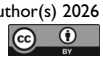
Check for updates

Article

Journal of Ambient Intelligence and Smart Environments
2026, Vol. 18(2) 234–248
© The Author(s) 2026

Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/18761364251396640
journals.sagepub.com/home/ais

Mobility effect on city pollution: A case study

Alexandre Juma[1], Luis B Elvas[1,2,3,4,5], Joao C Ferreira[1,2,3,4] and Luis Nunes[1,2]

Abstract

The present work reports the impacts on urban mobility and air quality in Lisbon, Portugal, of the imposed restrictions to curb the transmission of SARS-CoV-2 virus, which causes COVID-19 disease. We performed a data-driven approach over Lisbon Smart cities data, collected from several sources, such as traffic and pollution. During the first Portuguese emergency period (18-03-2020 to 03-05-2020) the sharp reductions in anthropogenic activities, most importantly road traffic, resulted in generally reduced criteria air pollutant concentration compared to an homologous baseline from 2013–2019 measured in the six air quality monitoring stations throughout the city. The most negatively impacted air pollutants were NO2, with a reduction of 54.35% in traffic stations and 28.62% in background stations. Google mobility indicator for local commerce was found to be the main anthropogenic activity indicator for Lisbon, with a moderate and positive correlation with NO_2 concentration (r=+0.54). A regressor ML pipeline was trained to predict NO_2 concentration with the available anthropogenic activity, weather, and air pollutant inputs from March/2020 to March/2021, achieving $R^2 = 0.925$ on the test set.

Keywords

urban mobility, air quality, covid-19, data mining, data analytics, machine learning, pollution

Received: 7 February 2025; accepted: 29 October 2025

1 Introduction

In April 2020, the month with the most severe levels of confinement, the consumption of liquid fuels in Portugal suffered unprecedented homologous reductions (Sugawara et al., 2021), −58.6% for Gasoline and −47.0% for Diesel, and electrical energy consumption had a homologous drop of −13.8%, whereas 69% of the production was of renewable sources, 14% was imported and the remaining 17% of non-renewable sources (REN - Consumo de energia elétrica recua 12% em abril). Changes in economic activity and telework promoted drastic changes in energy consumption (Hook et al., Aug. 2020), namely domestic energy consumption, which in April 2020 had a homologous increase of 31%, while industry energy consumption dropped 17% and services dropped 43%. In April 2020, the maritime port of Lisbon registered a decrease of 47.70% in ships docked (Estatísticas," Statistics Port of Lisbon), and Lisbon International Airport registered a drop in 95.88% of aircraft movements (Passenger traffic at the Lisbon Airport).

Throughout the most severe lockdown periods, government agencies, environmental associations and the press reported exceptional improvement of air quality in Lisbon and the cause was associated with the sharp reduction of anthropogenic pollutant activities. It is of utmost importance to study the phenomena of air pollution response to the variation of

[1]Instituto Universitário de Lisboa (ISCTE-IUL), ISTAR, Lisboa, Portugal
[2]ISTAR, Centro de Investigação em Ciências da Informação, Tecnologias e Arquitetura, Lisboa, Portugal
[3]Faculty of Logistics, Molde University College, Molde, Norway
[4]Inov Inesc Inovação – Instituto de Novas Tecnologias, Lisbon, Portugal
[5]Breast Cancer Research Program, Champalimaud Foundation, Lisbon, Portugal

Corresponding author:
Luis Nunes, Instituto Universitário de Lisboa (ISCTE-IUL), ISTAR, Lisboa, Portugal; ISTAR, Centro de Investigação em Ciências da Informação, Tecnologias e Arquitetura, Lisboa, Portugal.
Email: luis.nunes@iscte-iul.pt

anthropogenic activities to better understand and quantify its causal relation and support the decision process related to environmental policies.

In Portugal, in 2019, of all registered vehicles (Motorised road vehicles (No.) by Vehicle Type and Fuel Type), diesel engines account for 65.47%, gasoline engines for 32.46%, LPG engines for 0.83% and 1,24% for other motorization types where all types of electric vehicles are included (Associação de Utilizadores de Veículos Elétricos (UVE), Jan. 5, 2023). Moreover, in 2019, 62.02% of the Portuguese vehicle fleet is over 10 years (Motorised road vehicles (No.) by Vehicle Type and Fuel Type), whose engines implement worse European emission standards than newer engines [Commission Regulation (EU)].

The 2017 Mobility Survey for the Lisbon Metropolitan Area [Mobility and functionality of the territory in the Metropolitan Areas of Porto and Lisbon, 2017, p. 2017] reports that 60.8% of all trips are done using private passenger vehicles (cars and motorbikes), and the occupation ratio of passenger vehicles is 1,60 persons. On the other hand, only 15.8% of trips are done using public transportation and 23.5% of trips are done with soft transportation means from which only 0.5% of trips are done using bicycles.

The low electrification of the Portuguese vehicle fleet, the large share of diesel engines and older vehicles (pre-EURO 5/6), high urban road congestion levels and low adoption of public transport, make urban air pollution concentrations, namely NO_x and particulate matter (PM) pollutants, in an urban setting such as Lisbon a tough phenomenon to tackle (Hooftman et al., Feb. 2016).

In the Lisbon and Tagus Valley area, the main contributing anthropogenic sources of primary air pollutants are predominantly road transport vehicles (Coordenação et al., 2019) with NO_x contribution of 63%, Carbon Monoxide (CO) contribution of 78% and PM_{10} contribution of 62%, as well as the minor contribution from other forms of transportation like air and sea transport, which were all gravely impacted by restrictive mobility measures to control the COVID-19 pandemic.

Air pollution in urban settings is a major source of concern due to adverse effects on human health, leading to increased respiratory and cardiovascular disease development and premature death (Ghorani-Azam et al., Sept. 2016). Other impacts of air pollution include increases in sick leaves for workers and students, lower standards of living for vulnerable groups, such as asthmatic and elderly citizens, as well as an added cost burden to public health systems. WHO estimates 4.2 million yearly deaths were related to air pollution in 2016 (W. H. Organization, 2021), EEA estimates that 6.690 premature deaths in Portugal (2018) are caused due to three main air pollutants (NO_2, $PM_{2.5}$ and O_2) (E.E.A., 2018) and an 858 city study in Europe for the year of 2015 estimates that 1.837 premature yearly deaths in Lisbon Metropolitan Area are related to two main air pollutants (NO_2, $PM_{2.5}$) (Khomenko et al., 2021), ranking 116th worst position for NO_2 related premature deaths and 514th worst position for $PM_{2.5}$ related premature deaths.

The COVID-19 pandemic management by the Portuguese government and local Lisbon Metropolitan Area authorities included highly restrictive measures that severely reduced urban mobility in Lisbon for all transportation forms and for a large period of time. It is a once in a lifetime opportunity to measure and study the effects of reduced pollutant anthropogenic activities in an urban setting, which were previously only possible to simulate using atmospheric chemistry simulation techniques, in order to aid the definition and prioritization of air pollution reduction policies.

That said, we will use sensor data from the city and correlate mobility and pollution in order to perform a complete data analysis process. Once we understand the relationships between the variables, we will apply a machine-learning algorithm to our data to predict the concentration of NO_2.

City sensors allow measuring traffic and air quality. This data can be integrated with other sources to produce knowledge and check correlations among mobility patterns and pollution in a city.

2 Literature review

In order to search existing relevant scholarly literature, Google Scholar was used. Taking into consideration that Google Scholar is not a scholarly literature database, but rather an academic search engine, it does not have an editorial review board, and content quality is solely evaluated by means of specialized algorithms, so special care was taken to assess the source and quality of the papers, namely if they come from peer-reviewed journals as well as sorting out conference proceedings, reports, and other documents as grey literature.

Literature for three different relevant research topics was searched separately. The literature research topics are as follows:

1. COVID-19 pandemic restrictions impact on urban air pollutant concentration
   - Search query: (O3 OR NO2 OR SO2 OR CO OR PM2.5 OR PM10 OR C6H6) AND (Air Pollution) AND COVID-19

- Time filter: From 2020-01-01 until now
- Language: English, Portuguese
- Full-text articles analyzed: 17

2. Machine learning methods for air pollutant concentration prediction
   - Search query: (O3 OR NO2 OR SO2 OR CO OR PM2.5 OR PM10 OR C6H6) AND (Air Pollution) AND (Artificial Intelligence OR Machine Learning)
   - Time filter: From 2010-01-01 until now
   - Language: English, Portuguese
   - Full-text articles analyzed: 11

In addition to scholarly literature, several searches for grey literature, in the form of official technical and statistical reports, as well as other publications like fact sheets, were conducted in trusted and/or official Portuguese, European Union and International agencies related to the environment, demography, mobility, etc. (13/38 references). Finally, press articles related to pandemic management and the environment in Lisbon were also searched through Google News.

