Università degli studi di Milano-Bicocca

Data Science Lab for Smart Cities

Final Essay

Impact of Urban Traffic Restrictions on Air Quality: Predictive Modeling in Milan

Authors:
Alfio Leanza - 910235- a.leanza3@campus.unimib.it
Riccardo Tavecchio - 852211- r.tavecchio1@campus.unimib.it

July 11, 2025

Abstract

This study explores Milan's air quality, strongly affected by nitrogen dioxide (NO_2) emitted by traffic, is still one of the city's main environmental challenges. In this study we examine the trend of NO_2 from 2018 to May 2025 by integrating GIS information to distinguish three geographic areas: Area C, Area B, and the belt outside the municipality. The analysis considers temporal changes in concentrations before the activation of Area B (February 2019), after its introduction, and following its tightening on October 1, 2024, and relates these changes to spatial differences among the three areas. Downstream of the descriptive characterization, we build time series models with the aim of producing 12 month forecasts and quantifying the sensitivity of concentrations to traffic factors. The results provide a measure of the effectiveness of low-emission zone policies implemented in Milan and offer predictive tools useful for planning further sustainable mobility interventions.

Contents

1 Introduction

Air quality is one of the main environmental challenges for European cities; among regulated pollutants, nitrogen dioxide (NO_2) plays a key role because it is largely emitted by internal combustion engines, acts as a precursor to secondary aerosols and tropospheric ozone, and is associated with acute and chronic respiratory and cardiovascular health effects [1]. The World Health Organization updated its Air Quality Guidelines in 2021 [2], recommending an annual target value of 10 $\mu g/m^3$ to protect public health; current European legislation still sets a limit of 40 $\mu g/m^3$ [3] (with political agreement to reduce it to 20 $\mu g/m^3$ by 2030). In this context, Milan, a city with high private motorization and unfavorable orographic conditions for dispersion, is an ideal laboratory to study the evolution of NO_2 in the urban environment.

Since 2012, the Lombard capital has introduced road pricing measures (Area C) and, most importantly, the Low-Emission Zone Area B, active since February 25, 2019 [4] and progressively tightened since October 1, 2022, which restricts access to the most polluting vehicles on the entire Smart_city municipal territory. Understanding whether and how much these policies have affected NO_2 levels, distinguishing the inside of the LTZ from the outer belt, is crucial for calibrating further sustainable mobility interventions.

This paper analyzes the spatiotemporal trend of NO_2 in Milan from 2018 to 2025, with three objectives:

- describe pollutant trends in the three geographic domains-Area C, Area B and outer belt
- Quantify the effectiveness of Area B by comparing concentrations "inside" and "outside" the LTZ before and after key dates

- model construction for forecasting purposes

1.1 Health Risks of Nitrogen Dioxide (NO_2)

Nitrogen dioxide (NO_2) is one of the main air pollutants produced by the combustion of fossil fuels, particularly from vehicle traffic, heating systems, and the use of gas stoves. It represents a serious threat to human health, especially in urban areas. Biologically, NO_2 is a highly oxidizing gas that, once inhaled, penetrates deep into the respiratory tract, generating free radicals, inflammation, and oxidative stress [5]. These processes disrupt the normal functioning of the respiratory epithelium and impair the body's ability to defend itself against infections, thereby promoting the onset and worsening of various diseases.

Acute effects of NO_2 exposure include irritation of the eyes, nose, and throat, coughing, respiratory distress, and bronchitis [6]. In more severe cases, especially at high concentrations, it can cause pulmonary edema and become life-threatening. Even more concerning are the long-term effects: scientific studies show that chronic exposure—even at concentrations below current legal limits—is associated with an increased incidence of asthma (particularly in children), worsening conditions in individuals with chronic obstructive pulmonary disease (COPD), and a rise in mortality from respiratory and cardiovascular diseases [7].

Recent meta-analyses have quantified a significant increase in mortality for every 10 µg/m^3 [7] increase in NO_2 concentration. Other studies highlight risks to reproductive health and fetal development [**?**, 5], showing correlations between exposure during pregnancy and higher probabilities of miscarriage, low birth weight, and preterm birth. There is also growing concern about neuropsychological implications [5]: mounting evidence links NO_2 pollution to cognitive disorders in children and to increased symptoms of depression and anxiety in adults.

Another area of concern is indoor exposure. In many households, the use of gas stoves and heaters can lead to NO_2 concentrations equal to or greater than those found outdoors in the most polluted cities [8]. In such cases, continuous exposure can directly impact children's health by contributing to the development of new asthma cases and worsening pre-existing conditions. In the United States alone, it is estimated that around 50,000 childhood asthma [9] cases each year are attributable to gas stove use.

These findings make it clear that NO_2 is not merely an indicator of urban traffic levels, but a true toxic agent with systemic effects. For this reason, the World Health Organization has recently lowered the recommended annual average exposure limit from 40 µg/m^3 to just 10 µg/m^3 [2], highlighting that even levels once considered "safe" may cause harm to human health.

2 Data

2.1 EEA

For the analysis of air quality in the city of Milan, the first step was to acquire data through the API made available by the European Environment Agency (EEA) [10]. Specifically, we carried out data collection from January 1, 2018 to June 1, 2025, focusing on the concentration of nitrogen dioxide (NO_2) detected by the various monitoring stations located in the Milan metropolitan area and surrounding areas.

Choice of pollutant: NO_2 Nitrogen dioxide (NO_2) is a pollutant gas mainly produced by the combustion of fossil fuels, especially in vehicular traffic and industrial processes. It is known for its adverse effects on human health, particularly on the respiratory system, and contributes to the formation of acid rain and tropospheric ozone [11]. For this reason, it is a good indicator of urban air quality.

Station Location and Identification of Area B Each survey provided by the EEA is associated with geographic coordinates (latitude and longitude), which allowed us to accurately map the location of monitoring stations. We then compared the coordinates with the boundaries of the Area B Restricted Traffic Zone of the City of Milan [12] by implementing a binary variable named $inside\_areaB$, whose value is:

- True if the station is located within the boundaries of Area B;
- False if the station is located outside of it.

This distinction will be useful in subsequent analytical steps, to assess any differences in NO_2 exposure based on location versus LTZ.

Data cleaning The original dataset included a total of 1,036,783 rows. Before proceeding with the analysis, we performed a check for the presence of missing values. We found that 50,554 rows contained null values, corresponding to 4.8% of the total. Considering the low incidence of missing data, we decided to remove these rows directly, maintaining the consistency and quality of the dataset.

