers of municipalities.

4. Discussion

4.1. General Approach

We aimed at establishing if the SARD case numbers and the four-day horizon forecasts of PM$_{2.5}$ concentrations evolve similarly in Brazil to determine the ability to use the PM$_{2.5}$ forecasts to predict respiratory disorders in the Brazilian population and consequently the feasibility of building an early warning system for PM$_{2.5}$-related health problems. Simple correlation coefficients, accounting for both linear (Pearson correlation) and nonlinear (Spearman rank-based correlation) relationships between the SARD cases and the PM$_{2.5}$ concentrations corrected from the confounding effects of meteorological
conditions, were found to be sufficient for this purpose. The modeling of such relationships through, for instance, linear regression, generalized additive models (GAMs), or case-crossover models, was not considered. In fact, such models could provide additional information but would have been relevant for the study period only (from the 5th epidemiological week of 2017 to the 27th epidemiological week of 2019) due to the significant and frequent changes observed in the amplitudes of the PM$_{2.5}$ concentrations due to the updates of the IFS (Figure 1).

### 4.2. Early Warning System Feasibility

When analysing data, the application of a three-week centered moving average was justified by a potential time lag occurring between the PM$_{2.5}$ and SARD time series, in both directions. A 6- to 7-day lag was observed between exposure to PM and significant increases in ambulatory admissions of children [10], indicating that one week can separate the peaks of exposure and hospital admission. On the other hand, PM$_{2.5}$ concentrations with a four-day forecast horizon are based on the assimilation of information (AOD and occurrence of events including fire events) observed by satellites four days before. This could result in a delay of up to four days in the detection of PM$_{2.5}$ concentration peaks related to fire events, which could lead, in our study, to a one-week delay between the hospital admission and the predicted PM$_{2.5}$ peak.

Negative correlations found in the South region may appear counterintuitive. However, complex interactions between virus type, climate, and population behavior in this region can explain this result. A significant part of the SARD case dynamics and magnitude in the South could have been driven more by viral infections than by outdoor environmental factors, despite the meteorological variables explaining 78% ($R^2 = 0.78$) and 65% of the total variance of the SARD cases in the SREGICs of Rio Grande do Sul and Santa Catarina, respectively (values ranked first and third among all SREGIC). Respiratory syncytial virus (RSV) circulates particularly during the winter (July to September) in the South and Southeast regions, with test positivity being twice to three times higher than in the other regions for children under five years old [34]. RSV infections were found to be not associated with meteorological factors in Kentucky, the United States [35]. The authors argued, in particular, that patients were not exposed to particular outdoor temperature or humidity levels (and we can add particular outdoor PM concentration levels here), with people living more indoor during the winter season. Even if climate differs between Kentucky and the southern Brazilian region, comparable population behavior can be assumed, especially among young children that are the most affected by RSV. This may be the reason why outdoor environmental factors do not fully explain the SARD seasonality in the South region and, to a lesser extent, in the Southeast region of Brazil. However, PM$_{2.5}$ concentration and negative test rate being significantly correlated demonstrated that PM$_{2.5}$ tended to explain nonviral infections in these regions, and that PM$_{2.5}$ is an accurate indicator of SARD case increase.

The delimitation of the areas of exposure to environmental factors is probably the weakest point of the study and possibly contributed to reducing the significance of the results. This delimitation was based on the REGIC, aggregated by state. It was used to consider the recruiting basins of the hospitals but led to fragmented areas, which, for some states, had different environmental and socioeconomic contexts and different PM$_{2.5}$ concentration levels, pollution source types, population activities, etc., leading to different exposure risks. This was, for instance, the case for the São Paulo SREGIC, which includes a non-negligible part of the Mato Grosso state (Figure 3). Moreover, large SREGIC like those associated with Pará and Amazonas exhibited heterogeneous spatial distributions of PM concentrations. The mean, chosen as the summary value for SREGICs, tended to smoothing information both in space and time. However, by testing other summary statistics (median, max, and number of days of the epidemiological week above WHO reference value), the mean appeared the more robust to extreme values, significantly linked to SARD cases, and easy to interpret. Eventually, an early warning system would not consider such a high aggregation level (state level). As shown in the example provided in Appendix B, this system should be based on mean PM concentrations at
the municipality or autochthonous territory (Terra Indígena) level, or even at the 40 km pixel level, which would result in a more accurate estimation of exposure.

