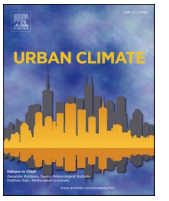
Urban Climate 55 (2024) 101876

Contents lists available at ScienceDirect

Urban Climate

journal homepage: www.elsevier.com/locate/uclim



Short term forecasting of persistent air quality deterioration events in the metropolis of Sao Paulo

L.V. Rizzo [a,*], A.G.B. Miranda [b]

[a] Physics Institute, Universidade de Sao Paulo, Sao Paulo, Brazil
[b] Department of Environmental Sciences, Universidade Federal de Sao Paulo, Diadema, Brazil

ARTICLE INFO

Keywords:
Air pollution
extreme events
Particulate matter
Ozone
Mitigation
Classification models

ABSTRACT

Air pollution is one of the main environmental problems in metropolitan areas. Negative impacts to human health are intensified when poor air quality conditions persist for many consecutive days, with gradual accumulation of pollutants in the atmosphere. Persistent air quality deterioration events are typically associated with occurrence of stagnant atmospheric conditions and reach the regional scale. In this study, data-driven models were developed to forecast the occurrence of persistent air pollution events of inhalable particulate matter (PM_{10}) and ozone (O_3) in the metropolis of Sao Paulo, Brazil. On average, 8 events per year were observed between 2005 and 2022, comprising 73 event days per year. The logistic regression method was used in a supervised learning framework. Daily timeseries of surface weather variables were used as predictors. In the case of PM_{10}, a consistent long-term decrease in the number events impacted the model performance. The PM_{10} model benefited from the restriction of the training set to recent years, with a significant increase in the model accuracy despite the reduction in the volume of data. The final models correctly reproduced the seasonal distribution of events, with overall accuracies of 0.92 and 0.87 for O_3 and PM_{10}, respectively, in 2022. Despite the fact that persistent exceedance events are relatively rare, the models were able to detect 81% and 97% of the event days in 2022, respectively for O_3 and PM_{10}. Daily maximum temperature was an important predictor, increasing the event odds by 483% (O_3) and 84% (PM_{10}). The classification models developed in this study can successfully forecast the occurrence of regional air pollution events concerning both primary and secondary air pollutants, which have different drivers for accumulation in the atmosphere. The models require simple input data and low computational resources, aiming to stimulate future usage by the general public and decision-makers, in order to mitigate exposure to harmful air pollutant concentrations.

1. Introduction

Air quality deterioration is one of the most relevant environmental problems in metropolitan areas. The combination of high population densities and gathering of air pollution emission sources frequently leads to a vast exposure to air pollutant concentrations above the standards, with negative impacts on health and quality of life (Molina et al., 2020; Gurjar et al., 2008). The association between poor air quality and cardio-respiratory diseases is clearly depicted in the literature (Anderson, 2009; Brook et al., 2010).

* Corresponding author.
E-mail addresses: lrizzo@usp.br (L.V. Rizzo), andre.miranda@unifesp.br (A.G.B. Miranda).

https://doi.org/10.1016/j.uclim.2024.101876
Received 19 November 2023; Received in revised form 15 March 2024; Accepted 19 March 2024
Available online 22 March 2024

Globally, review studies have shown that a 10 $\mu g/m^3$ increase in the concentration of fine particulate matter is associated with increased mortality risks due to cardiovascular disease (excess risk of 11%) and due to non-malignant respiratory diseases (excess risk of 3%) (Hoek et al., 2013). The Global Burden of Diseases Study attributed 7.6% of the total global mortality in 2015 to the long-term exposure to ambient air pollution (Cohen et al., 2017).

Exceedances of air quality standards are recurrent in megacities like New Delhi, Los Angeles and Sao Paulo (de Andrade et al., 2017; Parrish et al., 2011; Singh et al., 2021). Inhalable particulate matter (PM_{10}) and ozone (O_3) are air pollutants that frequently exceed air quality standards in the aforementioned megacities. While PM_{10} is mostly primary, i.e., directly emitted by air pollution sources, O_3 is a secondary pollutant, being produced in the atmosphere by photochemical reactions. In the Brazilian megacity of Sao Paulo, vehicles are the dominant source of air pollutants. Vehicles in Sao Paulo are fueled by a blend of fossil and biofuels like ethanol and biodiesel, resulting in a peculiar hydrocarbon composition and unique effects to the atmospheric chemistry (Kumar et al., 2016; Salvo and Geiger, 2014). Emission control measures succeeded on reducing primary air pollutant concentrations in Sao Paulo in the last decades, but challenges remain for the control of secondary air pollutants and unregulated emission sources (de Andrade et al., 2017).

In many countries, the regulation of emission sources is lenient, so that the air quality standards preconized by the World Health Organization are frequently not met (Gómez Peláez et al., 2020). However, the strength of emission sources is not the only factor that controls the accumulation of pollutants in the atmosphere. Weather conditions modulate air pollutant concentrations by affecting their dispersion, removal and photochemical production rates. While emission sources can be regulated, weather conditions are not under control, so that mitigation measures must be taken to avoid exposure to harmful concentrations in periods of unfavorable air pollution dispersion.

Unfavorable dispersion conditions are typically observed under the influence of high-pressure atmospheric systems, leading to dry and stagnant conditions following air pollution accumulation (Tai et al., 2010; Zhang et al., 2016; Oliveira et al., 2022). Mesoscale circulations driven by topography and land cover can also play a role (Ribeiro et al., 2018; Ning et al., 2019). Moreover, dry conditions increase the risk of vegetation fires, resulting in a positive feedback for air quality deterioration (Martins et al., 2018; Chen et al., 2017). Previous studies in the megacity of São Paulo identified synoptic weather patterns that favor the accumulation of air pollutants, like pre-frontal conditions and the influence of the South Atlantic Subtropical Anticyclone (SASA) (Oliveira et al., 2022; Sánchez-Ccoyllo and Andrade, 2002). High-pressure systems and atmospheric stagnation are regional scale phenomena, so that poor air quality conditions associated with their influence typically extend throughout whole metropolitan areas and may persist for several days. Acute exposure to high concentrations of air pollutants aggravates cardio-respiratory diseases in vulnerable sectors of the population (e.g., Brook et al., 2010). Previous studies have shown increased morbidity and hospital admissions following periods of intense and persistent air pollution episodes (Szyszkowicz et al., 2018; Zhang et al., 2014).

Given the impacts of persistent air quality deterioration events on human health, it is important to develop tools for their prediction in the short term, so that measures can be taken to mitigate exposure to harmful air pollutant concentrations. Real-time air quality forecasting can be pursued using deterministic (process-based) and statistical (data-driven) models. Deterministic models represent explicitly the major physical and chemical processes that influence air pollutant concentrations, being able to forecast spatial and temporal resolved concentrations (Zhang et al., 2012; Longo et al., 2013). To provide accurate forecasts, deterministic models require sophisticated prior knowledge, detailed information on emissions and meteorological fields and demand high performance computer systems. Statistical models, on its turn, explore relationships between air quality and weather variables, developing predictive models based on ordinary regression and classification techniques and machine learning methods, requiring less domain expertise and fewer computational resources. Statistical models are able to capture complex site-specific relationships between air pollutants and weather, often reaching higher accuracy compared to deterministic models (Zhang et al., 2012; Li et al., 2022). Whatever the methods applied, air quality forecasting models aim to provide timely information to protect the health of inhabitants through early warnings and control measures.