3 Impact of COVID-19 restrictions on urban air quality

The impact of sharp reductions of anthropogenic activities throughout many countries during COVID-19 pandemic associated lockdowns were widely analyzed as to its impact and contributions to primary and secondary air pollutant concentrations. Rana et al. (Rana et al., 2021) carried out a systematic literature review of the impact of COVID-19 restrictions on air quality in China using PRISMA guidelines where 35 studies out of 396 met the eligibility criteria and were thoroughly analyzed. All articles were published in 2020 and included articles using both satellite (12) and ground monitoring (23) air pollutant concentration data, with temporal comparison to both pre and post lockdown periods in 2020 (8), comparison to the same period in 2019 (15) and the remainder compared using period ranges from 2015 to 2019 (12). The most studied air pollutant was NO_2 (80%) followed by $PM_{2.5}$ (75%) and the authors estimate that during COVID-19 lockdown periods the mean reduction of air pollutant concentrations reported was 45.1% for NO2, 26.6% for PM2.5, 31.4% for PM10, 31.3% for SO2, 20.7% for CO and 21.7% for AQI.

Highlighting some additional studies, Bauwens et al. (Bauwens et al., 2020) used satellite observations to acquire NO2 vertical column measures from two satellite instruments, TROPOMI (TROPOspheric Monitoring Instrument) on board of Sentinel P-5 satellite and OMI (Ozone Monitoring Instrument) on board of Aura satellite, to reveal a sharp reductions of NO2 concentrations throughout the COVID-19 related lockdown phases in multiple cities throughout the world when compared to pre-lockdown periods and homologous periods in 2019. Notable examples include −43% to −57% in Wuhan, China, and −31% to −32% in Barcelona, Spain.

Connerton et al. (2020) used air quality data from ground monitoring stations in four megacities during the initial COVID-19 lockdown period in March 2020, and conducted a statistical analysis having measured pollutant concentration reductions, for instance in Paris, France, of 67% for CO, 39% for NO2 and 29% for PM2.5 when compared to a 2015–2019 air pollutant baseline. In order to extrapolate the contribution of the reduction of anthropogenic activities to the air pollutant measured concentration changes during the pandemic period, thus accounting for natural atmospheric phenomena that can disperse air pollutants or facilitate atmospheric chemical reactions that consumes said air pollutant, such as wind speed, air temperature and relative humidity, a general linear model was fitted with meteorology measures and a lockdown indicator variable as independent variables and air pollutant concentration as dependent variables, where it was found that while both anthropogenic activities and meteorology significantly influence air pollutant concentrations, anthropogenic activities reduction contribution was heavier.

Sicard et al. (2020) investigated the effect of sharp reductions of NOx emissions on O_3 concentrations during the COVID-19 related lockdowns in European and Chinese cities where it was suggested that local urban O_3 concentrations greatly increased as NOx emissions decreased due to reduced O_3 titration by NO thus raising the need of controlling VOC emissions to balance the NOx : VOC ratio which is key in tropospheric O_3 formation. For instance, during the 2020 COVID-19 related lockdown period in Wuhan, China daily mean O_3 concentration increased by 36% while NO_2 concentrations decreased 57%, when compared to a computed 2017–2019 baseline for the same period.

4 Machine learning based air pollutant concentration prediction

Advanced atmospheric chemical transport frameworks (E.P.A.) to model air pollution, such as CAMX or CMAQ, use as inputs a weather forecast model, an emission inventory for anthropogenic and biogenic sources and using a chemical transport model can output a 3D air pollution concentration map and compute source appointment analysis. On the other

hand, instead of a framework capable of simulating complex atmospheric chemical interactions, which can rely on outdated or incorrect estimated emission stocks, there are other data-driven techniques which use statistical and machine learning techniques to model air pollutant concentrations, capture the complex non-linear relationship of the several variables that determine the concentration of a given air pollutant at a specific place and time. Rybarczyk and Zalakeviciute (2018) have performed a systematic review of machine learning approaches to outdoor air pollution modeling following the PRISMA guidelines where 46 out of 103 papers published from 2010 to 2018 met the eligibility criteria. It was found that the volume of publications of applied machine learning techniques related to air quality has ramped up from 2016 onwards, mostly published with regards to northern hemisphere geographies, and are mostly related to criteria pollutants (NO_2, SO_2, CO, PM_{10}, $PM_{2.5}$ and O_3) or the air quality index (AQI), whereas the largest group of publications are related to identifying relevant predictor variables and modeling non-linear relationship of variables in air pollution. The main algorithm classes ordered by prevalence found in the publications were Ensembles (mainly tree-based predictors), ANNs, SVMs and LRs.

Further analyzing more recent publications, Vu et al. (2019) trained a Random Forest ML algorithm and compared prediction results for 2017 $PM_{2.5}$ concentrations in Beijing with those outputted by a WRF-CMAQ model and were able to produce a slightly more accurate value for the yearly mean $PM_{2.5}$ concentration, 61.8–62.4 $\mu g/m^3$ for the WRF-CMAQ model and 61.0 $\mu g/m^3$ for the RF ML model, whereas the observed value was 58.0 $\mu g/m^3$. At the month granularity for 2017, the WRF-CMAQ model concentration predictions ranged from 3% to 33.6% difference when compared to the observed values, a mean difference of 7.8%, whereas the RF ML model predictions ranged from 0.4% to 7.9%, a mean difference of 1.5%. Castelli et al. (2020) employed SVR (Support Vector Regression) algorithm to predict air pollutant concentrations such as NO_2, CO, SO_2, O_3 and $PM_{2.5}$ in California for years 2016 to 2018, using meteorological measures, pollutant concentration rolling means and timeseries features as predictors, achieving $R^2 = 0.937$ on the validation set using the RBF (Radial Basis Function) kernel for NO_2 forecasting. Luna et al. (2014) also employed an SVR (Support Vector Regression) algorithm to predict O_3 concentrations in Rio de Janeiro using ground monitoring stations data from 2011 and 2012, using as data features the chemical precursors, such as NO, NO_2, NO_x and CO, as well as meteorological factors, namely wind speed, solar radiation, air temperature and relative humidity. This model achieved $R^2 = 0.912$ on the validation set, having also trained an ANN (Artificial Neural Network) using the same data source to solve the same problem and achieving $R^2 = 0.915$ on the same validation set.

5 Data analysis

In order to pursue this work objectives and according to the literature review process, several data sources are required to properly model the urban air pollution phenomena: (1) Air pollutant concentration, (2) meteorological parameters and (3) anthropogenic activity direct or indirect indicators. In this sense, apart from data available online, several public entities and companies were directly contacted to acquire such data. From all the contacts and data gathered, a high-level summary of the most relevant data-sources used in the present work can be found in Juma (2021).

After the data extraction, transformation and loading process, the resulting analytical data model loaded into PowerBI was used to support all of the visualizations and consolidated data extractions for machine learning purposes. A simplified diagram of the data model can be found in Figure 1, the summary of the datasets in Table 1, and a more detailed description of all the datasets used in (Juma, Oct. 2021).

The present work mainly focuses on the months between March and July 2020, where the two main confinement periods occurred, to understand the impacts of restriction measures introduced by the authorities to manage the pandemic in the urban environment and mobility. The preemptive approach to the management of the pandemic by the Portuguese government, which included strict lockdowns, resulted in the limited spread of COVID-19 in the Lisbon and Tagus Valley health region as depicted in Figure 2. This region (a NUTS II unit until 2000) is often used as a standard for transport-related problems in the Lisbon area due to its high internal connectivity.