The different sources of PM, and thus its chemical composition and toxicity, could have been considered in this study, as Brazil is exposed to different pollution sources and types. The Amazon region is mostly affected by forest fires, and Central and Southern Brazil are mainly influenced by urban activities (industry and transportation). More generally, PM$_{2.5}$ concentration should be considered as one of the components of air quality indexes, which should also integrate meteorological conditions. This issue should be investigated in the future.

Nevertheless, we succeeded in demonstrating the possibility of using PM$_{2.5}$ concentration forecasts provided by CAMS to predict respiratory disorders in the Brazilian population and, consequently, of developing an early warning system in Brazil based on this product.

4.3. Alert Thresholds

When building an early warning system, alert threshold(s) associated with PM$_{2.5}$ concentrations must be defined. This definition was outside the scope of our initial objectives and will constitute the main issue of further works, but some considerations can be outlined here.

This definition is partly related to the ability of the CAMS products to provide an accurate absolute estimation of the PM$_{2.5}$ concentrations. CAMS is supposed to predict reliable absolute PM$_{2.5}$ concentration values, and accuracy and biases are regularly estimated (see, for example, Ref. [36]).

However, the predicted PM$_{2.5}$ concentration levels reported in this study for the Center West region seem to be underestimated considering that this region is particularly affected by fires. This demonstrates the need for a more specific evaluation of the products in South America, and in Brazil in particular, that would help identify local biases, associated correction factors, and reference values from a public health perspective. The impacts of the successive IFS upgrades on the amplitudes of the predictions (Figure 1) cast doubt on the product’s absolute accuracy. In particular, the IFS underwent a major upgrade on 9 July 2019 (IFS cycle 46R1), associated with changes in the vertical resolution (from 60 vertical levels for the previous models to 137 levels for the new one) and in the model components. This upgrade led to “pronounced improvement of surface PM$_{2.5}$ and PM$_{10}$ forecasts, especially during night time, by changes to the diurnal cycle of emissions and the introduction of injection heights for biomass burning emissions” (https://atmosphere.copernicus.eu/node/472). More specifically, overestimation of near-surface PM$_{2.5}$ during fire events was reduced by changes in the biomass-burning injection heights from the Global Fire Assimilation System (GFAS) and the updated diurnal cycle. Simultaneously, the anthropogenic Secondary Organic Aerosol (SOA) emission was updated with a diurnal cycle and regionally-varying ratio to carbon monoxide (CO) emissions, resulting in a significant reduction in night-time near surface PM$_{2.5}$ concentration in polluted regions [36]. These changes are important for Brazil due to the extent of biomass burning events, the presence of industrial activities, and mega-cities like São Paulo and Rio de Janeiro. As a consequence, a significant drop in the estimated PM$_{2.5}$ concentrations was observed, especially in the Center West, Southeast, and South regions.

However, another perspective on the alert threshold definition issue can be adopted. The results showed how the definition of different thresholds highly impacts the number of municipalities experiencing an alert (Figure 11). Accurate predictions associated with an alert threshold that meets national and international guidelines would theoretically be ideal but can potentially generate a high number of alerts that would be difficult for local authorities to manage, notably due to limited public service resources and population acceptability. This consequently could lead to the impossibility of actually and systematically implementing the necessary prevention and mitigation actions. Moreover, no threshold effect has been associated with the health-related impact of PM exposure [24]. In that sense, a pragmatical method is required to define local alert thresholds according to local public authority and population capacities. This definition would require a highly interdisciplinary approach involving epidemiology, environmental science, human science, and public health.
5. Conclusions

In this study, we demonstrated the possibility of using PM\textsubscript{2.5} concentration forecasts provided by CAMS to predict SARD cases in Brazil up to four days in advance and, consequently, to be part of an early warning system of the health-related impacts of PM. Through the relationship to different pollution types, this system could be effective for the whole Brazilian territory and beyond. PM has been proven to aggravate COVID-19 symptoms. The mobilization of researchers, public health actors, and the general population in the fight against the pandemic, particularly during the current environmental disasters related to deforestation and forest fires, could provide an opportunity to develop an alert system and implement both PM emission reduction actions and mitigation measures, such as the wearing of adapted masks. This also would contribute to achieving the United Nation Sustainable Development Goals 3.9: by 2030, substantially reduce the number of deaths and illnesses from hazardous chemicals and air, water, and soil pollution and contamination.