Statistical models for air quality are frequently used for explanatory purposes, i.e., aiming to investigate the relative importance of drivers and processes. A myriad of previous studies investigated the impact of weather variables on air pollutant concentrations using statistical models (e.g., Schuch et al., 2019; Porter et al., 2015; Tai et al., 2010). Other studies developed data-driven models to forecast air pollutant concentrations based on weather variables, usually showing a satisfactory performance for moderate concentrations, but failing to predict critical pollution episodes (Cordova et al., 2021; Zhang et al., 2012). The current study takes a different approach, by focusing on the prediction of regional-scale air pollution events, which are typically driven by stagnant weather conditions and have important consequences to human health.

The main objective of this study is to forecast the occurrence of persistent regional-scale air quality deterioration events based on surface weather variables, using a data-driven approach. Unlike previous studies, statistical models were developed to predict the occurrence air pollution events that persist for many days, instead of predicting pollutant concentrations. An assessment of persistent air quality deterioration events in the metropolis of Sao Paulo, Brazil, is presented, showing their temporal distribution between 2005 and 2022 and associations with meteorological conditions. Another unique aspect of the current study in the consideration of species with different drivers for accumulation in the atmosphere: PM_{10}, which is mostly a primary air pollutant, and O_3, a secondary air pollutant. The models developed here require simple input data and low computational resources, aiming to facilitate a future usage by the general public and decision-makers.

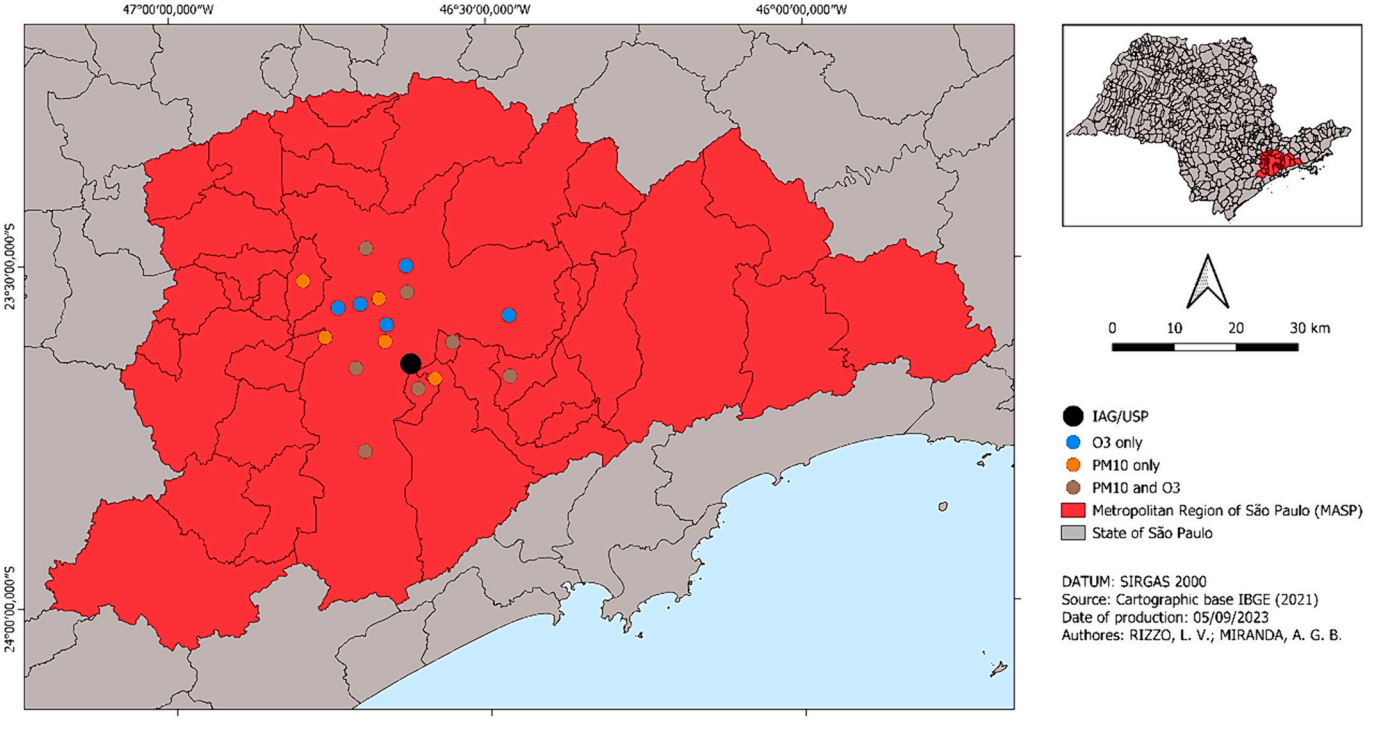


Fig. 1. Map of the MASP showing the location of the CETESB 17 air quality monitoring stations selected for this study and of the IAG/USP weather station.

Table 1
Description of variables in the dataset developed for this study. The dataset consists in daily timeseries between 2005 and 2022, comprising 6574 days.

	Abbreviation	Description	Metric	Unity	Data provider
Target variables	PM_{10} PEE	Binary variables for event and non-event days	–	–	This work
	O_3 PEE		–	–	This work
	SUM, FAL, WIN, SPR	Seasons (categorical)	SUM (DJF), FAL (MAM), WIN (JJA), SPR (SON)	–	–
Predictors	PBLH	Planetary boundary layer height	Hourly estimate at 12:00 local time	m	ERA5
	Ustar	Friction velocity	Hourly estimate at 12:00 local time	m/s	ERA5
	WS	Wind speed	Daily mean	m/s	IAG/USP
	WD	Wind direction (categorical)	Predominant direction	–	IAG/USP
	Insol	Sunshine duration	Daily integration	hours	IAG/USP
	Irrad	Global solar irradiance	Daily integration	W/m^2	IAG/USP
	Precip	Precipitation	Daily accumulation	mm/day	IAG/USP
	Press	Surface pressure	Daily mean	mbar	IAG/USP
	Tmean	Temperature	Daily mean	^oC	IAG/USP
	Tmax	Temperature	Daily maximum	^oC	IAG/USP
	RH	Relative humidity	Daily mean	%	IAG/USP

2. Materials and methods

2.1. Study area and datasets

The Metropolitan Area of Sao Paulo (MASP) is located in the Southeastern Brazil, extending through 8047 km^2 with a population of 21.9 million inhabitants and a fleet of 7.1 million vehicles (IBGE, 2023). The MASP is a global city, acting as an important hub in the global system of finance and trade (Parnreiter, 2019). Its vehicle fleet is singular because of the blend of gasoline and ethanol used as fuel, with singular impacts on the atmospheric chemistry and on the formation of secondary pollutants like O_3 (Salvo and Geiger, 2014). Climate is classified as subtropical humid (Cwa in Koppen's classification) (Alvares et al., 2013), characterized by a wet summer between December and February (1991–2020 climatological precipitation in the range 231–292 mm/month) and by a dry winter between June and August (32–60 mm/month) (INMET, 2023).

Data from 17 air quality monitoring stations in the MASP were used in this study (Fig. 1). The stations were selected based on the data coverage between 2005 and 2022 and on the localization of the stations, aiming to represent air quality conditions in the whole MASP. Table S1 in the supplementary material lists the air quality monitoring stations used in the current study. Hourly data on the concentration of PM_{10} and O_3 were acquired through the Sao Paulo State environmental agency website (CETESB, 2023). Moving averages of 24 h and 8 h were calculated for PM_{10} and O_3, respectively. The daily maximum moving average concentration was used to identify days with exceedance of the World Health Organization standards: 45 $\mu g.m^{-3}$ for PM_{10} and 100 $\mu g.m^{-3}$ for O_3 (WHO, 2021). The resulting daily dataset on exceedances in each air quality monitoring station was used to identify persistent exceedance events (PEE), as will be described in the following section.