6 Mobility analysis during the COVID-19 pandemic

According to the index published by a public transportation mobility app (Moovit App, 2020) depicted in Figure 3, the most affected period was 09-04-2020 to 16-04-2020, where the 7-day rolling percentual change against the baseline estimated a 78.30% decrease in public transport usage. Even after the first state of emergency ended on the 2nd of May 2020, demand for public transportation recovered only slightly to values far lower than those of pre-pandemic periods, potentially due to general practice of telework, large scale lay-offs, reduced offer, and capacity of PT equipment, inexistent tourism as well as fear of contagium due to the enclosed nature of the public transportation equipment. During the first national emergency period (18-03-2020 to 03-05-2020) the mean reduction in public transportation demand is estimated

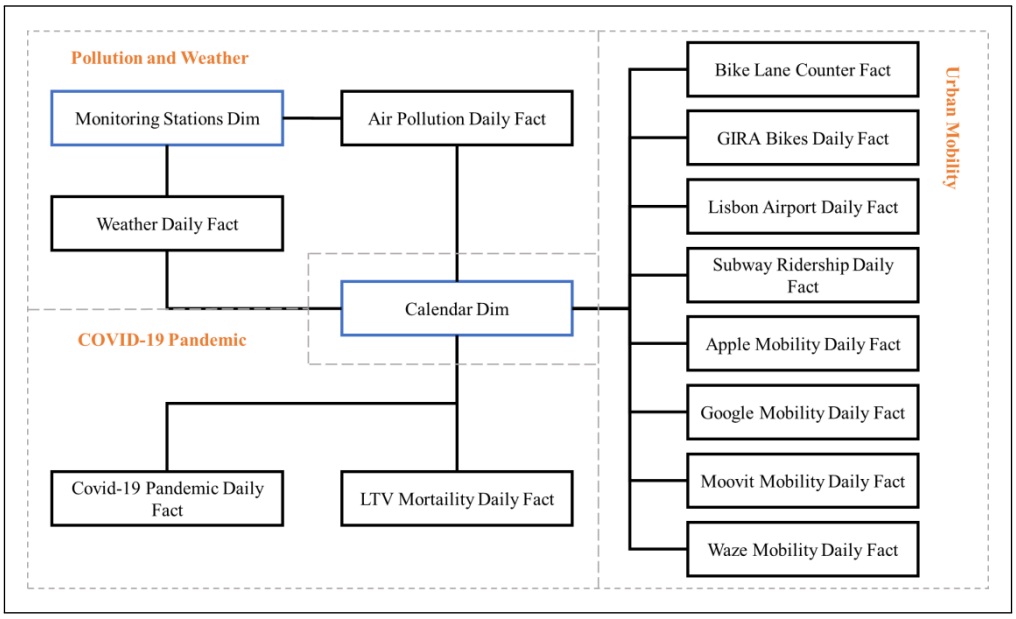


Figure 1. Simplified analytical data model built in the present work.

at 75.63% against the index baseline. The index would only come to reach the pre-pandemic baseline levels on the 13th of July 2021.

According to the Apple Mobility Trend Report (Apple Inc.), a mobility app which uses Apple Maps data from the navigation feature, the daily percentual change of direction requests in the city of Lisbon, Portugal against the baseline day of 13-01-2020 (a pre-lockdown Monday) provides details on other urban mobility means, namely walking and driving, which are two of the Apple Maps app modes.

According to this index, the day with largest reduction of walking navigation requests was on the 29th of March 2020 with a reduction of 92.60% and on the 12th of April for driving navigation requests with a reduction of 86.02% as depicted in Figure 4. These results must be interpreted cautiously since the baseline period for comparison is a pre-lockdown Monday and both the dates with largest reductions during the lockdowns are Sundays which are days where there's naturally reduced urban mobility associated to the weekends.

Nevertheless, the results are consistent with the high initial adherence to lockdowns by the population, where in the first national emergency state (18-03-2020 to 03-05-2020) walking trips reduced in 89% and driving trips reduced in 78.84% against baseline, while during the first national calamity state (04-05-2020 to 01-07-2020) walking trips reduced in 72% and driving trips reduced in 43.05% against baseline. From then on, lockdown erosion and continuous de-escalation of confinement contributed to a steady increase in both mobility patterns until the summer.

From the 1st of August 2020 onwards, the remaining 19 Lisbon City Hall Assembly parishes still in calamity state, which still had special mobility restrictions due to the virus incidence, joined the rest of the Lisbon parishes in the lower state of contingency. This appears to be correlated with a sharp increase in walking trips starting in early August, with additional potential contributions from tourism, which peaked in August 2020, and the vacation periods that might have been spent closer to home.

Waze App (Waze) follows a similar distribution to the Apple Driving index but with a negative offset that widens from the beginning of the lockdown period until the summer period, potentially due to the fact that values are in driven kilometers and not the number of trips. In contrast, it is possible that the average driven kilometers per trip during the lockdown periods were, overall, lower due to the more direct trips and non-leisure use of the private vehicle.

During the first national emergency period (18-03-2020 to 03-05-2020), it registered a 75.63% decrease in the daily driven kilometres registered with the Waze App while the first state of calamity (04-05-2020 to 01-07-2020), which was marked by lockdown erosion and continuous de-escalation of confinement, registered a 56.13% decrease against the baseline.

According to Google Community Mobility Index (Juma, Oct. 2021) a report on city mobility based on mobile phone data, depicted in Figure 5, the days with the sharpest reduction of activity in Grocery and Pharmacy (Commerce) was the 12th of April 2020 with a decrease in activity of 84%, for Workplace was the 10th of April 2020 with a decrease of 89%, for

Table 1. Datasets description.

Entity	Dataset	Grain	Period	Short description
IPMA	Meteorological parameters	Hour	2019-2021*	The measures include wind speed (m/s) and direction (°), relative humidity (%), air temperature (°C), sea level air pressure (hPa), total global radiation (KJ/m2) and accumulated precipitation (mm)
APA/EEA	Air Pollutants Concentration	Hour	2013–2021*	Concentration values ($\mu g/m^3$) for primary pollutants *NO*2 (Nitrogen Dioxide), *CO* (Carbon Monoxide), SO_2 (Sulphur Dioxide) and secondary pollutants *PM*10 Particulate Matter $< 10\mu$m of diameter, *PM*2.5 (Particulate Matter $< 2.5\mu$m of diameter) and *O*3 (Ozone).
Apple	Apple mobility trend report	Daily	2020–2021*	Daily percentual change of Apple Maps App walking and driving direction requests in the city of Lisbon, Portugal against the baseline day of 13–01–2020 (a pre-lockdown Monday).
Google	Community mobility report	Daily	2020–2021*	Per-weekday percentual change in user activity per location type (Retail & Recreation, Grocery & Pharmacy, Parks, Transit Stations, Workplace and Residential) of users with Google location services enabled in their mobile devices against a per-weekday baseline computed between 03–01–2020 and 06-02-2020 (pre-lockdown).
Moovit	Moovit Insights Public Transit Index	Daily	2020-2021*	Lisbon 7-day public transportation ridership percentual change, measured with the Moovit App, from a baseline week computed for the week of 08-01-2020 to 15-01-2020 (pre-lockdown).
Waze	COVID-19 Impact dashboard	Daily	2020-2021*	Percentual change of kilometers driven per day measured with the Waze App within the city of Lisbon against a pre-computed baseline scoped to the weekday for the period between 11-02-2020 and 25-02-2020 (pre-lockdown).
Metropolitano de Lisboa	Lisbon subway rides	Daily	2019–2020	Lisbon subway network daily cumulative rides (one-way trips) per type of pass (occasional, regular, child, etc)
ANA Portugal	Lisbon international airport traffic	Daily	2019–2020	Lisbon international airport (Humberto Delgado Airport) daily cumulative aircraft movements (inbound, outbound) and non-transferred passengers (passengers just transiting).
DGS	COVID-19 pandemic report	Daily	2020–2021*	COVID-19 Pandemic related health indicators communicated by public entities broken down by geographical context.
SICO/EVM	Lisbon and Tagus Valley Mortality	Daily	2020–2021*	Consolidated data on all deaths registered in Portugal broken down by death type, age group and by geography.
EMEL	GIRA bike ridership	Daily	2019–2020	Daily Lisbon GIRA shared bike program ridership.

Parks was the 5th of April 2020 with a decrease of 91%, for Retail and Leisure was the 12th of April 2020 with a decrease of 91% and for Public Transportation was the 10th of April 2020 with a decrease of 90% (Juma, Oct. 2021). On the other hand, for Residential activity there was a general increase peaking on the 10th of April 2020 with a maximum of +47%.

As we can imagine from the mobility report of Figure 5, Lisbon Subway as followed the same trend, being one of the three main public transportation modes used in the city, it showcases the dramatic effect of the lockdown in the city.

The YoY (a database with daily counts of subway trips in Lisbon) analysis should be done with care since from September 2019 onwards a change in the prices of recurrent titles (From April, your pass costs less - Metropolitano de Lisboa) increased the baseline usage of the subway network starting in April 2019. The first two months of 2020 where no restrictions were in place, compared to the homologous month in 2019, regular passes had an increase in ridership of 23%, free child passes 39%, and overall ridership of 13%.