Funding: This research was supported by the MARIONETTE3 project, funded by the Observations of Natural Environments and Global Changes (OMNCG) Research Federation of the Observatory of Universe Sciences of Reunion Island (OSU-Réunion) and Reunion University.

Acknowledgments: We want to thank the Copernicus Helpdesk at ECMWF for all the answers to our questions on CAMS.

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

Abbreviations

The following abbreviations are used in this manuscript:

AIC Akaike Information Criterion
AOD Aerosol Optical Depth
CAMS Copernicus Atmosphere Monitoring Service
CONAMA Conselho Nacional de Meio Ambiente
ECMWF European Centre for Medium-Range Weather Forecasts
EMAp Escola de Matematica Aplicada
ESA European Space Agency
EUMETSAT European Organisation for the Exploitation of Meteorological Satellites
FGV Fundação Getulio Vargas
Fiocruz Fundação Oswaldo Cruz
FTP File Transfer Protocol
GMAO Global Monitoring and Assimilation Service
HN minimum relative humidity
HX maximum relative humidity
IBAMA Instituto Brasileiro do meio ambiente e dos recursos naturais renováveis
IBGE Instituto Brasileiro de Geografia e Estatística
IFS Integrated Forecasting System
INMET Instituto nacional de meteorologia
INPE Instituto Nacional de Pesquisa Espacial
MODIS Moderate Resolution Imaging Radiospectrometer
NASA National Aeronautics and Space Administration
NRT Near Real Time
PM Particulate Matter
PMAp  Polar Multi-sensor Aerosol product
PREC  precipitation
PROCC Programa de computação científica
RSV   Respiratory syncytial virus
REGIC Regiões de Influência das Cidades
SREGIC Statewide REGIC
SARD  Severe Acute Respiratory Diseases
SARS  Severe Acute Respiratory Syndromes
SINAN Sistema de Informação de Agravos de Notificação
SOA   Secondary Organic Aerosol
TN    minimum temperature
TX    maximum temperature
UTC   Coordinated Universal Time
WHO   World Health Organization
**Appendix A. Data Sources and Description**

**Table A1.** Data used in the study: sources and description.

<table>
<thead>
<tr>
<th>Data</th>
<th>Spatial Resolution/ Spatial Unit</th>
<th>Source</th>
<th>Type of Access (File Format)</th>
<th>Question Addressed in the Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>PM$_{2.5}$ concentration reanalyses</td>
<td>~80 km</td>
<td>CAMS reanalyses, Copernicus Atmosphere Data Store <a href="https://ads.atmosphere.copernicus.eu/#!/home">https://ads.atmosphere.copernicus.eu/#!/home</a></td>
<td>HTTPS (GRIB or NETCDF)</td>
<td>Description of the past PM$_{2.5}$ concentration distribution, in space and time</td>
</tr>
<tr>
<td>Archived PM$_{2.5}$ concentration forecasts (every 3 h)</td>
<td>~40 km</td>
<td>CAMS near real-time database: <a href="https://apps.ecmwf.int/datasets/data/eams-nrealtime/levtype=sfc/">https://apps.ecmwf.int/datasets/data/eams-nrealtime/levtype=sfc/</a></td>
<td>ECMWF (GRIB or NETCDF)</td>
<td>Feasibility study of an early warning system; simulation of alert threshold definition</td>
</tr>
<tr>
<td>PM$_{2.5}$ hourly concentration forecasts up to 5 days after the current day</td>
<td>~40 km</td>
<td>CAMS Global forecast data <a href="https://confluence.ecmwf.int/display/CKB/FTP+access+to+CAMS+global+data">https://confluence.ecmwf.int/display/CKB/FTP+access+to+CAMS+global+data</a></td>
<td>FTP (GRIB)</td>
<td>Proposition for PM$_{2.5}$ concentration indicators to be included in an early warning system (Appendix F)</td>
</tr>
<tr>
<td>States of Brazil</td>
<td>State</td>
<td>brazilmaps R package</td>
<td>(R object)</td>
<td>Cartography of administrative units of Brazil</td>
</tr>
<tr>
<td>Municipalities of Brazil (Municipios)</td>
<td>Municipalities</td>
<td>brazilmaps R package</td>
<td>(R object)</td>
<td>Simulation of alert threshold definition</td>
</tr>
<tr>
<td>2019 population estimate in Brazil</td>
<td>State</td>
<td>brazilmaps R package</td>
<td>(R object)</td>
<td>SARD annual incidence</td>
</tr>
</tbody>
</table>
Appendix B. Examples of PM$_{2.5}$ Concentration Forecasts Obtained from CAMS NRT on Which an Early Warning System Could Be Based