The result of this phase is a first clean dataset, containing geo-referenced information on the concentration of NO_2 in the Milan territory, with labeling that distinguishes detections inside and outside Area B.

Feature	Description
`Samplingpoint`	Detection station name
`Pollutant`	Numerical code of the pollutant detected
`Start`	Start of the hourly interval of the survey
`End`	End of the hourly interval of the survey
`Value`	Detected value of the pollutant
`Unit`	Unit of measurement of the detected value
`AggType`	Time aggregation type (e.g., hourly, daily)
`Validity`	Data validity status (e.g., validated/not validated)
`Verification`	Status of data verification by the institution
`ResultTime`	Time at which the data was recorded in the system
`DataCapture`	Timestamp of data capture
`FkObservationLog`	Observation identification code in the log
`datetime`	Exact date and time of the survey
`EoI`	EoI code of the air quality station.
`lon`	Longitude of the station
`lat`	Latitude of the station
`geometry`	Geometry used for spatial mapping
`inside_areaB`	Boolean variable: True if the station is inside Area B, False otherwise

Table 1: Description of the variables contained in the first dataset

Ending up with the stations located in:

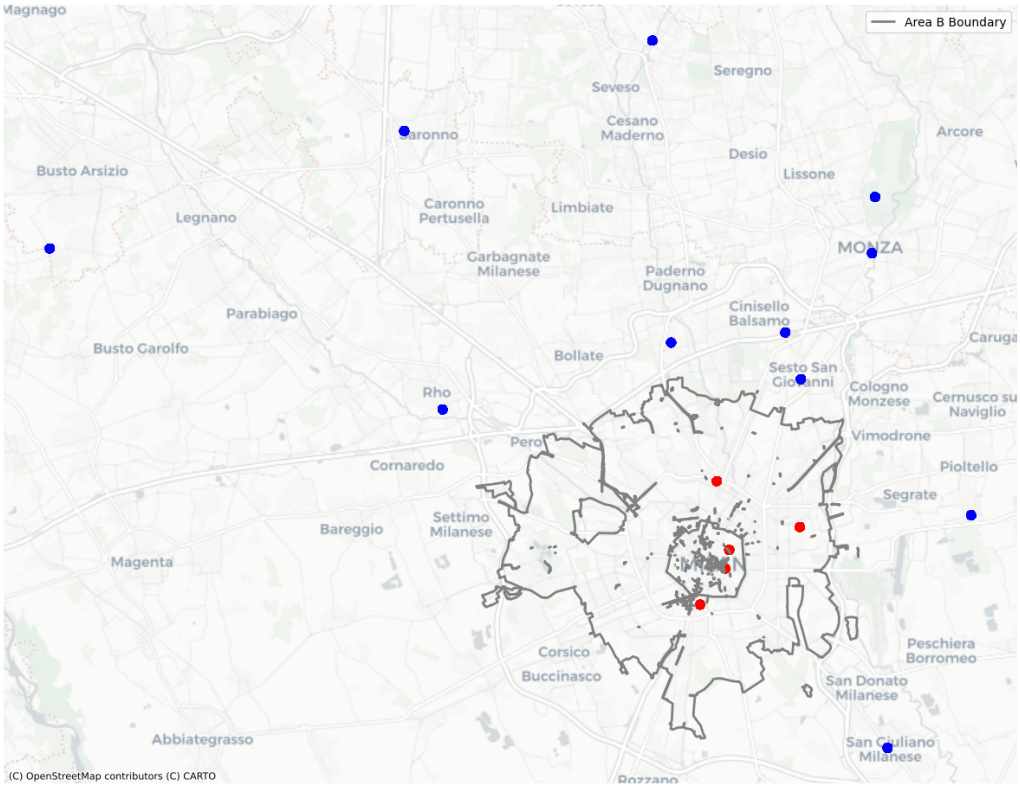


Figure 1: Stations' Map

2.2 ERA5

The second dataset was constructed from data provided by the ERA5 service, made available via API [13]. ERA5 is a major source of climate and weather data globally, based on atmospheric reanalysis produced by the European Centre for Medium-term Weather Forecasts (ECMWF). The time interval of the collected data was kept consistent with that of the first dataset, i.e., from January 1, 2018 to June 1, 2025, in order to ensure compatibility in subsequent integration steps.

The obtained dataset contains the following variables:

- datetime: the date and time of the survey;
- Samplingpoint: the name of the station with which the detection is associated;
- wind_speed: wind speed;
- RH: relative humidity.

The inclusion of the variables wind_speed and RH is motivated by their significant role in air quality dynamics:

- Wind speed: is a major determinant in the dispersion of air pollutants. More intense wind tends to dilute and transport pollutants, reducing their local concentration. Conversely, in calm wind conditions, pollutants may stagnate in the atmosphere, leading to higher concentration values [14].
- Relative humidity (RH): affects the chemical and physical processes that regulate the formation and behavior of pollutants. For example, it can promote the transformation of nitrogen dioxide (NO_2) into other reactive forms, or influence the formation of secondary particulate matter [15]. In addition, high humidity conditions are often associated with thermal inversion phenomena, which in turn can trap pollutants in the lower layers of the atmosphere [16].

Also in the second dataset is the variable Samplingpoint, which represents the name of the survey station. This variable will be used as a connecting key to perform the merge between the two datasets, allowing meteorological information to be merged with air quality information. This integration will make possible a more complete analysis, in which NO_2 concentrations can also be interpreted in light of local atmospheric conditions.

2.3 Merge

Once the two datasets were prepared, it was possible to merge them using the Samplingpoint variable, which is present in both sources. The merge operation produced a final integrated dataset, combining air quality data with meteorological data for each sampling point and timestamp.

In the final dataset, we chose to keep only those variables considered most relevant to the analysis. The columns selected are:

- datetime: date and time of the survey;
- Samplingpoint: measuring station name;
- Value: NO_2 concentration detected;
- inside_areaB: Boolean indicating whether the station is within Area B;
- wind_speed: wind speed;
- RH: relative humidity;
- hour: time of detection (extracted from datetime);
- dayofyear: day of the year (extracted from datetime, useful for seasonal analysis).

This selection allows the dataset to be kept light and focused on the essential components for the analysis of pollution in relation to weather and geographic location.