Surface weather variables were used as predictors in the models. Data from the weather station of the IAG/USP (World Meteorological Organization station # 83,004) was used in this study from 2005 to 2022. Daily statistics on the following surface weather variables were used: temperature relative humidity, surface pressure, wind speed and direction, global solar irradiation, sunshine duration and precipitation (Table *1*). Daily dominant wind direction data was provided in 17 categories, including the 4 cardinal directions, the 4 intercardinal directions, 8 subdivisions and 1 category for calm conditions, when the wind direction is undetermined. The IAG/USP weather station has been characterized in previous studies and is representative of the MASP weather conditions, at least concerning measurements of temperature, precipitation and relative humidity (Sugahara et al., 2012; Yamasoe et al., 2021). Since atmospheric stability is an important driver for the accumulation of atmospheric pollutants, two turbulence-related variables were included in this study: the planetary boundary layer height (PBLH) and the friction velocity (Ustar), acquired from the ECWMF/ERA5 global reanalysis (European Centre for Medium-Range Weather Forecasts), with a spatial resolution of 0.25^o (Hersbach et al., 2020). The daily data retrieved from ERA5 correspond to instant values at 12:00 local time interpolated from the original grid to the study region (23.45–23.70 S; 46.40–46.80 W). We opted for the reanalysis data due to the absence of a daily data product based on in situ observations covering the whole study period.

2.2. Definition of persistent exceedance events (PEE)

Periods of persistent air quality deterioration were identified between 2005 and 2022 in the MASP using the method proposed by Oliveira et al. (2022). The so-called persistent exceedance events (PEE) differ from ordinary exceedance counts because they occur simultaneously at multiple air quality monitoring stations and persist for many days. Therefore, PEE pertain to the regional scale, potentially impacting the health of millions of inhabitants in the metropolis. The daily timeseries on exceedances of PM_{10} and O_3 in 17 air quality monitoring stations were used to identify the occurrence of PEE based on the following criteria:

i. Exceedance in >50% of the air quality monitoring stations;
ii. Exceedances lasting for at least 5 consecutive days for PM_{10} and 3 consecutive days for O_3.

Data was processed using scripts developed in R programming language. Compared to the original dataset on PEE provided by Oliveira et al. (2022), the current study provides the following updates and improvements: inclusion of data from 7 additional air quality monitoring stations in the MASP; extension of the study period from 2005 to 2017 to 2005–2022; update of the WHO air quality standards, released in the year of 2021.

The application of the criteria described above resulted in a timeseries of binary values showing the occurrence (1) or absence (0) of PEE on a particular day, separately for PM_{10} and O_3. Those timeseries were used to train and test models aiming to predict the occurrence of PEE, as will be presented in the next section. Also, statistics were provided for the duration of the PM_{10} and O_3 events, as well as the interannual and seasonal frequency of occurrence.

2.3. Logistic models

Multivariate classification models were developed aiming to predict the occurrence of persistent exceedance events (target variable) based on surface weather conditions (predictors). In this study, logistic models were developed using weather variables as predictors, obtaining the probability p of a given day to be part of a persistent exceedance event. Logistic models are defined by the following equation:

$$ln\left(\frac{p}{1-p}\right) = b_0 + b_1x_1 + \ldots + b_nx_n \tag{1}$$

where p is the probability of occurrence, x_i are the predictor variables and b_i are the coefficients associated with each predictor, obtained by fitting the model to the observed data. The exponential of a coefficient b_i provides the odds ratio (OR_i) of an event day for a unit increase in the correspondent predictor x_i, given that the other predictors are fixed:

$$OR_i = e^{b_i} \tag{2}$$

The odds ratio value is explanatory, since it measures the effect of a given predictor on the increase (OR > 1) or decrease (OR < 1) of PEE odds. For more details, refer to Section 3 in the supplementary material.

The input data initially consisted in a timeseries containing 6163 lines, corresponding to each day in the period 2005–2021, excluding the missing data (<1%). The dataset was randomly divided into train (80%) and test (20%) subsets. The models were built based on the train subset, under a supervised learning framework. In addition to the test subset, data from the year of 2022 was taken aside for the final model evaluation. After obtaining the model coefficients, Eq. (1) was applied to the remaining 20% of the data (test set), providing the probability of a particular day to be classified as an event-day. Whenever the calculated probability was above the cutoff (refer to Section 3 in the supplementary material), that particular day was classified as an event-day. The model performance was evaluated by comparing the classification provided by the model against the observed data. Three metrics were used to evaluate the model performance. The sensitivity of a prediction measures how able the model was to detect the positive events, being defined as:

$$Sens = \frac{TP}{TP + FN} \tag{3}$$

The specificity of a prediction measures the ability of the model to correctly predict the negative events, defined as:

$$Spec = \frac{TN}{TN + FP} \tag{4}$$

Finally, accuracy is defined as the percentual ratio of correct responses (i.e., TP + TN):

$$A = \frac{TP + TN}{N} \tag{5}$$

where N is the number of days in the subset.

The dataset contained 13 variables, including the target variable (occurrence of PM_{10} PEE) followed by the predictors: 10 numerical weather variables and 2 categorical variables for wind direction and season (Table 1). A similar dataset was built for O_3 PEE. Numerical variables were normalized by subtracting the mean and dividing by the standard deviation.

Different configurations were tested in the development of the predictive models for PM_{10} and O_3 events. The choice of the best set of predictors was based on the stepwise feature selection method combined with a multicollinearity analysis using the variance inflation factor (VIF), adopting the Akaike Information Criterion (AIC). The final version of the models included only the statistically significant predictors ($p < 0.05$) after the removal of multicollinearity. Tests were also performed with subsets of the data to evaluate the impact of imbalance on model performance. The dataset is imbalanced because the positive events were relatively rare during the study period, with a prevalence of 15% for PM_{10} event days and 5% for O_3 event days. Four different model setups with different subsets of the original dataset were tested for each pollutant (Tables S2 and S3). The first model setup contained data from all seasons, with the prevalence values reported above. The following model setups used subsets of the data in order to artificially increase the prevalence of PEE (refer to Section 1 in the supplementary material). In the results section, only the best model configurations will be

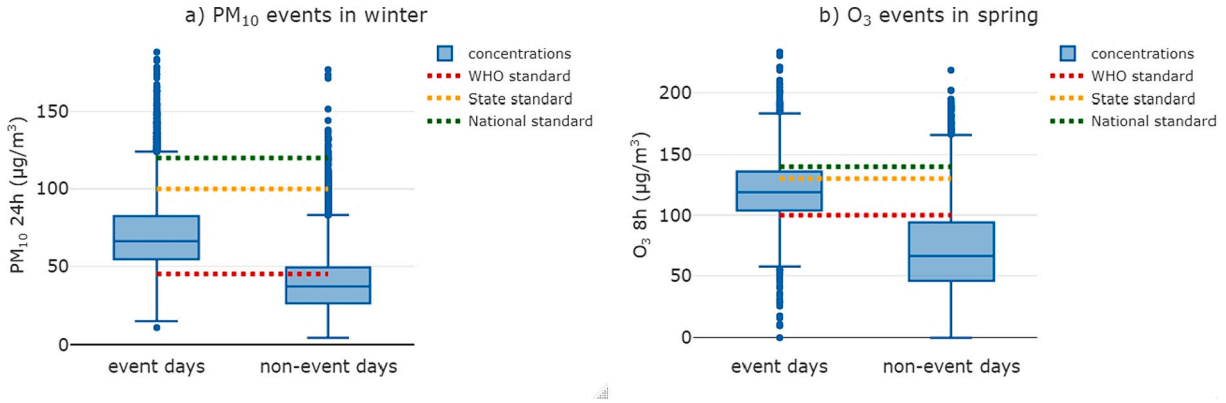


Fig. 2. Distribution of concentrations in event and non-event days for a) PM_{10} daily maximum of 24 h moving average in the winter b) O_3 daily maximum of 8 h moving average in the spring. The boxplots contain data from 17 air quality stations between 2005 and 2022. Horizontal lines show the concentration standards recommended by the WHO and legislated by Brazilian governmental agencies in national and state levels.