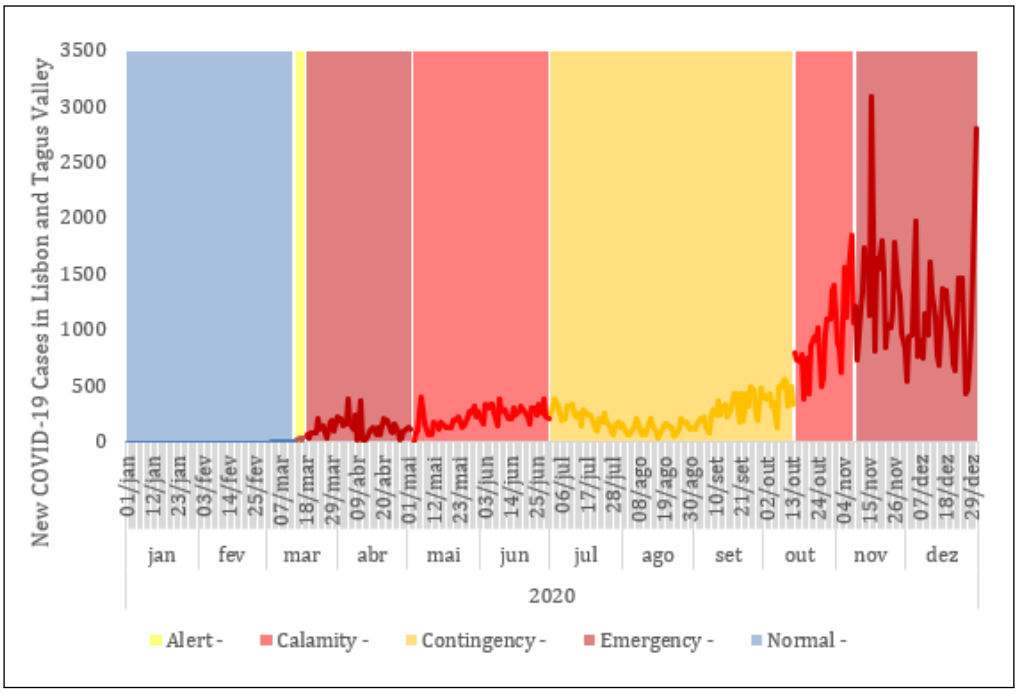


Figure 2. Main phases of the pandemic management in the Lisbon Metropolitan Area.

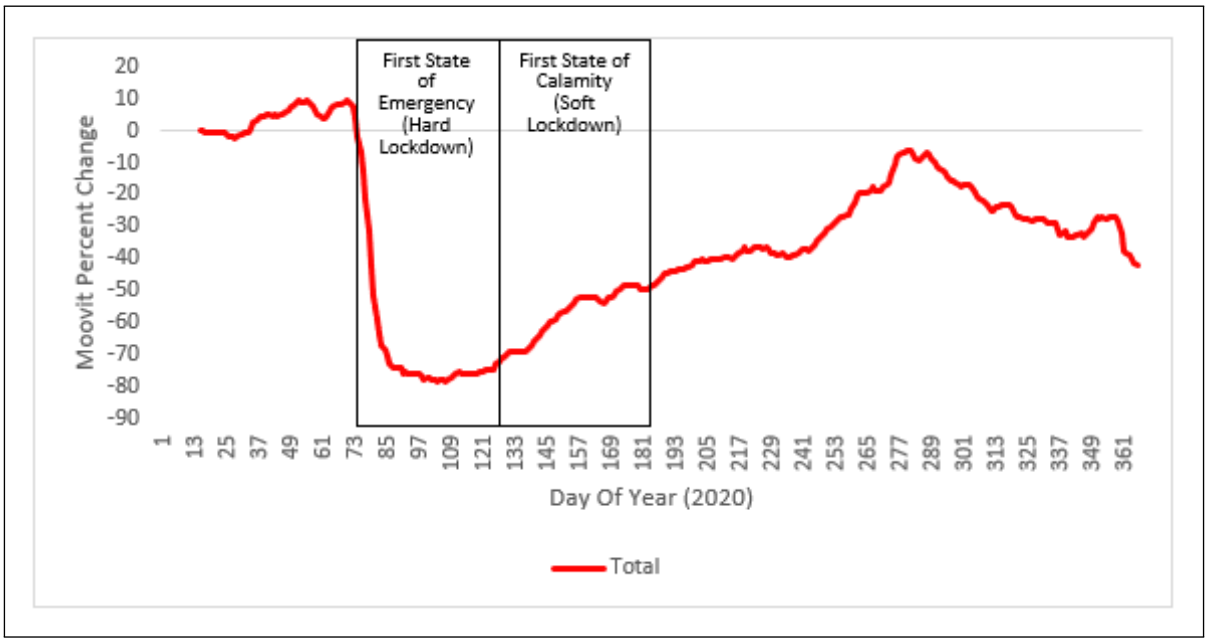


Figure 3. Moovit Insights Public Transportation Index for Lisbon 2020.

The most impacted month was April/2020, with a homologous decrease in total ridership of 84% when compared to 2019. For the same month, when compared to January/2020 baseline, there's an 85% decrease in total ridership. The most impacted type of trip was the occasional title which measured a homologous ridership decrease of 92% and a decrease of 89% when compared to January/2020. This was expected since leisure and tourist ridership were virtually halted during the initial lockdowns.

Bike trips fell by 68.47% when compared to the same period in 2019 but recovered during the first state of calamity (04-05-2020 to 01-07-2020), measuring an increase of 22.59% in ridership when compared to the same period in 2019.

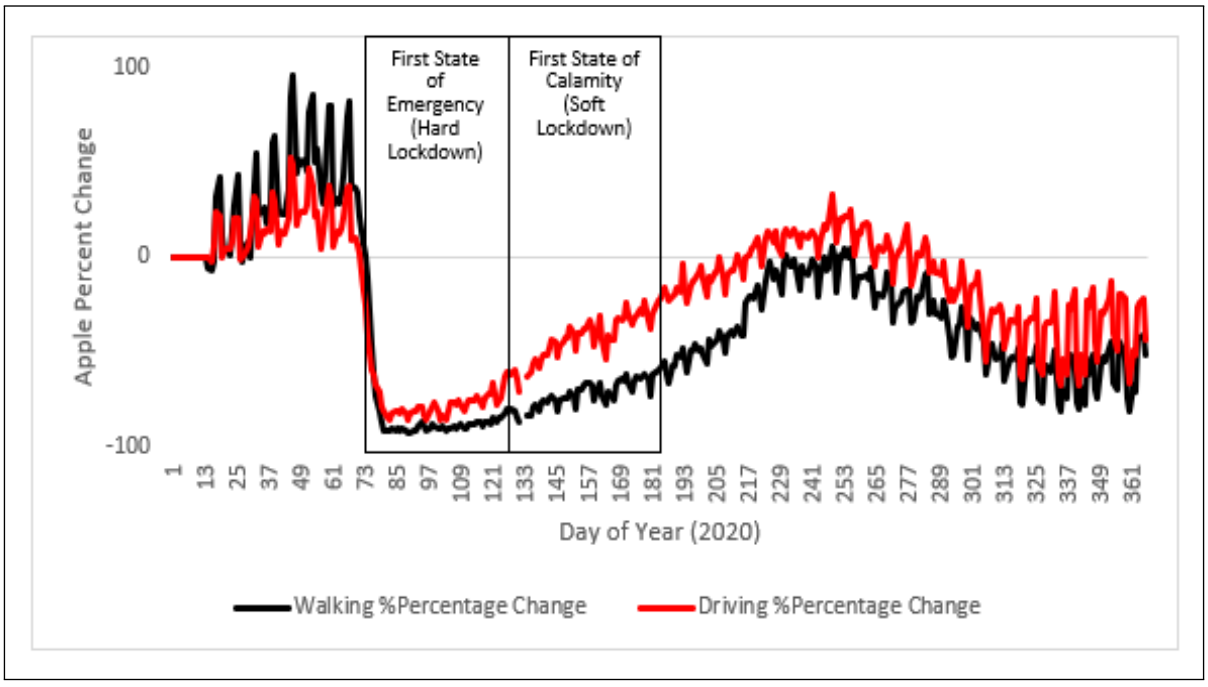


Figure 4. Apple Mobility Trend Report for Lisbon.