During the merge, the loss of 361,096 (35%) rows and some detection stations from the initial dataset (based on EEA data) was observed. Finally obtaining these stations:

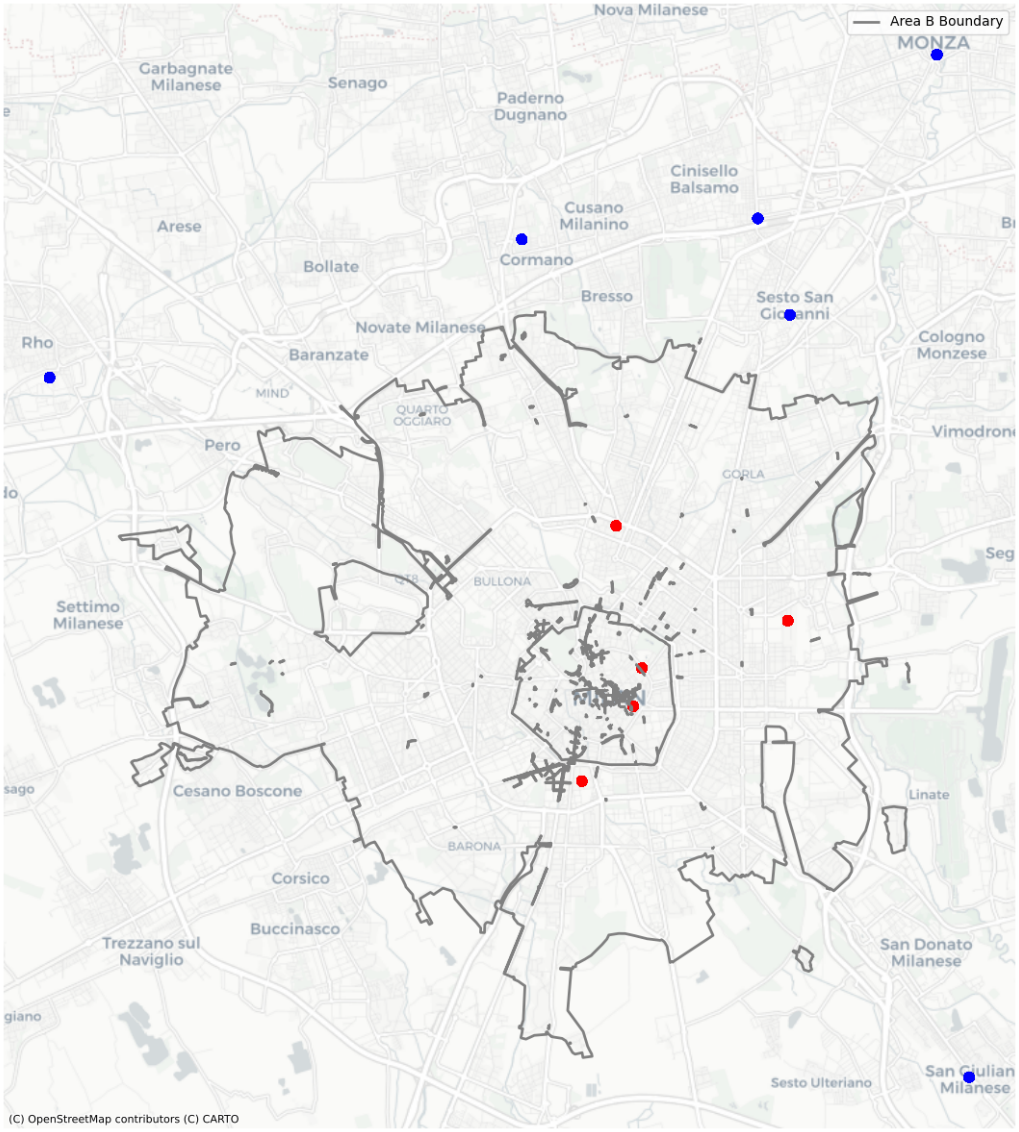


Figure 2: Final Stations' Map

Considering that the focus of the analysis is primarily on the city of Milan and, in particular, on the effect of Area B, the loss of external detections can be considered acceptable. Although they contain potentially useful information for broader or comparative analyses, such data are not essential in the specific context of this study.

3 Analysis

In this section we are going to evaluate several models

3.1 Temporal analysis

The objective of this section is to evaluate the effectiveness of traffic restrictions introduced in Milan (Area B from February 25, 2019 and subsequent tightening from October 1, 2024) in reducing the urban population's exposure to air pollutants, particularly nitrogen dioxide (NO_2). Monitoring air quality trends at environmental policies is crucial for three reasons:

- Public Health The NO_2 is associated with acute and chronic respiratory effects; quantifying any decreases allows estimating health benefits and avoided costs.

- Policy evaluation Measurable indicators allow one to understand whether traffic restrictions achieve their stated purpose and, if so, to calibrate future interventions.

For this reason we divided the time series into three regulatory periods:

- before Area B
- Area B – phase 1
- Area B – phase 2

and calculated the daily average concentration of NO_2 at inside (within Area B boundaries) and outside (control areas) stations for each phase

Period	**Inside (µg/m³)**	**Outside (µg/m³)**
pre_areaB	45.32	50.16
areaB_phase1	37.10	37.19
areaB_phase2	36.15	37.97

Table 2: Average PM concentrations for different periods (µg/m³)

Observations In the area outside Area B the mean concentration drops from 45.3 µg/m³ to 37.1 µg/m³ (-18%).

Inside Area B the drop is even more pronounced in absolute terms (-13 µg/m³) but similar in relative terms (-26%).

Outside Area B the average rises from 37.1 → 36.2 µg/m³ (-3%), an indication that the second set of restrictions had an additional impact in the outside area as well, albeit smaller.

Inside increases instead from 37.2 → 38.0 µg/m³ (+2%).

The results indicate that the introduction of Area B coincided with a substantial drop in average NO_2 levels both inside and outside the zone. In contrast, the further tightening in 2024 shows skillful refinement: air remains outside the regulated area while worsening slightly inside.

3.2 Statistical Model

3.2.1 Time series analysis on seasonality

To analyze the time trend of NO_2 concentration and evaluate the possibility of future predictions, a time series-based model was constructed. Specifically, it was decided to separate the study area into two distinct series:

- inside: represents the monthly average concentration of NO_2 detected by stations located inside Area B;
- outside: represents the monthly average concentration detected outside Area B.

This subdivision makes it possible to assess possible seasonal differences, trends and different behaviors depending on the geographical location with respect to the restricted area.