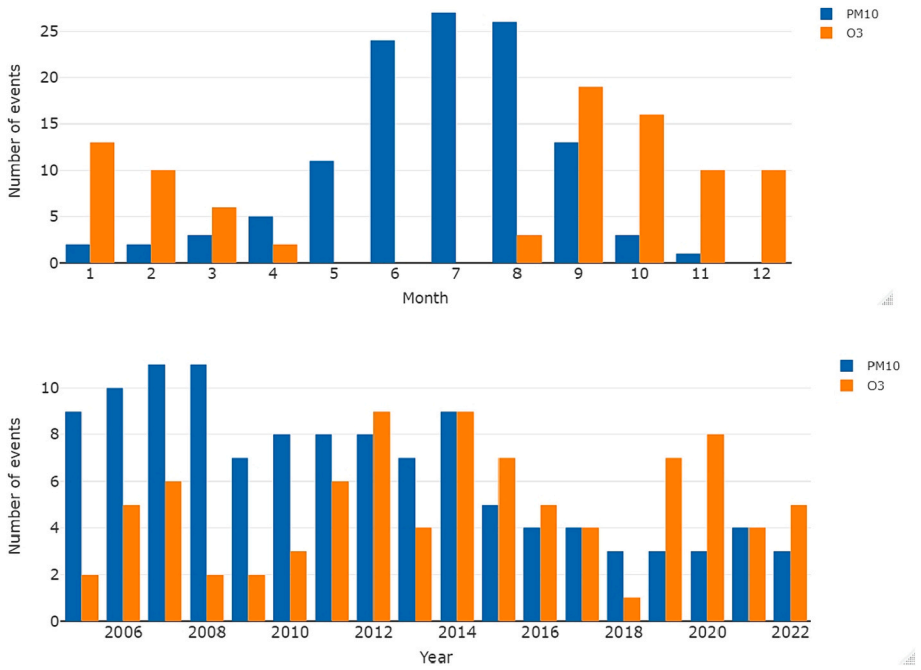


Fig. 3. Distribution of the number of persistent exceedance events for PM_{10} and O_3 grouped by month (top) and along the years (bottom) during the period of study (2005–2022).

presented.

3. Results and discussion

3.1. Characterization of persistent exceedance events of PM_{10} and O_3

Between 2005 and 2022, 117 PEE were detected for PM_{10} and 89 for O_3. PM_{10} events lasted longer, persisting along 5 to 24 consecutive days, while O_3 events lasted between 3 and 13 consecutive days (Fig. S3). On average, there were 4 PM_{10} events/year in the MASP, comprising 55 event days per year. In the case of O_3, 4 events/year were observed, comprising 18 event days per year. In event days, median PM_{10} and O_3 concentrations were approximately 80% higher compared to non-event days. Tables S6 and S7 in the supplementary material show the complete list of events. Fig. 2 shows the distribution of the daily maximum of moving average concentrations of PM_{10} and O_3, considering data from 17 air quality monitoring stations. In event days during the peak season, >75% of concentration records exceeded the air quality standards preconized by WHO. Brazilian air quality standards, both in National and

Table 2
Metrics for evaluation of the model performance for the prediction of PM_{10} and O_3 PEE for the train set (supervised learning), for the test set (initial evaluation) and for a subset including data only for the year of 2022 (final evaluation).

	PM_{10} PEE			O_3 PEE		
	Train set	Test set	2022 set	Train set	Test set	2022 set
Accuracy	0.82	0.81	0.79	0.90	0.91	0.92
Sensitivity	0.88	0.84	1.00	0.81	0.85	0.81
Specificity	0.81	0.80	0.79	0.91	0.91	0.92
AUC	0.91	–	–	0.94	–	–
Cutoff probability	0.136	–	–	0.128	–	–

State levels, are less restrict compared to WHO's. As a consequence, the official statistics on air pollution exceedances may not assure air quality safety conditions, considering the WHO guidelines. Please note that the classification as a non-event day does not guarantee clean air quality conditions in the whole MASP. Localized and short-lived air pollution episodes can still occur in non-event days.

PM_{10} events were more frequent in the austral winter (June–August), despite they occurred all through the year (Fig. 3). On the other hand, O_3 events predominated in the austral spring (September–November) and were rarely observed during the winter. The seasonal behavior of events is in accordance with the typical variability of PM_{10} and O_3 concentrations in the MASP. During the austral winter, the Brazilian southeast region, including the MASP, is affected by the climatological zonal expansion of the South Atlantic Subtropical Anticyclone (SASA), favoring the occurrence of dry and stable atmospheric conditions (Reboita et al., 2019). The winter dry conditions winter hinder air pollution dispersion and removal in the MASP, especially in the case of air pollutants of predominant primary origin like PM_{10}, NO_2 (nitrogen dioxide) and CO (carbon monoxide) (Carvalho et al., 2015). The occurrence of PEE of PM_{10} in the MASP is typically associated with pre-frontal conditions, characterized by subsidence, low wind speeds and dry conditions, leading to a progressive accumulation of air pollutants (Oliveira et al., 2022).

In the case of O_3, the stable atmospheric conditions typical of the winter were not sufficient to trigger the occurrence of persistent exceedance events, as is evident in Fig. 3 (top). O_3 is a secondary air pollutant, being photochemically produced in the atmosphere and having as precursors nitrogen oxides (NOx) and volatile organic compounds (VOCs) (Schuch et al., 2019; Targino et al., 2019). Oliveira et al. (2022) have shown that stable atmospheric conditions are necessary but not sufficient to drive the accumulation of O_3 concentrations in the regional scale. They showed that the occurrence of O_3 PEE were associated with greater anomalies and absolute values of temperature and radiation compared to the conditions observed during the PM_{10} PEE.

Hence, the atmospheric conditions that favor the occurrence of PM_{10} and O_3 events in the MASP have common features related to atmospheric stability, but distinct aspects concerning conditions for photochemical production. Fig. S4 shows significant differences in the daily maximum temperature observed during event and non-event days, with higher values in the event days of PM_{10} and O_3. Likewise, other surface weather variables showed significant differences between event and non-event days. Relative humidity was significantly lower during PM_{10} events, while irradiation was above the 3rd quartile during O_3 events (Fig. S4). It indicates that synoptic and mesoscale atmospheric conditions that favor the occurrence of regional air quality deterioration events in the MASP are captured in the variability of surface weather variables. The interplay between PEE occurrence and surface weather conditions motivated the development of predictive models for air quality deterioration events, to be discussed in the next section.

Nevertheless, it is important to point out that atmospheric conditions are not the only driver for the occurrence of PEE. The variability of air pollutant emissions also plays an important role. Fig. 3 (bottom) shows that PM_{10} events were more frequent in the past, considering the period of study. This is in accordance with the long term decreasing trend on PM_{10} concentrations in the MASP, typically ranging between -1 to -3 $\mu g.m^{-3}.y^{-1}$ since the 1990's (Carvalho et al., 2015). Long term reductions in PM_{10} and other primary pollutants concentrations in the MASP were attributed to successful programs implemented by National and State governmental agencies that gradually constrained the emission of pollutants by stationary and mobile sources (de Andrade et al., 2017). According to the emission inventory of the Sao Paulo State Environmental Agency, vehicular emissions of PM decreased from 2500 to 920 ton/year in the MASP from 2006 to 2022 (CETESB, 2023). The decreasing trend in PM_{10} concentrations was followed by the occurrence of PM_{10} PEE (Fig. 3). That is, similar weather conditions resulted in a higher frequency of PM_{10} events in the past compared to nowadays.