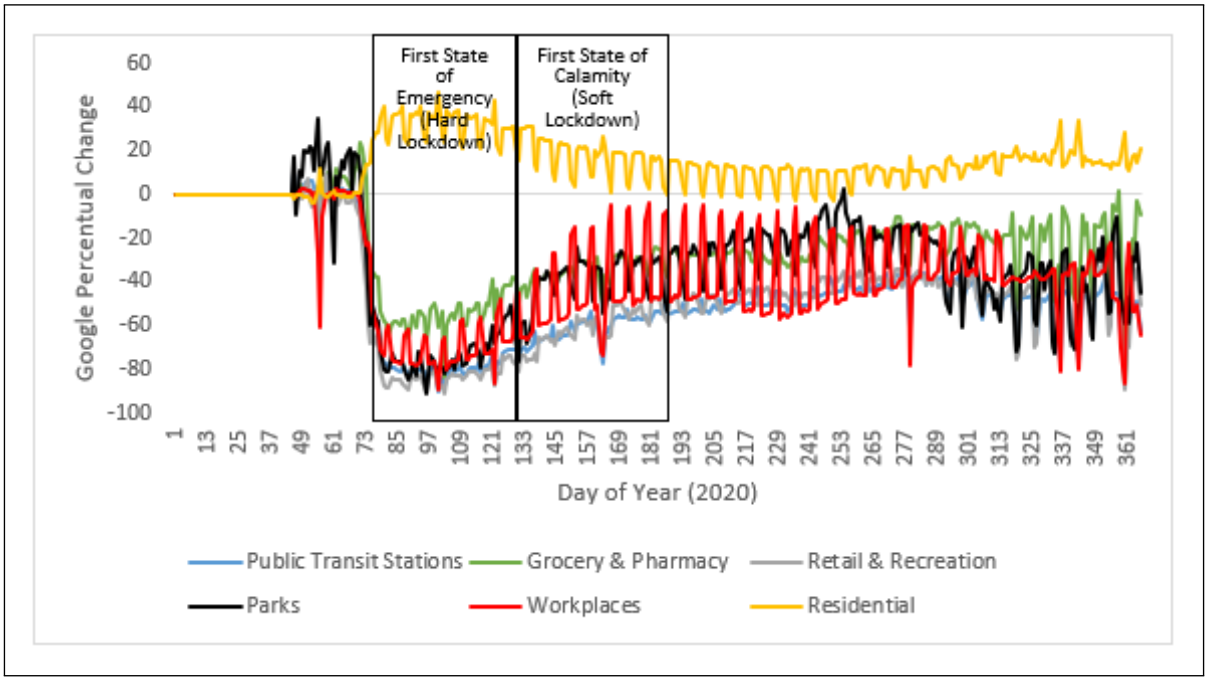


Figure 5. Google Community Report for Lisbon 2020.

After the first state of emergency, the usage of the shared GIRA bike during the first state of calamity increased in the greater order of magnitudes than other types of urban transportation, namely the subway, which still measured a 72.56% retraction against the same period in 2019. This could be attributed to increased demand for soft transportation means due to fear of contagion or because GIRA bikes might also be used as an outdoor leisure activity and not only as a means of transportation.

International non-resident business or leisure visitors make their way to Lisbon from their home countries primarily by means of air transportation (S.E.F./G.E.P.F., 2020). Besides aircraft emitted air and noise pollution, visitors temporarily enlarge Lisbon population and, therefore the human footprint in terms of anthropogenic pollutant activities and additional pressure on public transportation systems.

With regards to the Lisbon's Airport data, the total non-transferred passengers during COVID-19 restrictions had a homologous reduction of 99.11% in the 15th week of the year and was measured to be only 4.423 passengers in 2020 (2020-04-05 to 2020-04-11) against 498.565 passengers in 2019 (2019-04-07 to 2019-04-13). Regarding the number of total aircraft movements, it had a homologous reduction of 96.44% in the same analysis period wherein 2020 there was only 152 movements against 4273 movements in 2019, and some of the few flights were related to citizen repatriation efforts.

There was a partial recovery from July 2020 onwards, being this related to the lifting of self-imposed travel restrictions by the Portuguese government and additional relaxation of travelling rules from August 2020 onwards also had an effect (Portugal, 2020). The overall volume of passengers was also affected by travel restrictions imposed by the countries of origin of tourists or expatriates that visit Portugal (England, Germany, France, Italy, Spain, The Netherlands, etc.) and reciprocate restrictions of not allowing Portuguese citizens to travel to their countries.

By computing the Pearson Correlation of both public transportation mobility indexes for the year 2020 with an independent public transportation variable (subway ridership), it is found that Google Mobility Index for Public Transport correlates better than the Moovit COVID-19 Public Transport Impact Index ($r = 0.99$ vs $r = 0.91$) and is thus a mobility index that correctly models public transport commuter routines in the city. This could be useful to gauge changes in public transportation usage and to be used as a proxy indicator for public transportation usage in, for instance, the training of machine learning models.

A visual summary of the main impacts on urban mobility indicators in Lisbon during the first national emergency state period from 18-03-2020 to 03-05-2020 and during the first national calamity state period from 04-05-2020 to 01-07-2020 can be found in Figure 6.

While the present work used in-situ measures acquired by ground air quality monitoring stations, one can use satellite observations to have a high-level view of the impact of the confinement measures on Nitrogen dioxide (NO_2), a pollutant emitted directly and indirectly by internal combustion vehicles. During the first national emergency state lockdown and the first half of the first calamity state (March-June), where the most restrictive measures affecting anthropogenic activities were in-place, we can observe a clear reduction on the density of the NO_2 vertical column measured by the TROPOMI instrument aboard Sentinel-P5 satellite when compared to the same period in 2019 as shown in Figure 7.

7 No_2 predictive analysis

Apart from expert analysis of weather and pollutant emission data using traditional statistical analysis to assess the relationship between independent variables and the dependent variable, and, in the case of air pollution modelling, the usage of advanced and complex atmospheric chemical transport simulation models, it is also possible to use Machine Learning techniques to model physical and chemical phenomena involved in air pollution. In this work, the AutoML framework TPOT (Tree-Based Pipeline Optimization Tool) (TPOT Home Page - Data Science Assistant, 2021) is used to discover a performant and interpretable machine learning model capable of explaining the relationship between independent and dependent variables as well as their strength. This automatic parameter optimization (AutoML) tool generates a tree-based search by splitting one parameter value at every branch and evaluating effectiveness with small trainings.

The present work intends to train a sufficiently accurate model capable of identifying the most relevant independent data features used to model the NO_2 urban concentration phenomena and to identify any relation between independent features that might be of interest.

The list of initial raw data features used to model NO_2 concentration was selected from the datasets described in Table 1, taking into consideration known variables that potentiate the formation and destruction of the NO_2 air pollutant also described in the literature.

To perform this initial analysis, TPOT genetic algorithm optimizer ran 100 generations and 1000 retained individuals in every generation with a total of 101.000 model fits evaluations using sklearn (EUROCONTROL, Jul.–Sep. 2020) KFold cross-validation with 5 folds on top of 75% of the original data ($N = 381$) whereas the remainder 25% were kept for model performance testing purposes. In this initial run no specific template was used and TPOT ran with full autonomy, using all the 30 processors in the default Regressor configuration (Pennsylvania), to build a pipeline with whatever stages and with whatever hyperparameters that maximizes the NO_2 model accuracy.

This initial TPOT execution took roughly six days and two hours in a D4a_v4 Azure Virtual Machine (4 vCores, 16GB RAM and SSD Volume) using all the machine cores, keeping a cache of intermediate pipeline results to avoid unnecessary