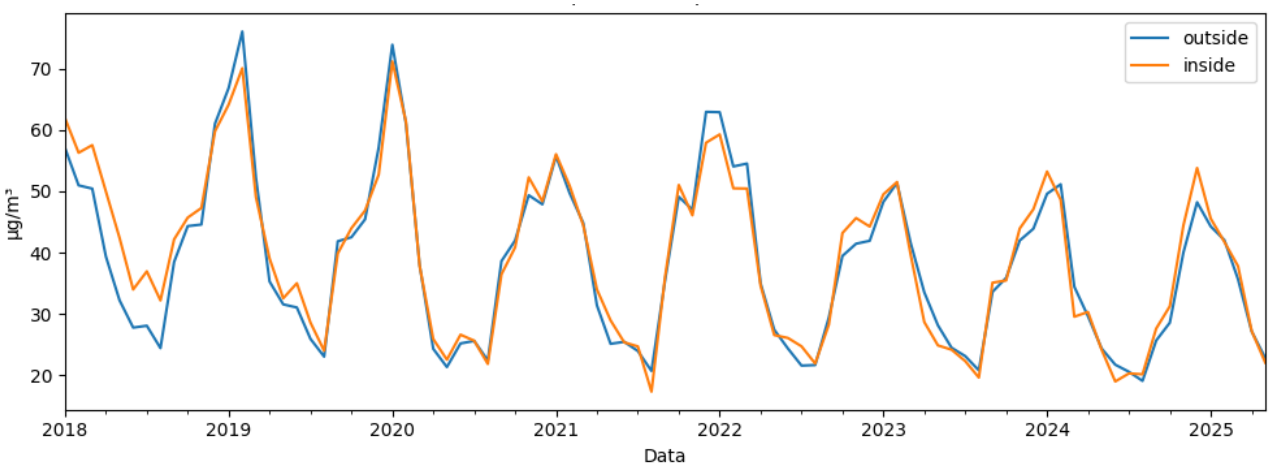


Figure 3: Series

The first visualization of the two time series showed a well-defined seasonality, with recurring peaks and troughs at seasonal changes (winter/summer), as shown in the graph shown in Figure 3.

To properly construct and evaluate the model, a temporal partitioning of the dataset was performed:

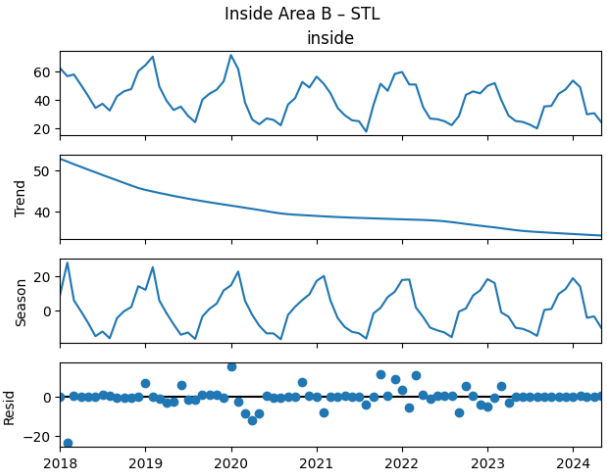
- The training set was defined by including data up to the end of May 2024.
- The test set includes the entire year from June 2024 to May 2025.

At first, it was considered to use all data up to December 31, 2024 as the training set, reserving 2025 for testing. However, the fact that the data available for 2025 stopped in June made the test unrepresentative in terms of the annual cycle. For this reason, it was preferred to advance the end of the training to May 2024, thus ensuring a full calendar year for the testing phase.

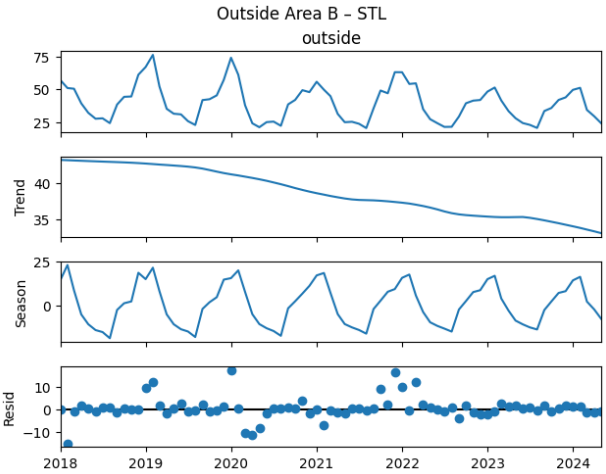
This choice allows the model to learn on the basis of a full seasonality and to be tested with data also distributed over all seasons, a key condition for properly assessing forecasting performance over time.

3.2.2 STL check

To build a reliable time series model, it is essential to work with stationary series, that is, series in which mean, variance and seasonality are constant over time or have been appropriately removed by transformations. This section describes the steps taken to obtain stationary of the two series of interest (inside and outside) in preparation for modeling.



(a) STL_inside



(b) STL_outside

Logarithmic Transformation From the graphical analysis of the 4a and 4b images, a seasonal adjustment of order 12 was applied right away and followed by a differentiation for trend removal but this led to high

variance graphs with many peaks, with therefore problem in stationary in variance, consequently it was decided as a first step to apply a logarithmic transformation so as to fix this component.

It is important to apply it before any other transformation, since the logarithmic function is not defined for negative values, and differentiation could generate values less than or equal to zero.

Deseasonalization Next, the two series were seasonally adjusted by applying a differentiation of order 12, corresponding to the annual component of seasonality (monthly data):

This operation removes the periodic fluctuations observed in the data, which are regularly repeated each year;

It is particularly relevant in this highly seasonally influenced context (winter with higher NO_2 concentration, summer with lower values).

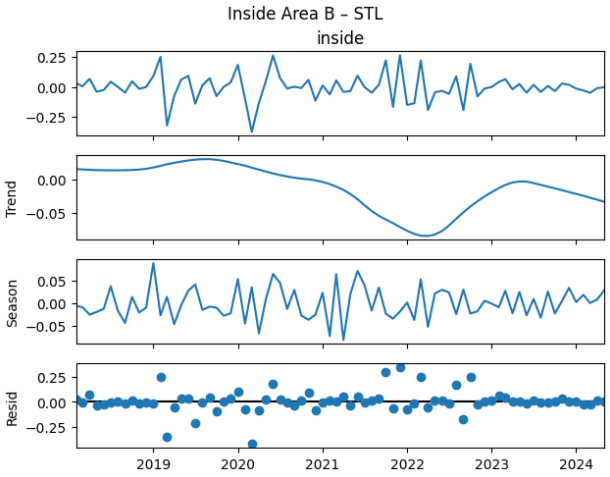
Trend removal After adjusting the variance and seasonality component, it was observed that the series still had residual trends, indicative of non-stationarity in the mean.