Figures A1 and A2 show the daily mean PM$_{2.5}$ concentration forecasts for the next four days from the current date (15 September 2020) at pixel (Figure A1) and municipality (Figure A2) levels. The municipality of Cujubim (RO) is supposed to exceed the WHO AQG for the daily mean PM$_{2.5}$ concentration on each of the next four days, with a predicted concentration above 50 $\mu$g m$^{-3}$ on 18 September 2020. This municipality regularly experiences forest fires. Four farmers were arrested for illegal deforestation and fires in August 2020 (https://globoplay.globo.com/v/8796016/), and the municipality was among those most affected by fires in the state of Rondônia in the first week of September 2020, along with the municipality of Porto Velho (https://g1.globo.com/ro/rondonia/natureza/amazonia/noticia/2020/09/08/rondonia-acumula-mais-de-mil-focos-de-queimadas-nos-primeiros-7-dias-de-setembro-de-2020-alta-e-de-70percent.ghtml).

The other municipalities predicated to exceed the WHO AQGs for the next four days after the 15 September 2020 were Minaçu (GO), Itamarati (AM), Alvorada (BA, for two consecutive days), Eirunepé (AM), Envia (AM), and Atalaia do Norte (AM).

![Figure A1. Daily predicted mean concentration of PM$_{2.5}$ in Brazil, for 15–18 September, 2020 at the pixel level.](image-url)
Figure A2. Daily predicted mean concentration of PM$_{2.5}$ in Brazil per municipality for 15–18 September 2020 at the municipality level.
Appendix C. PM$_{2.5}$ Concentration Forecasts and Weekly Case Number of SARD

Figure A3. PM$_{2.5}$ concentration forecasts (archived CAMS NRT) provided four days before the current date and averaged per epidemiological week (black line) and weekly case number of SARD as a function of the exam status: cases with positive viral infection laboratory test (red); cases with negative viral infection laboratory test (orange); cases awaiting test results (yellow); cases without laboratory information (blue); cases without test (light blue); others (grey). Data are presented by SREGIC from 29 January 2017 to 6 July 2019, i.e., from the 5th epidemiological week of 2017 to the 27th epidemiological week of 2019.
Figure A4. PM$_{2.5}$ concentration forecasts (archived CAMS NRT) provided four days before the current date and averaged per epidemiological week (black line) and weekly case number of SARD as a function of the exam status: cases with positive viral infection laboratory test (red); cases with negative viral infection laboratory test (orange); cases awaiting test results (yellow); cases without laboratory information (blue); cases without test (light blue); others (grey). Data are presented by SREGIC from 29 January 2017 to 6 July 2019, i.e., from the 5th epidemiological week of 2017 to the 27th epidemiological week of 2019.
Figure A5. PM$_{2.5}$ concentration forecasts (archived CAMS NRT) provided four days before the current date and averaged per epidemiological week (black line) and weekly case number of SARD as a function of the exam status: cases with positive viral infection laboratory test (red); cases with negative viral infection laboratory test (orange); cases awaiting test results (yellow); cases without laboratory information (blue); cases without test (light blue); others (grey). Data are presented by SREGIC from 29 January 2017 to 6 July 2019, i.e., from the 5th epidemiological week of 2017 to the 27th epidemiological week of 2019.
Appendix D. PM$_{2.5}$ Concentration Forecasts and Weekly Case Number of SARD (in %)