The distribution of persistent exceedance events for PM_{10} and O_3 along the years is clearly decoupled (Fig. 3, bottom), signalizing that different processes drive the accumulation of these pollutants in the atmosphere. While the PM_{10} events showed a significant decreasing trend, the number of O_3 events oscillated without a clear tendency. Previous studies also reported the absence of significant long term trends in the concentration of O_3 in the MASP (Schuch et al., 2019; Carvalho et al., 2015). Tropospheric O_3 photochemical production depends on weather conditions, abundancy and variety of chemical precursors (NOx and VOCs) in a non-linear way. According to Santolaya et al. (2019), about 50% of the temporal variability of O_3 concentrations in the MASP can be explained by weather conditions. Changes in precursor emissions can also affect O_3 concentrations. Previous studies reported impacts in O_3 concentrations due to a shift from ethanol to gasoline between 2010 and 2012 in the MASP, associated with changes in VOC emission patterns (Salvo and Geiger, 2014).

3.2. *Predictive models for PM_{10} and O_3 events*

Given that the behavior of surface weather variables showed a clear association with persistent air quality deterioration events in

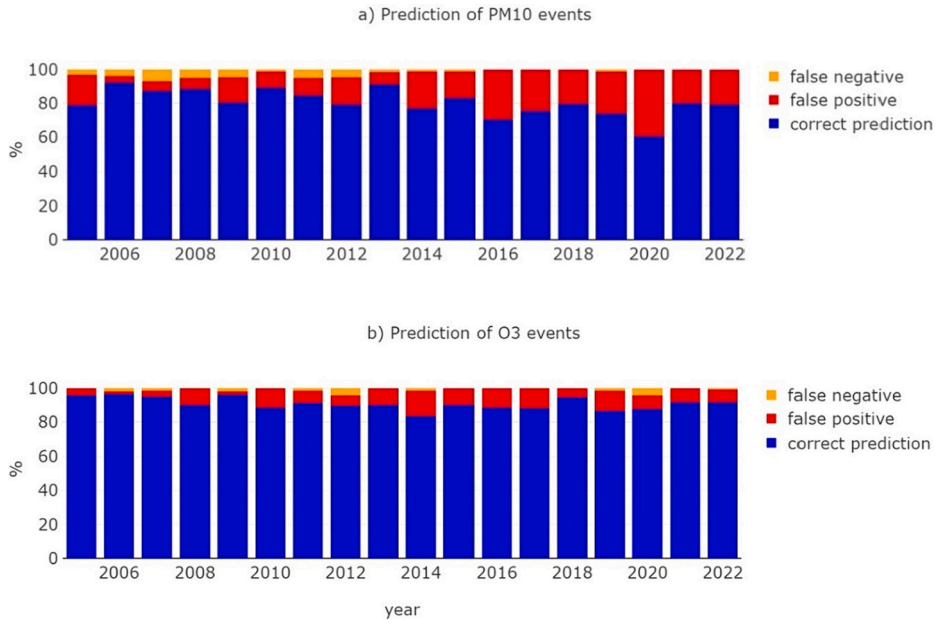


Fig. 4. Performance of model predictions for the occurrence of PM_{10} (a) and O_3 (b) events in each year, considering the test set and data from the whole year of 2022. The percentage of correct predictions correspond to the model accuracy for the identification of event and non-event days.

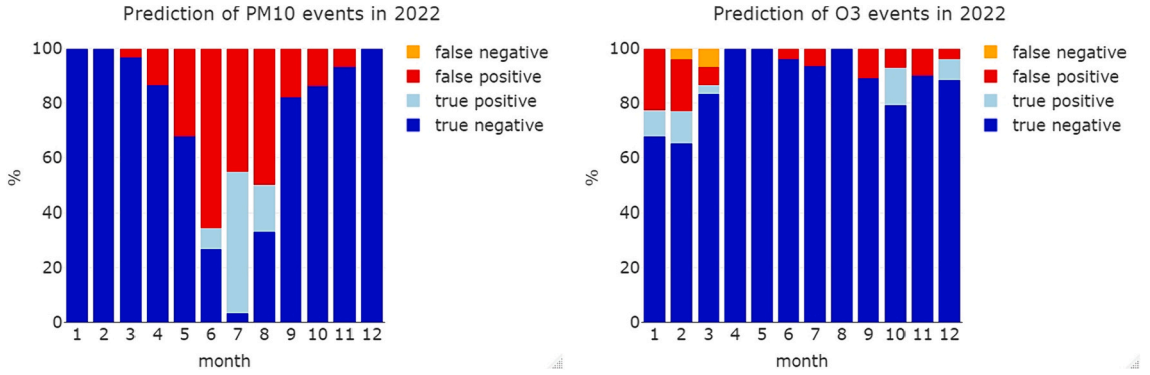


Fig. 5. Final evaluation of models for the prediction of PM_{10} (left) and O_3 (right) event days in the year of 2022.

the MASP, classification models were developed to predict their occurrence and to investigate the relative importance of predictors. Logistic models were developed separately for PM_{10} and O_3 events, initially using the same set of precursors (.

Table *1*). As explained in the methods section, different subsets of predictors and configurations were tested until reaching the final version of the models. Refer to Tables S2 and S3 for a comparison of the performance of different model configurations. The O_3 model was improved by the exclusion of data from the austral winter, when rare events were observed (Fig. *3*a). It reduced the imbalance between positive and negative event days, increasing the prevalence of O_3 event days from 5% to 8%, resulting in the highest accuracy among the tested configurations (Table S3). On the other hand, the PM_{10} model did not benefit from prevalence increase using subsets of data (Table S2), so that data from the four seasons were used in this case.

Table 2 shows metrics to evaluate the overall performance of the PM_{10} and O_3 models. The accuracy of the models (Eq. (5)) considering the test set was 0.81 and 0.91, respectively for PM_{10} and O_3 (Table 2), showing that both models had a suitable performance in the prediction of PEE. No signs of overfitting were observed, since the accuracy was similar in the train and test sets. The sensitivity was close to 0.85 in both models, indicating a satisfactory ability to correctly predict event days, despite the data imbalance. The specificity, i.e., the model skill to correctly predict non-event days, was lower for the PM_{10} model compared to the O_3 model, showing a tendency of the PM_{10} model to overestimate the number of event-days.

Fig. 4 shows the percentage of correct predictions, false positives (i.e., incorrect classification of a day as an event-day) and false negatives (i.e., incorrect classification of a day as a nonevent-day), aggregated by year, for the PM_{10} and O_3 models. Note that the percentage of false positives increased along the years for the PM_{10} model (Fig. 4a). This can be explained by the long-term decreasing trends in PM_{10} concentrations and PEE in the MASP. The model was trained with data randomly picked in the period 2005–2021. In the

Table 3
Set of predictors in the PM_{10} model, including: the standard deviation of observations (σ), the odds ratio (OR) and respective confidence interval (2.5–97.5%), the model coefficient and the corresponding *p*-value.