To correct for this component:

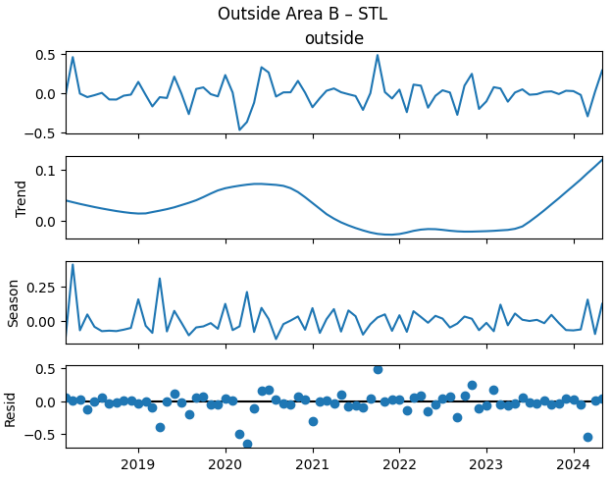
- A differentiation of order 1 was applied to the inside series, which was sufficient to remove the linear growth/decrease component;
- For the outside series, a differentiation order of 2 was necessary because of the greater complexity of the observed trend.

Graphical analysis of the transformed series after each step confirmed the progressive stabilization of the structural components. At the end of the transformations, the series were found to be stationary and thus suitable for the next modeling step.

After all transformations, the resulting graphs are as follows **??** and **??**



(a) STD series — inside



(b) STD series — outside

3.2.3 ARIMA

After transforming the series to ensure stationarity, the correlograms were analyzed by graphing the autocorrelation functions (ACF) and partial autocorrelation functions (PACF).

The objective of this step was to identify the optimal order for the parameters p (autoregressive) and q (moving average) of the ARIMA model. The analysis was conducted on the two transformed series (log + seasonal adjustment), without applying differentiation preliminarily, as the d parameter related to differentiation will be integrated directly into the definition of the ARIMA model.

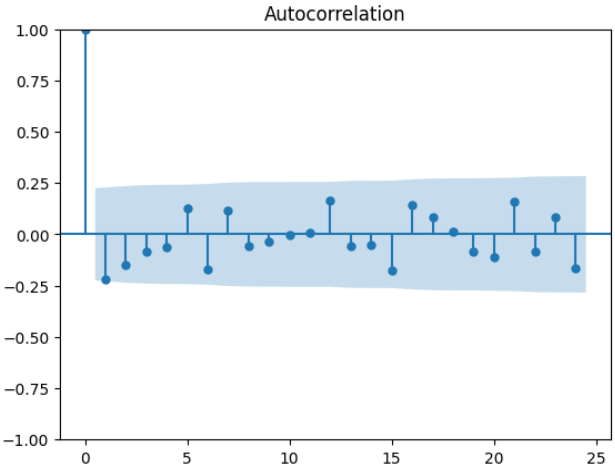


Figure 6: ACF in

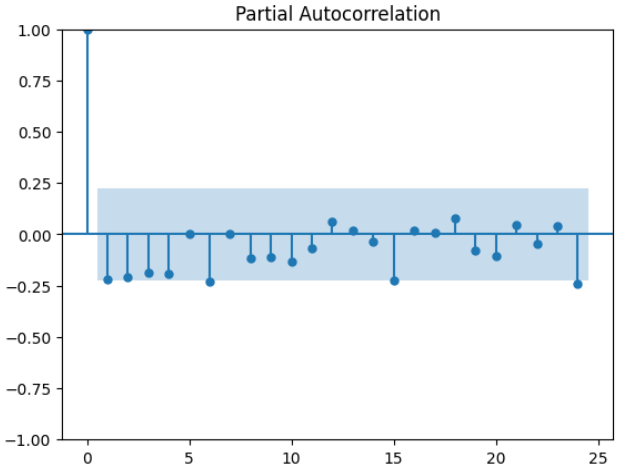


Figure 7: PACF in

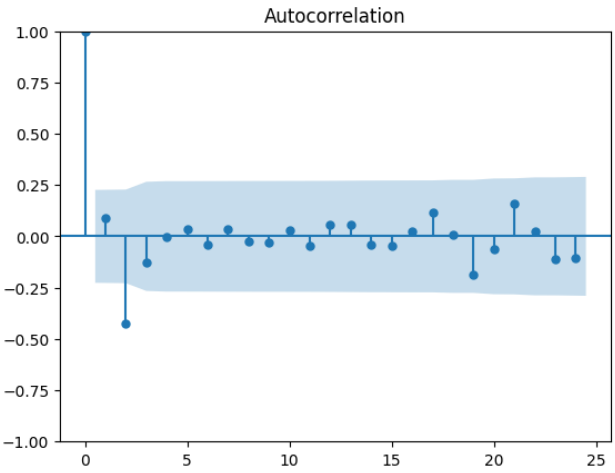


Figure 8: ACF out

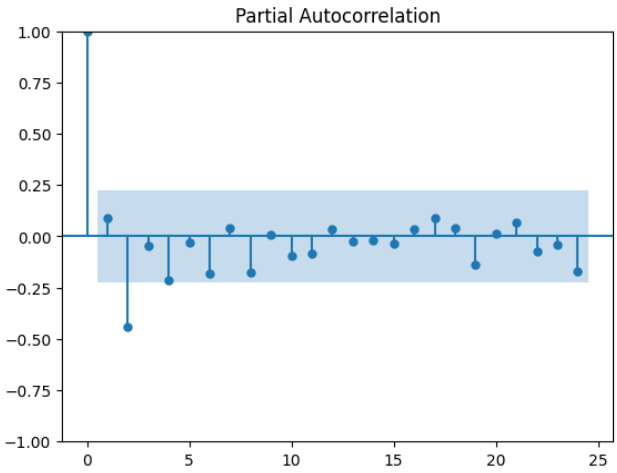


Figure 9: PACF out

Results of ACF/PACF analysis The graphical analysis showed, for both series (inside and outside), similar behavior:

ACF and PACF cancel out after the first lag, indicating the presence of an autoregressive and a moving average component, both of order 1;

This behavior is consistent with a simple but effective ARIMA model capable of capturing the information contained in short-term correlations.