Figure A6. PM$_{2.5}$ concentration forecasts (archived CAMS NRT) provided four days before the current date and averaged per epidemiological week (black line) and weekly case number of SARD expressed as a percentage of the total number of reported cases and as a function of the exam status: cases with positive viral infection laboratory test (red); cases with negative viral infection laboratory test (orange); cases awaiting test results (yellow); cases without laboratory information (blue); cases without test (light blue); others (grey). Data are presented by SREGIC from 29 January 2017 to 6 July 2019, i.e., from the 5th epidemiological week of 2017 to the 27th epidemiological week of 2019.
Figure A7. PM$_{2.5}$ concentration forecasts (archived CAMS NRT) provided four days before the current date and averaged per epidemiological week (black line) and weekly case number of SARD expressed as a percentage of the total number of reported cases and as a function of the exam status: cases with positive viral infection laboratory test (red); cases with negative viral infection laboratory test (orange); cases awaiting test results (yellow); cases without laboratory information (blue); cases without test (light blue); others (grey). Data are presented by SREGIC from 29 January 2017 to 6 July 2019, i.e., from the 5th epidemiological week of 2017 to the 27th epidemiological week of 2019.
Figure A8. PM$_{2.5}$ concentration forecasts (archived CAMS NRT) provided four days before the current date and averaged per epidemiological week (black line) and weekly case number of SARD expressed as a percentage of the total number of reported cases and as a function of the exam status: cases with positive viral infection laboratory test (red); cases with negative viral infection laboratory test (orange); cases awaiting test results (yellow); cases without laboratory information (blue); cases without test (light blue); others (grey). Data are presented by SREGIC from 29 January 2017 to 6 July 2019, i.e., from the 5th epidemiological week of 2017 to the 27th epidemiological week of 2019.
Appendix E. Rainfall and Weekly Case Number of SARD

Figure A9. Rainfall averaged per epidemiological week, and weekly case number of SARD as a function of the exam status: cases with positive viral infection laboratory test (red); cases with negative viral infection laboratory test (orange); cases awaiting test results (yellow); cases without laboratory information (blue); cases without test (light blue); others (grey). Data are presented by SREGIC, from 29 January 2017 to 6 July 2019, i.e., from the 5th epidemiological week of 2017 to the 27th epidemiological week of 2019.
Figure A10. Rainfall averaged per epidemiological week, and weekly case number of SARD as a function of the exam status: cases with positive viral infection laboratory test (red); cases with negative viral infection laboratory test (orange); cases awaiting test results (yellow); cases without laboratory information (blue); cases without test (light blue); others (grey). Data are presented by SREGIC, from 29 January 2017 to 6 July 2019, i.e., from the 5th epidemiological week of 2017 to the 27th epidemiological week of 2019.
Figure A11. Rainfall averaged per epidemiological week, and weekly case number of SARD as a function of the exam status: cases with positive viral infection laboratory test (red); cases with negative viral infection laboratory test (orange); cases awaiting test results (yellow); cases without laboratory information (blue); cases without test (light blue); others (grey). Data are presented by SREGIC, from 29 January 2017 to 6 July 2019, i.e., from the 5th epidemiological week of 2017 to the 27th epidemiological week of 2019.
Appendix F. Temperature and Weekly Case Number of SARD