	Unity	σ	OR	2.5%	97.5%	Coefficient	p-value
Intercept	–	–	0.07	0.05	0.09	−2.68	$< 10^{-3}$
Fall*	–	–	0.70	0.51	0.97	−0.35	0.031
Summer*	–	–	0.12	0.06	0.20	−2.14	$< 10^{-3}$
Winter*	–	–	3.32	2.37	4.69	1.20	$< 10^{-3}$
Ustar	m/s	0.11	0.81	0.70	0.93	−0.21	0. 002
WS	m/s	1.76	0.51	0.46	0.60	−0.65	$< 10^{-3}$
Irrad	W/m^2	6.16	0.53	0.43	0.66	−0.63	$< 10^{-3}$
Precip	mm/day	11.13	0.58	0.38	0.82	−0.54	0.006
Press	mbar	3.47	1.48	1.28	1.73	0.39	$< 10^{-3}$
Tmax	°C	4.63	1.84	1.53	2.24	0.61	$< 10^{-3}$
RH	%	8.58	0.32	0.27	0.34	−1.13	$< 10^{-3}$

* In comparison to Spring.

first years of the time series, the PM_{10} events were more frequent, so that the model was biased toward positive outcomes. The thresholds on weather variables may have changed along the years in the case of PM_{10} events, so that similar weather conditions produced a different number of event days at the present time compared to the past. Changes in the PM_{10} emission patterns also affected the model performance in 2020, when PM_{10} concentrations decreased due to mobility restrictions due to the COVID-19 pandemics (Rudke et al., 2021). The O_3 model, on its turn, was less sensitive to changes in the emission patterns. The O_3 model also produced false positives, but in a smaller proportion and without a clear long-term trend (Fig. 4b).

Fig. 5 shows the performance of the models separated by month for the year of 2022. Data from 2022 was totally new for the models, since the training stage did not include any data from this year. The PM_{10} model produced many false positives, overestimating the number of event days, particularly in the winter. It can be explained by the biased training of the PM_{10} model with data that do not represent the current air quality conditions in the MASP. In the past, PM_{10} concentrations were higher, so that even moderate stagnant weather conditions were enough to trigger a PM_{10} persistent exceedance event. Nowadays, similar weather conditions may not trigger an event, since the concentrations have decreased. The overestimation was more intense in the winter, because this season is more prone to stagnant weather conditions in the MASP. Also, there were changes in in the seasonal distribution of PM_{10} events. While in the long-term PM_{10} events were detected all through the year (Fig. 3), in last 7 years (2016–2022) the events were concentrated between June and September (Fig. S6). The PM_{10} model was improved using a train dataset restricted to recent years (2016–2021), despite the decrease in the volume of data available for the model learning (refer to Section 5 in the supplementary material). There was a significant improvement in the PM_{10} model predictions, reducing the overestimation of event days during the winter (Fig. S7). Restricting the model training to recent years resulted in 87% of correct classifications into event and non-event days in 2022, with a sensitivity of 94% and a specificity of 85%, which is a promising result.

The O_3 model had a better performance compared to PM_{10}, although it also produced false positives and has missed a couple of event days (false negatives). O_3 events did not show a long-term trend (Fig. 3). As a consequence, the train dataset from the 2005–2021 period had a greater diversity of weather conditions and event outcomes, providing better conditions for model learning in the case of O_3.

Overall, the results show that regional air pollution events can be satisfactorily predicted based on surface weather variables, both for primary (PM_{10}) and secondary (O_3) air pollutants. The models were able to predict >80% of the event days in 2022. This is a notable achievement, given that regional air pollution events result from the complex interplay between weather conditions, emission patterns, atmospheric removal processes and chemistry. The current discussion also shows the influence of long-term trends (non-stationarity) on the performance of data-based models. Machine leaning classification methods could better overcome the challenges posed by the non-stationarity of PM_{10} events, if associated with training set modification and ensemble strategies (Hoens et al., 2012).

The models developed here require ordinary surface weather variables typically available in meteorological stations and weather forecast products. The models could be published online in a dashboard, assimilating weather forecast predictions in real time, providing information for the general public on the short-term probability of poor air quality conditions. As such, the simple models developed in this study could assist decision makers in taking timely actions to reduce exposition during extreme air pollution events, either through warning and advice for the exposed population or through temporary restriction of emissions. The models could also support the strategic planning of health services, since critical air pollution events may result in increased demand for care, leading to work overload and financial impacts (Elliot et al., 2016).

Among the limitations of the models, they are site-specific and cannot be generalized to other areas. Nevertheless, the methodology can be applied to other urban areas, given that relatively long time series of air pollutant concentrations and surface meteorological variables are available. The models developed here may fail in periods of atypical emission patterns, since they were built over the relationships between air pollution and weather conditions that occurred in the past.

3.3. Interpretation of model predictors

More than accurate predictions, the models developed here can provide insights about the relative importance of predictors

Table 4
Set of predictors in the O_3 model, including: the standard deviation of observations (σ), the odds ratio (OR) and respective confidence interval (2.5–97.5%), the model coefficient and the corresponding p-value.

	Unity	σ	OR	2.5%	97.5%	Coefficient	p-value
Intercept	–	–	0.0015	0.0008	0.0029	−6.48	$< 10^{-3}$
Spring*	–	–	6.81	3.87	12.51	1.92	$< 10^{-3}$
Summer*	–	–	3.05	1.71	5.66	1.11	$< 10^{-3}$
Ustar	m/s	0.11	0.52	0.40	0.67	−0.65	$< 10^{-3}$
WS	m/s	1.76	0.40	0.31	0.53	−0.90	$< 10^{-3}$
Irrad	W/m^2	6.16	1.93	1.43	2.61	0.66	$< 10^{-3}$
Press	mbar	3.47	1.50	1.15	1.96	0.40	0.003
Tmax	°C	4.63	5.83	4.06	8.50	1.76	$< 10^{-3}$
RH	%	8.58	0.52	0.41	0.64	−0.66	$< 10^{-3}$

* In comparison to Fall.

contributing to the occurrence of a regional air pollution event. The model coefficients obtained for the PM_{10} and O_3 models are shown in Table 3 and Table 4, respectively, represented as odds ratios (Eq. (2)). Season was one of the most important predictors for both models. The chance of occurrence of a PM_{10} event day was 3.32 times greater in winter compared to spring (Table 3). In the case of O_3 events, the chance was 6.81 times greater in spring compared to fall (Table 4), which is consistent with the seasonal variability of events (Fig. 3). As expected, temperature had positive effects on the occurrence of PM_{10} and O_3 events. An isolated increase in the daily maximum temperature by one standard deviation (4.63 °C) would increase the chance of occurrence of an event by 84% and 483%, respectively for PM_{10} and O_3. This result suggest that persistent air quality deterioration events can be more frequent under global warming scenarios, as reported in the literature (Lacressonnière et al., 2016).

In the case of O_3, solar irradiation also contributed to increase the chance of events. On the other hand, weather variables like relative humidity and wind speed had a negative impact on the chance of events. For example, an isolated increase in RH by one standard deviation (8.58%) would decrease the chance of occurrence of an event day by 68% and 48%, respectively for PM_{10} and O_3. Ustar, which is a proxy for atmospheric mixing and turbulence intensity, was also an important predictor, decreasing the chance of air quality deterioration events. Other predictors listed in Table 1 (Tmean, Insol, PBLH, WD) were not included in any of the models, either because of collinearity or because they did not reach statistical significance. The sign of the contribution of each weather variable is in agreement with the physical mechanisms that drive the accumulation of pollutants in the lower atmosphere (Oliveira et al., 2022; Sánchez-Ccoyllo and Andrade, 2002).

4. Conclusion

Persistent air quality deterioration events are relatively frequent in the subtropical megacity of Sao Paulo, with a prevalence of 55 PM_{10} event days per year and 18 O_3 event days per year, on average. Short-term forecast models were developed to predict the occurrence of PM_{10} and O_3 persistent exceedance events. Results show that regional air pollution events can be satisfactorily predicted based on a simple set of surface weather variables, both for primary (PM_{10}) and secondary (O_3) air pollutants. The models were able to predict >80% of the event days in 2022. This is a notable achievement, given that extreme concentrations are relative rare and that regional air pollution events result from the complex interplay between weather conditions, emission patterns, and chemistry. Given the decreasing long-term trend in the number of PM_{10} events, restriction of the training data to recent years significantly improved the performance of the PM_{10} model. Daily maximum temperature was an important predictor, such that an isolated increase of 4 °C increased the chance of events by 84% and 483%, respectively for PM_{10} and O_3. This result suggests that extreme air pollution events could become more frequent under global warming if the current emission patterns continue. The short-term forecast models developed here are simple to use and do not require expertise in atmospheric sciences. If available online, the models could assist decision makers in taking timely actions to reduce exposition during extreme air pollution events, either through warning and advice for the exposed population, temporary restriction of emissions, or planning health care services.