Choice of model Based on these observations, the following models were chosen:

- ARIMA(1,1,1) for the inside series: a single differentiation was sufficient to achieve stationarity in the mean, as verified in the previous step;
- ARIMA(1,2,1) for the outside series: a differentiation of order 2 was maintained, which was necessary to completely eliminate the trend.

These models are found to be adequate to capture the essential dynamics of the time series and are now ready to be used in the forecasting phase.

3.2.4 Predictions

Once the ARIMA models were trained on the two time series (‘inside‘ and ‘outside‘), predictions were made on the test set, i.e., the interval between June 2024 and May 2025. The objective was to evaluate the effectiveness of the model in capturing the future trend of NO_2 concentration, also considering the seasonal variability already observed in the historical data.

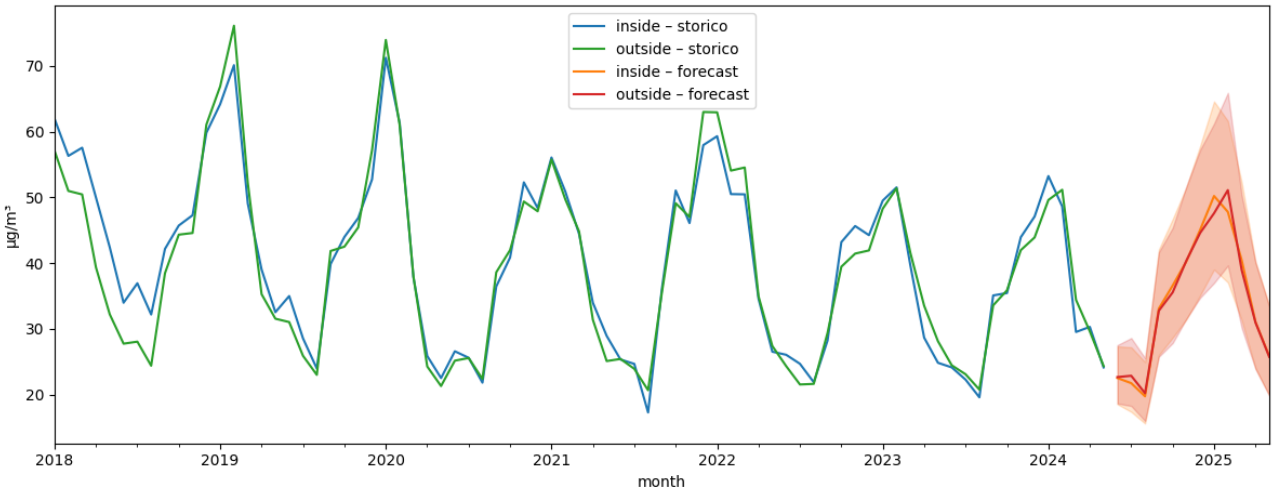


Figure 10: Forecasts 2025

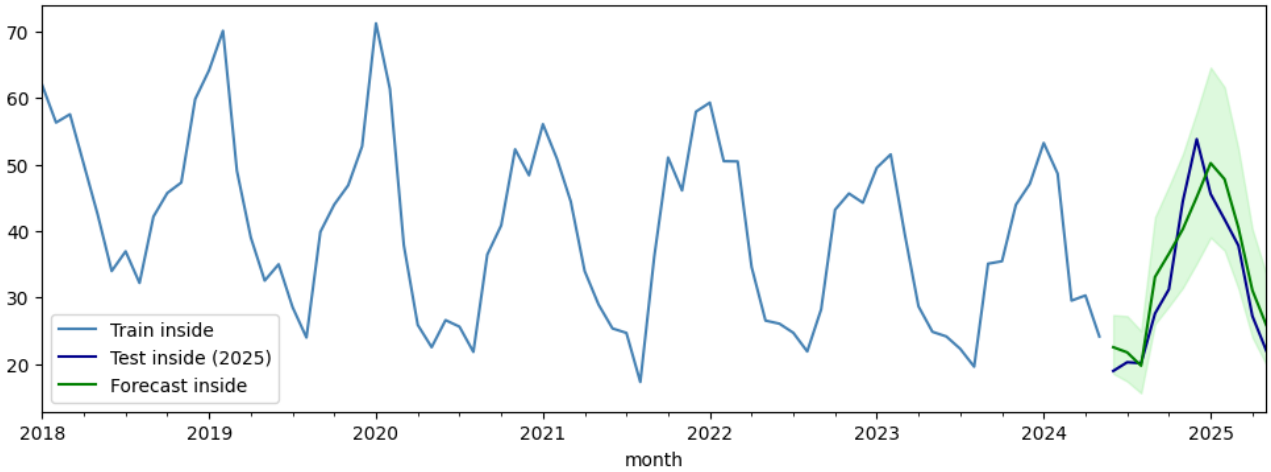


Figure 11: Forecasts 2025 inside

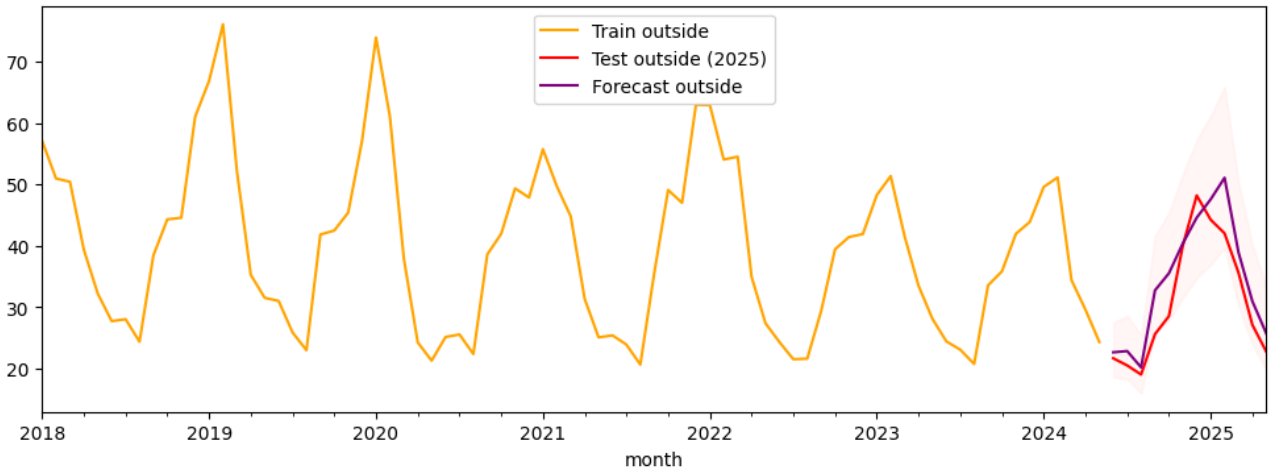


Figure 12: Forecasts 2025 out

To quantify the goodness of predictions, the following indicators were used:

- **MAE (Mean Absolute Error)**

- **RMSE (Root Mean Squared Error)**
- **MAPE (Mean Absolute Percentage Error)**
- **Accuracy**: forecast accuracy calculated as 100% − MAPE.