Figure A12. Minimum and maximum temperatures averaged per epidemiological week and weekly case number of SARD as a function of the exam status: cases with positive viral infection laboratory test (red); cases with negative viral infection laboratory test (orange); cases awaiting test results (yellow); cases without laboratory information (blue); cases without test (light blue); others (grey). Data are presented by SREGIC, from 29 January 2017 to 6 July 2019, i.e., from the 5th epidemiological week of 2017 to the 27th epidemiological week of 2019.
Figure A13. Minimum and maximum temperatures averaged per epidemiological week and weekly case number of SARD as a function of the exam status: cases with positive viral infection laboratory test (red); cases with negative viral infection laboratory test (orange); cases awaiting test results (yellow); cases without laboratory information (blue); cases without test (light blue); others (grey). Data are presented by SREGIC, from 29 January 2017 to 6 July 2019, i.e., from the 5th epidemiological week of 2017 to the 27th epidemiological week of 2019.
Figure A14. Minimum and maximum temperatures averaged per epidemiological week and weekly case number of SARD as a function of the exam status: cases with positive viral infection laboratory test (red); cases with negative viral infection laboratory test (orange); cases awaiting test results (yellow); cases without laboratory information (blue); cases without test (light blue); others (grey). Data are presented by SREGIC, from 29 January 2017 to 6 July 2019, i.e., from the 5th epidemiological week of 2017 to the 27th epidemiological week of 2019.
Appendix G. Relative Humidity and Weekly Case Number of SARD

Figure A15. Minimum and maximum relative humidity averaged per epidemiological week and weekly case number of SARD as a function of the exam status: cases with positive viral infection laboratory test (red); cases with negative viral infection laboratory test (orange); cases awaiting test results (yellow); cases without laboratory information (blue); cases without test (light blue); others (grey). Data are presented by SREGIC, from 29 January 2017 to 6 July 2019, i.e., from the 5th epidemiological week of 2017 to the 27th epidemiological week of 2019.
Figure A16. Minimum and maximum relative humidity averaged per epidemiological week and weekly case number of SARD as a function of the exam status: cases with positive viral infection laboratory test (red); cases with negative viral infection laboratory test (orange); cases awaiting test results (yellow); cases without laboratory information (blue); cases without test (light blue); others (grey). Data are presented by SREGIC, from 29 January 2017 to 6 July 2019, i.e., from the 5th epidemiological week of 2017 to the 27th epidemiological week of 2019.
Figure A17. Minimum and maximum relative humidity averaged per epidemiological week and weekly case number of SARD as a function of the exam status: cases with positive viral infection laboratory test (red); cases with negative viral infection laboratory test (orange); cases awaiting test results (yellow); cases without laboratory information (blue); cases without test (light blue); others (grey). Data are presented by SREGIC, from 29 January 2017 to 6 July 2019, i.e., from the 5th epidemiological week of 2017 to the 27th epidemiological week of 2019.
Appendix H. Residuals of the Multivariate Linear Regression Models for PM$_{2.5}$ Concentrations and Case Number of SARD

Figure A18. $uPM$ (black line) and $uSARD$ (red line) by SREGIC, from 29 January 2017 to 6 July 2019, i.e., from the 5th epidemiological week of 2017 to the 27th epidemiological week of 2019.
Figure A19. $uPM$ (black line) and $uSARD$ (red line) by SREGIC, from 29 January 2017 to 6 July 2019, i.e., from the 5th epidemiological week of 2017 to the 27th epidemiological week of 2019.

Figure A20. $uPM$ (black line) and $uSARD$ (red line) by SREGIC, from 29 January 2017 to 6 July 2019, i.e., from the 5th epidemiological week of 2017 to the 27th epidemiological week of 2019.
Appendix I. Residuals of the Multivariate Linear Regression Models for PM$_{2.5}$ Concentrations and SARD Negative Test Rate

Figure A21. $uPM$ (black line) and $uNR$ (red line) by SREGIC from 29 January 2017 to 6 July 2019, i.e., from the 5th epidemiological week of 2017 to the 27th epidemiological week of 2019.
Figure A22. $uPM$ (black line) and $uNR$ (red line) by SREGIC from 29 January 2017 to 6 July 2019, i.e., from the 5th epidemiological week of 2017 to the 27th epidemiological week of 2019.

Figure A23. $uPM$ (black line) and $uNR$ (red line) by SREGIC from 29 January 2017 to 6 July 2019, i.e., from the 5th epidemiological week of 2017 to the 27th epidemiological week of 2019.
References


Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.