Funding

This work was supported by the Sao Paulo State Research Foundation (FAPESP) [grant number 2021/14342–7] and by the Brazilian National Council for Scientific and Technological Development (CNPq) [grant numbers 405179/2021–9 and 304819/2022–0].

CRediT authorship contribution statement

L.V. Rizzo: Writing – original draft, Supervision, Software, Methodology, Funding acquisition, Conceptualization. **A. G. B. Miranda:** Writing - Original Draft, Software, Formal analysis, Visualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to

influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

The authors would like to thank the Sao Paulo State Environmental Agency (CETESB) and the Weather Station of the Institute of Astronomy, Geophysics and Atmospheric Science of the University of São Paulo for making available the air quality and meteorological observations.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.uclim.2024.101876.

References

Alvares, C.A., Stape, J.L., Sentelhas, P.C., De Moraes Gonçalves, J.L., Sparovek, G., 2013. Köppen's climate classification map for Brazil. Meteorol. Z. 22, 711–728. https://doi.org/10.1127/0941-2948/2013/0507.

Anderson, H.R., 2009. Air pollution and mortality: A history. Atmos. Environ. 43, 142–152. https://doi.org/10.1016/j.atmosenv.2008.09.026.

Brook, R.D., Rajagopalan, S., Pope, C.A., Brook, J.R., Bhatnagar, A., Diez-Roux, A.V., Holguin, F., Hong, Y., Luepker, R.V., Mittleman, M.A., Peters, A., Siscovick, D., Smith, S.C., Whitsel, L., Kaufman, J.D., 2010. Particulate matter air pollution and cardiovascular disease: an update to the scientific statement from the american heart association. Circulation 121, 2331–2378. https://doi.org/10.1161/CIR.0b013e3181dbece1.

Carvalho, V.S.B., Freitas, E.D., Martins, L.D., Martins, J.A., Mazzoli, C.R., de Andrade, M.F., 2015. Air quality status and trends over the metropolitan area of São Paulo, Brazil as a result of emission control policies. Environ. Sci. Pol. 47, 68–79. https://doi.org/10.1016/j.envsci.2014.11.001.

CETESB, 2023. https://cetesb.sp.gov.br/ar/qualar/ last access: 1 June.

Chen, J., Li, C., Ristovski, Z., Milic, A., Gu, Y., Islam, M.S., Wang, S., Hao, J., Zhang, H., He, C., Guo, H., Fu, H., Miljevic, B., Morawska, L., Thai, P., Lam, Y.F., Pereira, G., Ding, A., Huang, X., Dumka, U.C., 2017. A review of biomass burning: emissions and impacts on air quality, health and climate in China. Sci. Total Environ. 579, 1000–1034. https://doi.org/10.1016/j.scitotenv.2016.11.025.

Cohen, A.J., Brauer, M., Burnett, R., Anderson, H.R., Frostad, J., Estep, K., Balakrishnan, K., Brunekreef, B., Morawska, L., Iii, C.A.P., Shin, H., Straif, K., Shaddick, G., Thomas, M., Van Dingenen, R., Van Donkelaar, A., Vos, T., Murray, C.J.L., Forouzanfar, M.H., 2017. Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution : an analysis of data from the global burden of diseases study 2015. Lancet 6736, 1–12. https://doi.org/10.1016/S0140-6736(17)30505-6.

Cordova, C.H., Portocarrero, M.N.L., Salas, R., Torres, R., Rodrigues, P.C., López-Gonzales, J.L., 2021. Air quality assessment and pollution forecasting using artificial neural networks in metropolitan Lima-Peru. Sci. Rep. 11, 24232. https://doi.org/10.1038/s41598-021-03650-9.

de Andrade, M.F., Kumar, P., de Freitas, E.D., Ynoue, R.Y., Martins, J., Martins, L.D., Nogueira, T., Perez-Martinez, P., de Miranda, R.M., Albuquerque, T., Gonçalves, F.L.T., Oyama, B., Zhang, Y., 2017. Air quality in the megacity of São Paulo: evolution over the last 30 years and future perspectives. Atmos. Environ. 159, 66–82. https://doi.org/10.1016/j.atmosenv.2017.03.051.

Elliot, A.J., Smith, S., Dobney, A., Thornes, J., Smith, G.E., Vardoulakis, S., 2016. Monitoring the effect of air pollution episodes on health care consultations and ambulance call-outs in England during march/April 2014: A retrospective observational analysis. Environ. Pollut. 214, 903–911. https://doi.org/10.1016/j.envpol.2016.04.026.

Gómez Peláez, L.M., Santos, J.M., de Almeida Albuquerque, T.T., Reis, N.C., Andreão, W.L., de Fátima Andrade, M., 2020. Air quality status and trends over large cities in South America. Environ. Sci. Pol. 114, 422–435. https://doi.org/10.1016/j.envsci.2020.09.009.

Gurjar, B.R., Butler, T.M., Lawrence, M.G., Lelieveld, J., 2008. Evaluation of emissions and air quality in megacities. Atmos. Environ. 42, 1593–1606. https://doi.org/10.1016/j.atmosenv.2007.10.048.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R.J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., Thépaut, J.N., 2020. The ERA5 global reanalysis. Q. J. R. Meteorol. Soc. 146, 1999–2049. https://doi.org/10.1002/qj.3803.

Hoek, G., Krishnan, R.M., Beelen, R., Peters, A., Ostro, B., Brunekreef, B., Kaufman, J.D., 2013. Long-term air pollution exposure and cardio- respiratory mortality: a review. Environ. Health 12, 43. https://doi.org/10.1186/1476-069X-12-43.

Hoens, T.R., Polikar, R., Chawla, N.V., 2012. Learning from Streaming Data with Concept Drift and Imbalance: An Overview. https://doi.org/10.1007/s13748-011-0008-0, 1 April.

IBGE, 2023. https://cidades.ibge.gov.br/ last access: 30 April.

INMET, 2023. https://portal.inmet.gov.br/normais# last access: 4 September.

Kumar, P., de Fatima Andrade, M., Ynoue, R.Y., Fornaro, A., de Freitas, E.D., Martins, J., Martins, L.D., Albuquerque, T., Zhang, Y., Morawska, L., 2016. New directions: from biofuels to wood stoves: the modern and ancient air quality challenges in the megacity of São Paulo. Atmos. Environ. 140, 364–369. https://doi.org/10.1016/j.atmosenv.2016.05.059.

Lacressonnière, G., Foret, G., Beekmann, M., Siour, G., Engardt, M., Gauss, M., Watson, L., Andersson, C., Colette, A., Josse, B., Marécal, V., Nyiri, A., Vautard, R., 2016. Impacts of regional climate change on air quality projections and associated uncertainties. Clim. Chang. 136, 309–324. https://doi.org/10.1007/s10584-016-1619-z.

Li, Y., Guo, J.E., Sun, S., Li, J., Wang, S., Zhang, C., 1 March 2022. Air Quality Forecasting with Artificial Intelligence Techniques: A Scientometric and Content Analysis. https://doi.org/10.1016/j.envsoft.2022.105329.