Metric	Inside	Outside
MAE	4.19	3.76
RMSE	4.68	4.55
MAPE	12.86%	12.36%
Accuracy	87.14%	87.64%

Table 3: Performance evaluation of the ARIMA model for the two series

The results obtained show a good level of accuracy for both models, with MAPE values around 12%. In particular:

- The forecasts for the stations **inside Area B** are slightly less accurate than those outside, but still satisfactory;
- The metrics suggest that the chosen ARIMA models are adequate to capture seasonality and residual trend in the two series;
- The mean absolute error (MAE) of less than 5 $\mu g/m^3$ in both cases makes the models potentially useful for operational applications as well.

These results confirm the validity of the approach taken for modeling NO_2 concentrations as a function of time.

A forecast from May 2025 to June 2026 was also made to see what the future trend might look like.

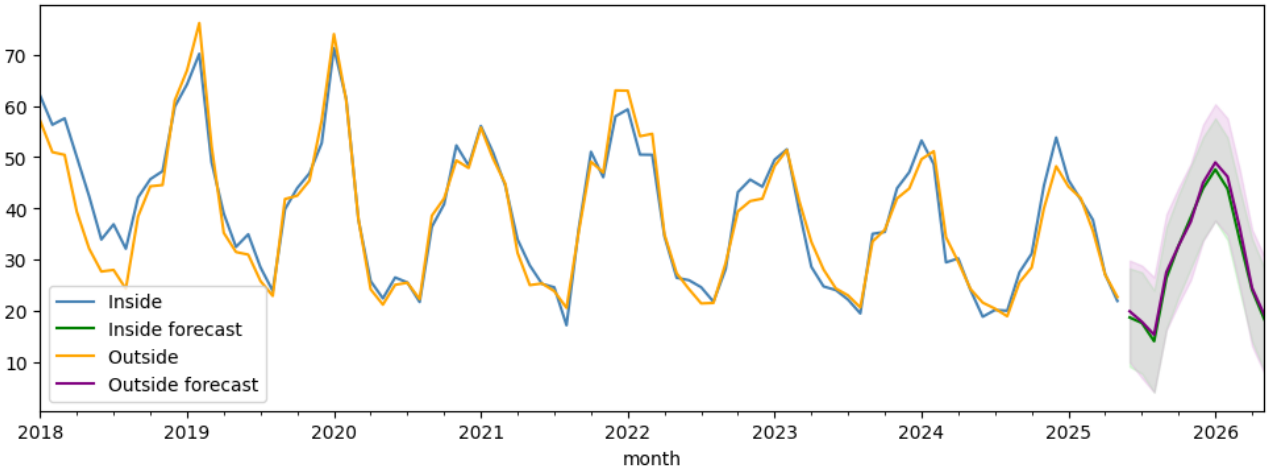


Figure 13: Forecasts 2026

3.3 Machine Learning Model

In addition to the time-series model, a machine learning model was also used to have a comparison, which, unlike a canonical statistical approach (such as ARIMA or linear autoregressions), easily integrates heterogeneous variables, captures nonlinear relationships, and updates its predictions recursively by exploiting newly produced outputs. This makes it possible to combine historical indicators, weather factors, and calendar variables without having to impose assumptions of linearity or stationarity, resulting in more robust projections when the context undergoes rapid changes or policies intervene that alter the underlying dynamics.

3.3.1 Feature Engineering

Hourly measurements were aggregated to daily average by station and then to overall average (**AreaB_mean**). Information holes were found on some stations subsequently filled by replacing each missing value with the average of the corresponding day of the year.

In addition, the following were created:

- lag of the one-, seven-, and fourteen-day NO_2;
- moving averages over seven- and thirty-day windows computed only on data prior to the day under consideration;
- trigonometric representations of annual and weekly seasonality;
- "weekend" flag,
- "normative" flag, of the type: ante Area B (0), first tightening 2019 (1), tightening 2022 (2), and future tightening 2025 (3).
- NO_2 value of the previous day.

3.3.2 Modeling

The algorithm chosen is XGBoost, embedded in a pipeline that standardizes numerical variables and encodes categorical ones with One-Hot Encoding. The sample was divided temporally: training until December 31, 2024, and testing from January to May 2025. A grid search with TimeSeriesSplit identified the optimal set of hyper-parameters.

In a second iteration, the least informative calendar variables (month and day of the week) were removed, obtaining a very slight improvement in performance.

3.3.3 Results and forecasts

.

Over the validation period, the second iteration of the model obtained an RMSE of 6.76 µg/m^3 and an MAE of 5.33 µg/m^3.

The weights that each variable contributed to the predictions were measured and the highest contribution comes from the NO_2 value of the previous day, followed by the "weekend" flag. In addition, it was also possible to have the previous day's concentration value for the forecasts by setting recursivity on the model, so each forecast is also based on the previous one.