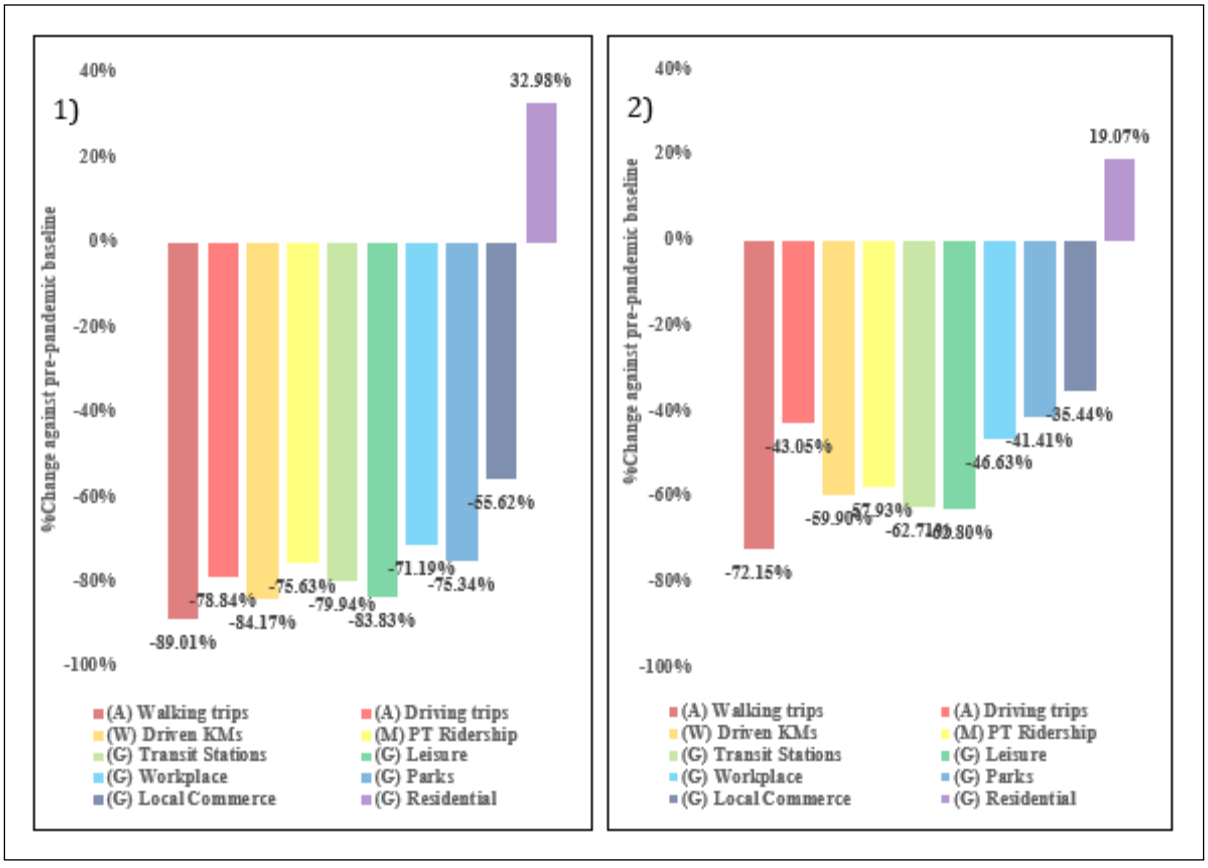


Figure 6. Impact on urban mobility indicators in Lisbon during the first national emergency state period from 18-03-2020 to 03-05-2020 (1) and during the first national calamity state period from 04-05-2020 to 01-07-2020 (2), against a pre-pandemic baseline.

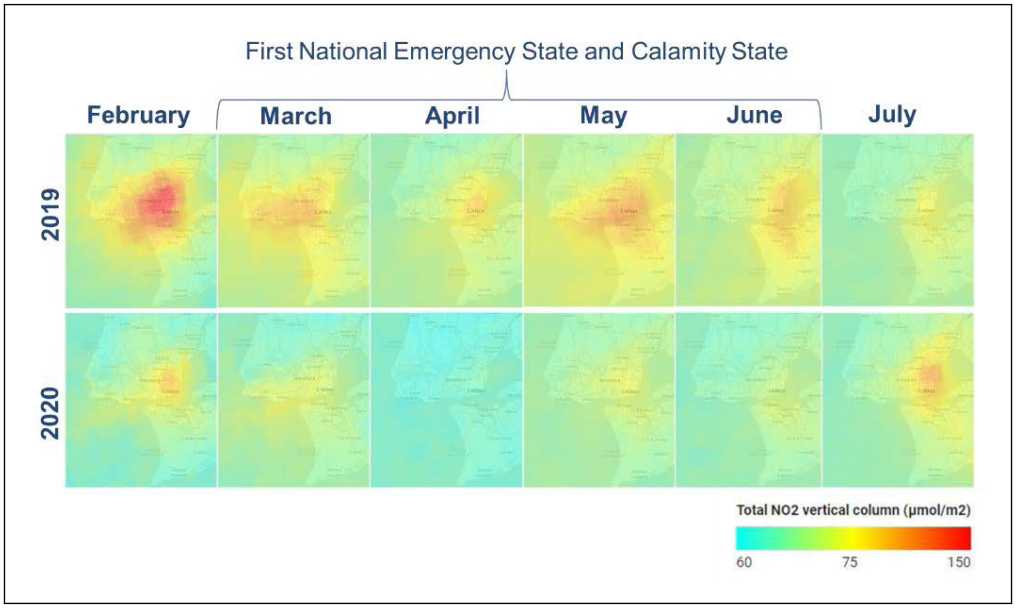


Figure 7. Lisbon and Tagus Valley Monthly Average of Total NO2 vertical column (μmol/m2) measured by the TROPOMI instrument onboard Sentinel-P5 satellite.

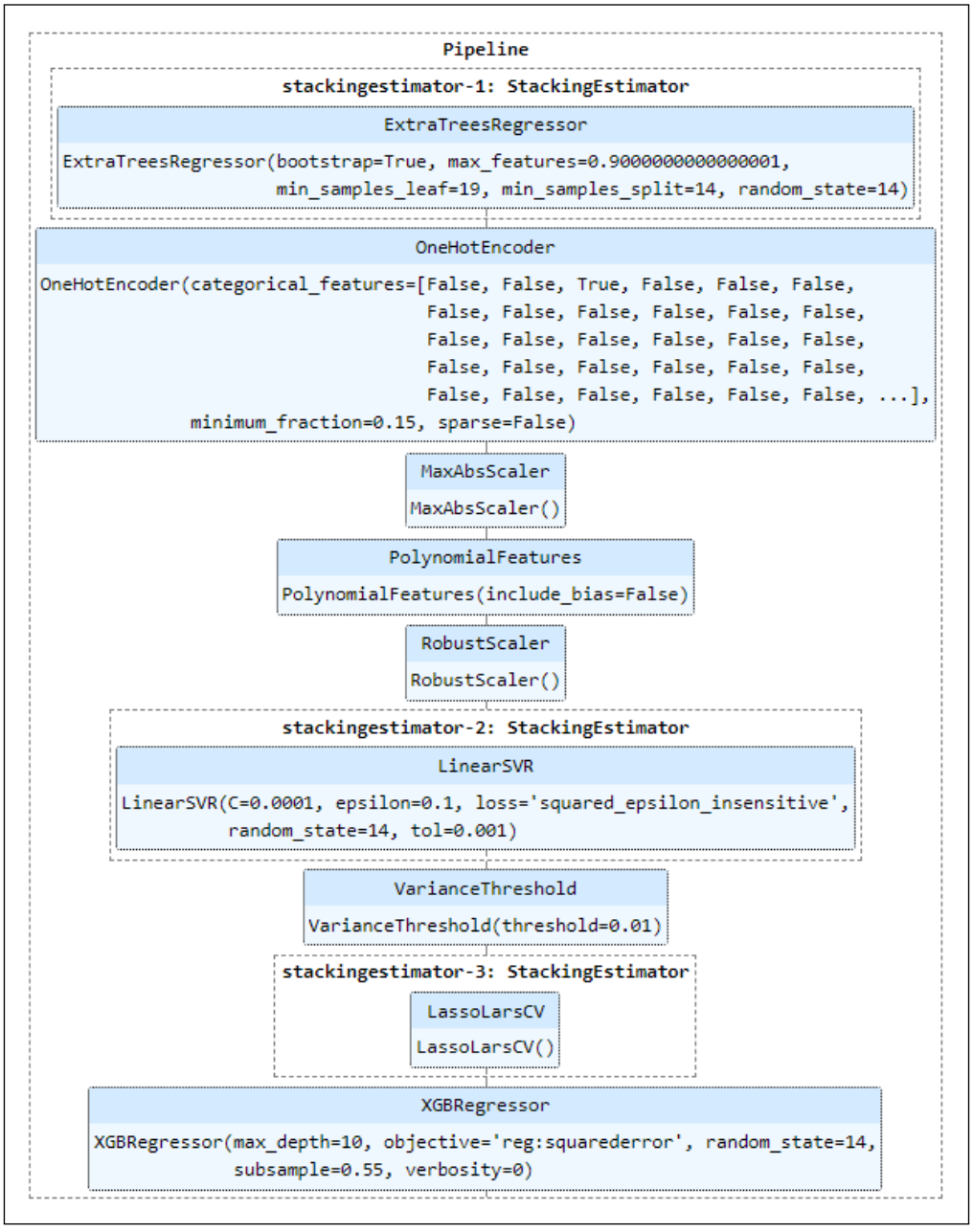


Figure 8. Exploratory AutoML based NO2 prediction pipeline stages and hyperparameters.

recomputations and periodically checkpointing the optimization process progress. This execution resulted in a complex pipeline with nine stages, four pre-processors, one feature-selector and four stacked regressors, with a resulting average KFold Cross-Validation score on the training set of 12.87 (MSE) and 11.92 (MSE) on the 25% of records that was held for model testing. Having a similar MSE on both the CV (train + validate set) and test set is a good sign that the model is not overfitting or underfitting. While another pipeline performed better on the training set with MSE = 12.32, the reported pipeline performed better on the test set (Figure 8 and Table 2) and was therefore selected as the reference NO_2 prediction regressor ML pipeline (Figure 9).