Longo, K.M., Freitas, S.R., Pirre, M., Marécal, V., Rodrigues, L.F., Panetta, J., Alonso, M.F., Rosário, N.E., Moreira, D.S., Gácita, M.S., Arteta, J., Fonseca, R., Stockler, R., Katsurayama, D.M., Fazenda, A., Bela, M., 2013. The chemistry CATT-BRAMS model (CCATT-BRAMS 4.5): a regional atmospheric model system for integrated air quality and weather forecasting and research. Geosci. Model Dev. 6, 1389–1405. https://doi.org/10.5194/gmd-6-1389-2013.

Martins, L.D., Hallak, R., Alves, R.C., de Almeida, D.S., Squizzato, R., Moreira, C.A.B., Beal, A., da Silva, I., Rudke, A., Martins, J.A., 2018. Long-range transport of aerosols from biomass burning over southeastern South America and their implications on air quality. Aerosol Air Qual. Res. 18, 1734–1745. https://doi.org/10.4209/aaqr.2017.11.0545.

Molina, L.T., Zhu, T., Wan, W., Gurjar, B., 2020. R.: impacts of megacities on air quality: challenges and opportunities. In: Oxford Research Encyclopedia of Environmental Science. Oxford University Press. https://doi.org/10.1093/acrefore/9780199389414.013.5.
Ning, G., Yim, S.H.L., Wang, S., Duan, B., Nie, C., Yang, X., Wang, J., Shang, K., 2019. Synergistic effects of synoptic weather patterns and topography on air quality: a case of the Sichuan Basin of China. Clim. Dyn. 53, 6729–6744. https://doi.org/10.1007/s00382-019-04954-3.
Oliveira, M.C.Q.D., Drumond, A., Rizzo, L.V., 2022. Air pollution persistent exceedance events in the Brazilian metropolis of Sao Paulo and associated surface weather patterns. Int. J. Environ. Sci. Technol. 19, 9495–9506. https://doi.org/10.1007/s13762-021-03778-1.
Parnreiter, C., 2019. Global cities and the geographical transfer of value. Urban Stud. 56, 81–96. https://doi.org/10.1177/0042098017722739.
Parrish, D.D., Singh, H.B., Molina, L., Madronich, S., 2011. Air quality progress in north American megacities: A review. Atmos. Environ. 45, 7015–7025. https://doi.org/10.1016/j.atmosenv.2011.09.039.
Porter, W.C., Heald, C.L., Cooley, D., Russell, B., 2015. Investigating the observed sensitivities of air-quality extremes to meteorological drivers via quantile regression. Atmos. Chem. Phys. 15, 10349–10366. https://doi.org/10.5194/acp-15-10349-2015.
Reboita, M.S., Ambrizzi, T., Silva, B.A., Pinheiro, R.F., da Rocha, R.P., 2019. The South Atlantic subtropical anticyclone: present and future climate. Front. Earth Sci. 7, 1–15. https://doi.org/10.3389/feart.2019.00008.
Ribeiro, F.N.D., Oliveira, A.P.D., Soares, J., Miranda, R.M.D., Barlage, M., Chen, F., 2018. Effect of sea breeze propagation on the urban boundary layer of the metropolitan region of Sao Paulo, Brazil. Atmos. Res. 214, 174–188. https://doi.org/10.1016/j.atmosres.2018.07.015.
Rudke, A.P., Martins, J.A., de Almeida, D.S., Martins, L.D., Beal, A., Hallak, R., Freitas, E.D., Andrade, M.F., Foroutan, H., Baek, B.H., De, T.T., 2021. How mobility restrictions policy and atmospheric conditions impacted air quality in the state of São Paulo during the COVID-19 outbreak. Environ. Res. 198 https://doi.org/10.1016/j.envres.2021.111255.
Salvo, A., Geiger, F.M., 2014. Reduction in local ozone levels in urban São Paulo due to a shift from ethanol to gasoline use. Nat. Geosci. 7, 450–458. https://doi.org/10.1038/NGEO2144.
Sánchez-Ccoyllo, O.R., Andrade, M.F., 2002. The Influence of Meteorological Conditions on the Behavior of Pollutants Concentrations in São Paulo, Brazil. Environ. Pollut. vol. 116, 257–263.
Santolaya, C., Oliveira, M.C.Q.D., Rizzo, L.V., Miraglia, S.G.E.K., 2019. Contribution of chemical and meteorological factors to tropospheric ozone formation in Sao Paulo. Rev. Bras. Ciências Ambient. 90–104 https://doi.org/10.5327/Z2176-947820190577.
Schuch, D., de Freitas, E.D., Espinosa, S.I., Martins, L.D., Carvalho, V.S.B., Ramin, B.F., Silva, J.S., Martins, J.A., de Fatima Andrade, M., 2019. A two decades study on ozone variability and trend over the main urban areas of the São Paulo state, Brazil. Environ. Sci. Pollut. Res. 26, 31699–31716. https://doi.org/10.1007/s11356-019-06200-z.
Singh, V., Singh, S., Biswal, A., 2021. Exceedances and trends of particulate matter (PM2.5) in five Indian megacities. Sci. Total Environ. 750, 141461 https://doi.org/10.1016/j.scitotenv.2020.141461.
Sugahara, S., da Rocha, R.P., Ynoue, R.Y., da Silveira, R.B., 2012. Homogeneity assessment of a station climate series (1933-2005) in the metropolitan area of São Paulo: instruments change and urbanization effects. Theor. Appl. Climatol. 107, 361–374. https://doi.org/10.1007/s00704-011-0485-x.
Szyszkowicz, M., Kousha, T., Castner, J., Dales, R., 2018. Air pollution and emergency department visits for respiratory diseases: A multi-city case crossover study. Environ. Res. 163, 263–269. https://doi.org/10.1016/j.envres.2018.01.043.
Tai, A.P.K., Mickley, L.J., Jacob, D.J., 2010. Correlations between fine particulate matter (PM2.5) and meteorological variables in the United States: implications for the sensitivity of PM2.5 to climate change. Atmos. Environ. 44, 3976–3984. https://doi.org/10.1016/j.atmosenv.2010.06.060.
Targino, A.C., Harrison, R.M., Krecl, P., Glantz, P., de Lima, C.H., Beddows, D., 2019. Surface ozone climatology of south eastern Brazil and the impact of biomass burning events. J. Environ. Manag. 252, 1–12. https://doi.org/10.1016/j.jenvman.2019.109645.
WHO, 2021. Executive summary. In: World HEALTH Organization Global Air Quality Guidelines: Particulate Matter (PM2.5 and PM10), Ozone, Nitrogen Dioxide, Sulfur Dioxide and Carbon Monoxide.
Yamasoe, M.A., Rosário, N.M.É., Almeida, S.N.S.M., Wild, M., 2021. Fifty-six years of surface solar radiation and sunshine duration over São Paulo, Brazil: 1961–2016. Atmos. Chem. Phys. 21, 6593–6603. https://doi.org/10.5194/acp-21-6593-2021.
Zhang, Y., Bocquet, M., Mallet, V., Seigneur, C., Baklanov, A., December 2012. Real-Time Air Quality Forecasting, Part I: History, Techniques, and Current Status. https://doi.org/10.1016/j.atmosenv.2012.06.031.
Zhang, Z., Wang, J., Chen, L., Chen, X., Sun, G., Zhong, N., Kan, H., Lu, W., 2014. Impact of haze and air pollution-related hazards on hospital admissions in Guangzhou, China. Environ. Sci. Pollut. Res. 21, 4236–4244. https://doi.org/10.1007/s11356-013-2374-6.
Zhang, Y., Ding, A., Mao, H., Nie, W., Zhou, D., Liu, L., Huang, X., Fu, C., 2016. Impact of synoptic weather patterns and inter-decadal climate variability on air quality in the North China plain during 1980–2013. Atmos. Environ. 124, 119–128. https://doi.org/10.1016/j.atmosenv.2015.05.063.