During this first run, the pipeline cross-validation score evolved rapidly during the first phases of the optimization process progress. It slowed down on later phases with each generation taking longer to complete which is expectable since the average complexity of the pipelines increases over time with additional pipeline stages, unions, and stacks that take longer to compute. The time required to complete a generation stabilized at around generation 40 since the optimizer

Table 2. Metrics for the prediction of NO_2 concentration using AutoML.

Metric	Result
Mean Absolute Error (MAE):	2.77
Mean Squared Error (MSE):	11.93
Root Mean Squared Error (RMSE):	3.45
Mean Squared Log Error (MSLE):	0.01
R^2 Score:	0.92

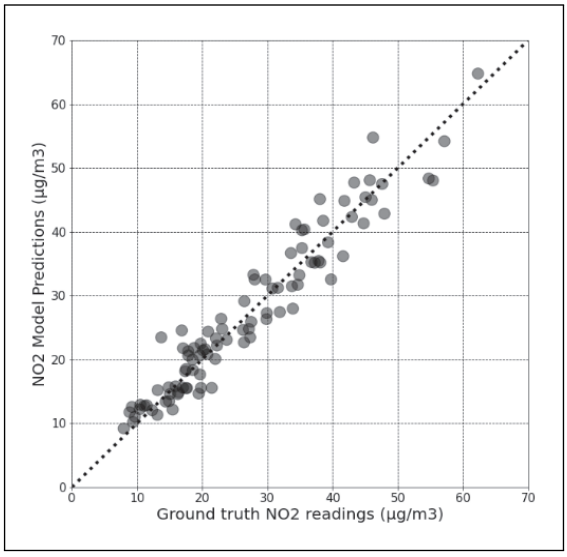


Figure 9. Exploratory AutoML based NO2 prediction performance.

focused on a similar cost and complex search space for the remainder of the optimization process. It took 87 generations in 100 to achieve the best score, which took 122 h in a total of 146 h, and it took 36 generations in 100 to achieve 80% of the optimization from a min-max perspective, which took 32 h in a total of 146 h. A depiction of how this process unfolded throughout the optimization process can be found in Figure 10 with the highlighted area representing the fulfilment of 80% of the training CV score optimization.

8 Conclusions

As for the impact of COVID-19 pandemic related restrictions in Lisbon on the urban mobility, it is estimated that during the first national emergency period (18-03-2020 to 03-05-2020) public transportation ridership has suffered a homologous drop of 75% to 80% while during the first calamity state period (04-05-2020 to 01-07-2020) it is estimated to have dropped 57% to 62%. Regarding the usage of private vehicles, during the first national emergency period, it is estimated to have suffered a reduction of 78% to 84% while during the first calamity state period it is estimated to have dropped 43% to 59%. Residential area mobility activity has increased 33% during the first national emergency period, decreasing to 19% during the first calamity state period. The urban mobility indicators used in the present work correlate moderately with the NO_2 air pollutant which is usually associated to anthropogenic activity in the city, whereas the local commerce indicator (Google) has the strongest Pearson correlation ($r = 0.54$) and has also been identified as the main anthropogenic data feature contributing to NO_2 concentration by the trained NO_2 concentration prediction pipeline.

Regarding the impact of COVID-19 pandemic related restrictions in Lisbon on the air quality, during the first national emergency period (18-03-2020 to 03-05-2020) and subsequent calamity state period (04-05-2020 to 01-07-2020) the main criteria air pollutants have generally decreased in both urban background and urban traffic stations when compared to a 2013-2019 baseline with the exception of Ozone (O_3) in urban traffic stations which has increased. With a sharp reduction in anthropogenic activities, most importantly in road traffic, and with depressionary weather suboptimal for NO_2 build-ups in the first phase, NO_2 registered a 54.35% drop in the first phase and 40.39% in the second phase in urban traffic stations, while in background stations it dropped 28.62% and 22.99%, respectively.

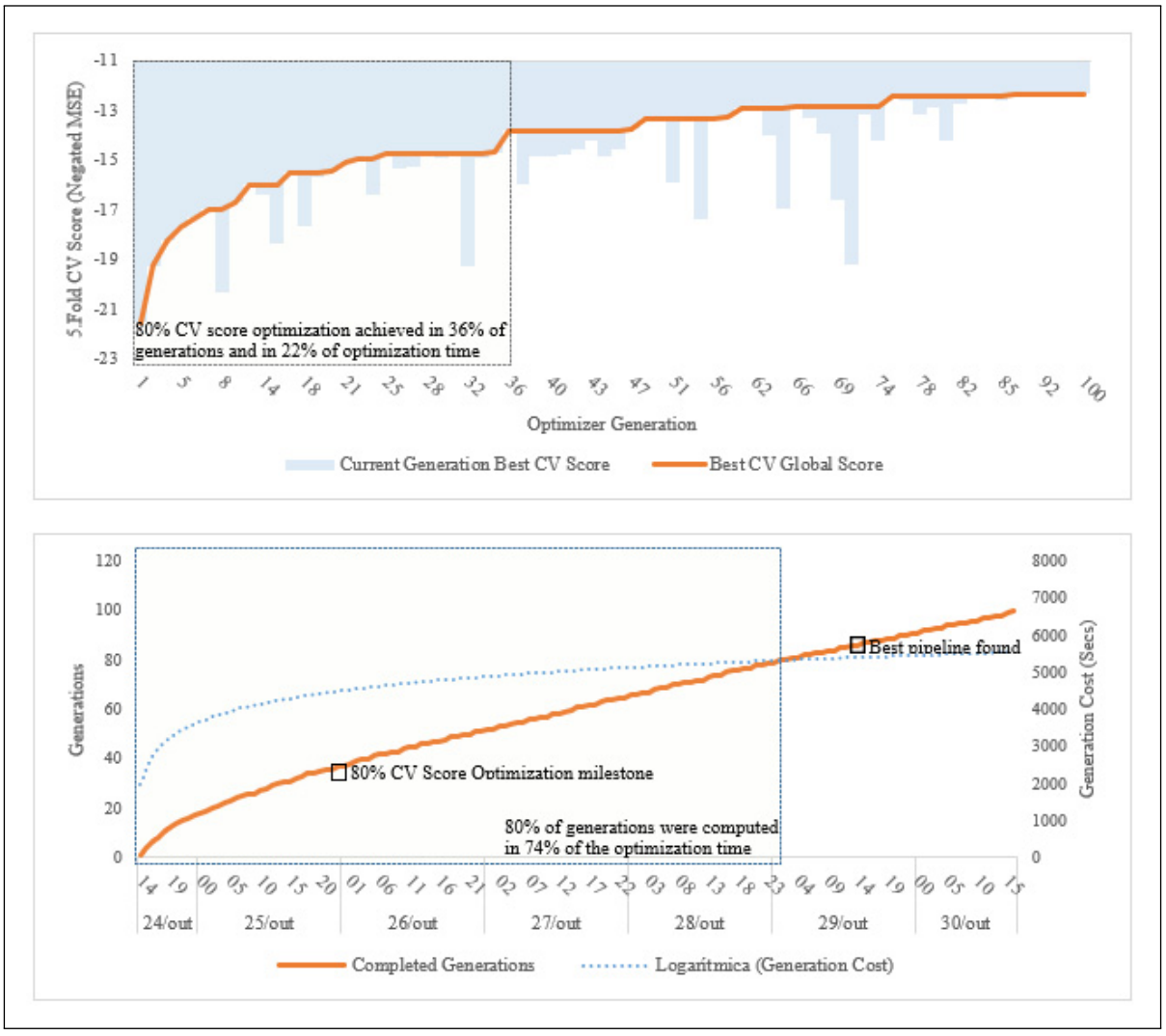


Figure 10. TPOT AutoML training efficiency.

These results are consistent with (Olson et al.; Pedregosa et al., 2011) taken in Wuhan, China, in similar periods, where the daily mean of O_3 concentration increased by 36% and NO_2 concentrations decreased 57%.

An AutoML framework (TPOT) was used to build, train, and optimize a regressor ML pipeline to predict NO_2 concentration with the available anthropogenic activity, weather, and air pollutant inputs from March/2020 to March/2021, achieving $R^2 = 0.925$ out of the box on the test set. This is an acceptable result for an AutoML approach when compared to other recently purpose-built NO_2 prediction models such as $R^2 = 0.920$ (Sicard et al., 2020), $R^2 = 0.890$ (Srivastava et al., May 2021) and $R^2 = 0.937$ (Castelli et al., 2020; Zhan et al., 2020).

Acknowledgments

This work was partially supported by Fundação para a Ciência e a Tecnologia, I.P. (FCT) [ISTAR Projects: UIDB/04466/2020 and UIDP/04466/2020].

ORCID iDs

Luis B Elvas https://orcid.org/0000-0002-7489-4380
Joao C Ferreira https://orcid.org/0000-0002-6662-0806
Luis Nunes https://orcid.org/0000-0001-7072-0925

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was partly funded through national funds by FCT - Fundação para a Ciência e Tecnologia, I.P. under project UIDB/04466/2020 & UIDP/04466/2020 (ISTAR).

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

Apple Inc. "Mobility Trends Reports," accessed: Jan. 2021. [Online]. Available: https://covid19.apple.com/mobility.

Associação de Utilizadores de Veículos Elétricos (UVE) (Jan. 5, 2023) Annual battery-electric and plug-in hybrid electric vehicle sales in Portugal between 2019 and 2022," Statista, Lisbon, Portugal, [Online]. Available: https://www.statista.com/statistics/1380313/portugal-annual-plug-in-electric-vehicle-sales/.

Bauwens M, Compernolle S, Stavrakou T, et al. (2020) Impact of coronavirus outbreak on NO2 pollution assessed using TROPOMI and OMI observations. *Geophys. Res. Lett.* 47(11): e2020GL087978.

Castelli M, Clemente FM, Popovič A, et al. (2020) A machine learning approach to predict air quality in California. *Complexity* 2020: 1–23.

Commission Regulation (EU). No 459/2012 of 29 May 2012 amending Regulation (EC) No 715/2007 of the European Parliament and of the Council and Commission Regulation (EC) No 692/2008 as regards emissions from light passenger and commercial vehicles (Euro 6) Text with EEA relevance, vol. 142. 2012. Accessed: Nov. 18, 2021. [Online]. Available: http://data.europa.eu/eli/reg/2012/459/oj/eng.

Connerton P, Assunção JV, Miranda RM, et al. (2020) Air Quality during COVID-19 in Four Megacities: Lessons and Challenges for Public Health. *Int. J. Environ. Res. Public. Health* 17(14): 5067.

Coordenação C, Lisboa DR and Tejo V. (2019) "Assessment of ambient air quality in the Lisbon and Tagus Valley region." [Online]. Available: http://www.ccdr-lvt.pt/files/acd442d5056c22381724d85aa9ffe3f179b49fe3.pdf.

E.E.A. (2018) "Air quality in Europe — 2018 report." [Online]. Available: https://www.eea.europa.eu/publications/air-quality-in-europe-2018.

E.P.A. "Photochemical Air Quality Modeling," *EPA*, [Online]. Available: https://www.epa.gov/scram/photochemical-air-quality-modeling.

Estatísticas," Statistics Port of Lisbon. Accessed: Nov. 18, 2021. [Online]. Available: https://www.portodelisboa.pt/estatisticas.

EUROCONTROL. (Jul.–Sep. 2020) "COVID-19 Impact on the European Air Traffic Network," [Online]. Available: https://www.eurocontrol.int/covid19.

From April, your pass costs less - Metropolitano de Lisboa. Accessed: Nov. 23, 2021. [Online]. Available: https://www.metrolisboa.pt/2019/03/18/novos-passes-aml/.

Ghorani-Azam A, Riahi-Zanjani B and Balali-Mood M (Sept. 2016) Effects of air pollution on human health and practical measures for prevention in Iran. *J. Res. Med. Sci. Off. J. Isfahan Univ. Med. Sci.* 21: 65.

Hooftman N, Oliveira L, Messagie M, et al. (Feb. 2016) Environmental Analysis of Petrol, Diesel and Electric Passenger Cars in a Belgian Urban Setting. *Energies* 9(2): 84.

Hook A, Court V, Sovacool BK, et al. (Aug. 2020) A systematic review of the energy and climate impacts of teleworking. *Environ. Res. Lett.* 15(9): 093003.

Juma ASV, (Oct. 2021). *Impact of COVID-19 Pandemic Restrictions in Urban Mobility and Air Pollution in Lisbon*. Portugal, M.Sc. thesis. ISCTE – Instituto Universitário de Lisboa.

Khomenko S, Cirach M, Pereira-Barboza E, et al. (2021) Premature mortality due to air pollution in European cities: A health impact assessment. *Lancet Planet. Health* 5(3): e121–e134.

Luna A, Paredes M, Oliveira G, et al. (2014) Prediction of ozone concentration in tropospheric levels using artificial neural networks and support vector machine at Rio de Janeiro, Brazil. *Atmos. Environ.* 98: 98–104.

Mobility and functionality of the territory in the Metropolitan Areas of Porto and Lisbon, (2017). Accessed: Nov. 18, 2021. [Online]. Available: https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_publicacoes&PUBLICACOESpub_boui=349495406&PUBLICACOESmodo=2.

Moovit App (2020) "Moovit Insights Public Transit Index – Lisbon," accessed: Jan. 2021. [Online]. Available: https://moovitapp.com/insights/ Google, "Google Community Mobility Reports.". [Online]. Available: https://www.google.com/covid19/mobility/.

Motorised road vehicles (No.) by Vehicle Type and Fuel Type." Accessed: Nov. 18, 2021. [Online]. Available: https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_indicadores&contecto=pi&indOcorrCod=0007244&selTab=tab0&xlang=pt.

Olson R, Bartley N, Urbanowicz R, et al. (2016) Evaluation of a Tree-based Pipeline Optimization Tool for Automating Data Science. In: Proc. GECCO '16: Genetic and Evolutionary Computation Conference Companion, Denver, CO, USA, pp. 485–492. doi: 10.1145/2908961.2908993.

Passenger traffic at the Lisbon Airport. Accessed: Nov. 22, 2021. [Online]. Available: https://www.pordata.pt/Portugal/Tr%C3%A1fego+de+passageiros+nos+principais+aeroportos+Lisboa++Porto+e+Faro-3248.

Pedregosa F, Varoquaux G, Gramfort A, et al. (2011) Scikit-learn: Machine learning in python. *Journal of Machine Learning Research* 12: 2825–2830.

Pennsylvania U. "TPOT - Default Regression configuration." [Online]. Available: https://github.com/EpistasisLab/tpot/blob/master/tpot/config/regressor.py.

Portugal G. (2020) "Air traffic rules to and from Portugal from 1 August." [Online]. Available: https://www.portugal.gov.pt/pt/gc22/comunicacao/comunicado?i=regras-de-trafego-aereo-de-e-para-portugal-a-partir-de-1-de-agosto.

Rana RH, Keramat SA and Gow J (2021) A Systematic Literature Review of the Impact of COVID-19 Lockdowns on Air Quality in China. *Aerosol Air Qual. Res.* 21: 200614

REN - Consumo de energia elétrica recua 12% em abril. Accessed: Nov. 18, 2021. [Online]. Available: https://www.ren.pt/pt-PT/media/comunicados/detalhe/consumo_de_energia_eletrica_recua_12__em_abril_3.

Rybarczyk Y and Zalakeviciute R (2018) Machine learning approaches for outdoor air quality modelling: A systematic review. *Appl. Sci.* 8: 2570.

S.E.F./G.E.P.F. (2020) *Immigration, Borders and Asylum Report 2019*. Oeiras: SEF.

Sicard P, De Marco A, Agathokleous E, et al. (2020) Amplified ozone pollution in cities during the COVID-19 lockdown. *Sci. Total Environ.* 735: 139542.

Srivastava P, Kumar A, Bauddh M, et al. (May 2021) Artificial intelligence modeling approach for forecasting of air pollutants: A review. *Environmental Pollution* 271: 116380.

Sugawara H, Ishidoya S, Terao Y, et al. (2021) Anthropogenic CO2 emissions changes in an urban area of Tokyo, Japan, due to the COVID-19 pandemic: A case study during the state of emergency in April–may 2020. *Geophys. Res. Lett.* 48(15): e2021GL092600.

TPOT Home Page - Data Science Assistant (2021) Epistasis Lab at Cedars Sinai, Accessed: Dec. 11, 2021. [Online]. Available: https://github.com/EpistasisLab/tpot.

Vu TV, Shi Z, Cheng J, et al. (2019) Assessing the impact of clean air action on air quality trends in Beijing using a machine learning technique. *Atmospheric Chem. Phys.* 19(17): 11303–11314.

Waze. "Waze COVID-19 Impact Dashboard," Waze, [Online]. Available: https://www.waze.com/en-GB/covid19. [Acedido em 30 09 2021].

W. H. Organization (2021) "Ambient (outdoor) air pollution fact sheet." [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health.

Zhan F, Wu Y and Deng H (2020) Prediction of NO_2 concentration based on an improved deep learning model. *Atmosphere* 11(11): 1165.