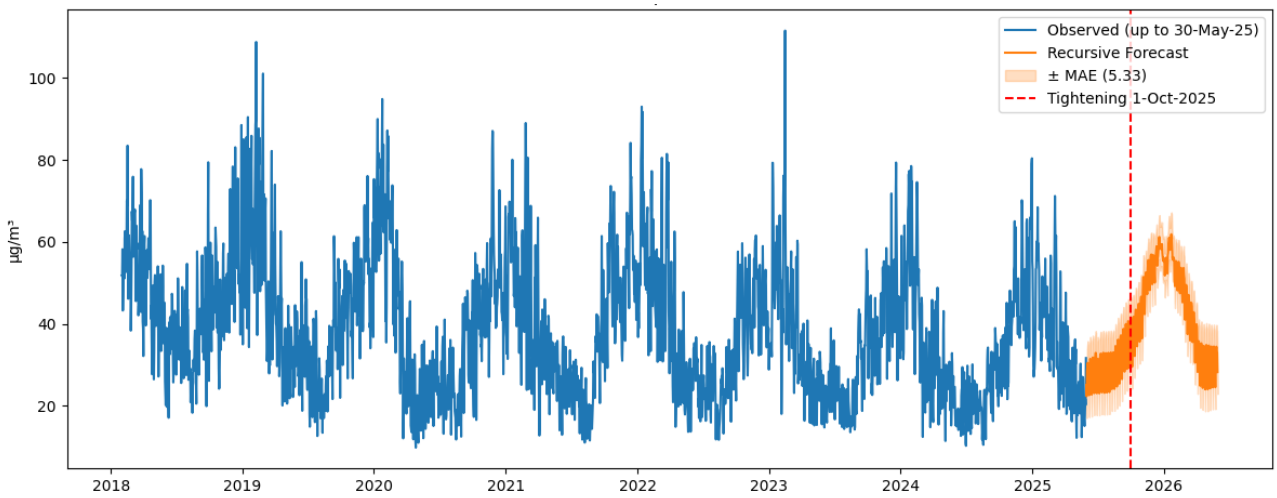


Figure 14: Forecasts 2026 XGBoost

The graph consists of a band of uncertainty represented by the MFA obtained in validation to emphasize possible fluctuations.

The projection suggests a winter peak between December 2025 and February 2026. The model follows the trend/seasonality of past observations suggesting a decline in NO_2 concentrations given by the upcoming tightening of the bans on Oct. 1, 2025, a sign that the ban on the most polluting vehicles could lower the bottom line. Summer 2026 is expected to be stably below 30 µg/m^3, thus well below the annual EU limit of 40 µg/m^3.

Finally, the average error of about 5 µg/m^3 appears small compared to the observed daily amplitude (20-80 µg/m^3) and makes the model suitable for annual scenarios

3.4 Comparison

In this section we compare the performance of the two forecasting approaches taken: the ARIMA statistical model, which is classical for time series, and the XGBoost machine learning model, which is more flexible in dealing with heterogeneity and nonlinearity.

Methodological Approach The ARIMA model exclusively exploits autocorrelation in the time series, relying on assumptions of stationarity and linearity. It is particularly suitable for stable contexts with predictable dynamics over time. In contrast, XGBoost is a supervised learning algorithm capable of integrating a variety of explanatory features, capturing nonlinear relationships and adapting to more rapid context variations such as new environmental policies.

Performance comparison . To compare the models, we consider error metrics over the test period (June 2024 - May 2025 for ARIMA, January - May 2025 for XGBoost):

Although ARIMA shows slightly better performance in terms of absolute error (MAE and RMSE), it should be noted that:

- The ARIMA model is specific to the inside and outside series and does not integrate weather information or policy signals.
- The XGBoost model, while having a slightly higher mean error (MAE = 5.33), shows greater fit and robustness due to the inclusion of exogenous factors.
- The ML approach also allowed for a recursive forecast that more realistically reflects the expected downward trend following the tightening of restrictions in 2025.

Final considerations . The ARIMA model is effective for short-term forecasts in stable and well-structured contexts, while XGBoost is better suited for dynamic scenarios with multiple factors at play. From the perspective of urban planning and evaluating the effectiveness of new policies, the machine learning approach is potentially more useful because of its flexibility and ability to incorporate context signals.

4 Conclusion

Summary of results . The comparison between ARIMA and XGBoost showed distinct advantages and limitations. ARIMA shows higher accuracy in stable contexts devoid of exogenous variables, with lower average MAE and RMSE. However, XGBoost stands out for its ability to integrate exogenous features, providing better predictive robustness in dynamic scenarios such as urban ones.

Future Perspectives . The advancement of smart cities requires a systems approach: smart infrastructure, shared public policies, advanced analytical tools and active citizenship. In this context, the synergy between statistical models (ARIMA) and machine learning (XGBoost) is an effective strategy to generate insights useful for urban planning, resilience and sustainability.

The integration of predictive algorithms with transparent governance and "smart urban design" enables a shift from isolated solutions to integrated systems capable of significantly improving the quality of life in the cities of the future.

References

[1] X. Chen, L. Qi, S. Li, and X. Duan, "Long-term no_2 exposure and mortality: A comprehensive meta-analysis," *Environmental Pollution*, vol. 341, p. 122971, 2024.

[2] World Health Organization, "Who global air quality guidelines: Particulate matter ($pm_{2.5}$ and pm_{10}), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide," Geneva, 2021. [Online]. Available: https://www.who.int/publications/i/item/9789240034228

[3] European Parliament and Council of the European Union, "Directive (eu) 2024/2881 of 23 october 2024 on ambient air quality and cleaner air for europe (recast)," *Official Journal of the European Union*, L 2024/2881, 2024. [Online]. Available: https://eur-lex.europa.eu/eli/dir/2024/2881/oj/eng

[4] Comune di Milano, "Deliberazione della giunta comunale n. 1366 del 2 agosto 2018: Istituzione della zona a traffico limitato "area b"," Milano, 2018. [Online]. Available: https://delibere.comune.milano.it/documents/173794

[5] Health and Environment Alliance, "Science review: The health impacts of nitrogen dioxide (no_2)," 2023. [Online]. Available: https://www.env-health.org/science-review-the-health-impacts-of-nitrogen-dioxide-no2

[6] American Lung Association, "Nitrogen dioxide: Health effects of nitrogen dioxide pollution," 2025, webpage, accessed 11 Jul 2025. [Online]. Available: https://www.lung.org/clean-air/outdoors/what-makes-air-unhealthy/nitrogen-dioxide

[7] S. Huang, H. Li, M. Wang, Y. Qian, K. Steenland, W. M. Caudle, Y. Liu, J. Sarnat, S. Papatheodorou, and L. Shi, "Long-term exposure to nitrogen dioxide and mortality: A systematic review and meta-analysis," *Science of the Total Environment*, vol. 776, p. 145968, 2021.

[8] World Health Organization, *WHO Guidelines for Indoor Air Quality: Selected Pollutants – Nitrogen Dioxide*. Geneva: World Health Organization, 2010. [Online]. Available: https://www.ncbi.nlm.nih.gov/books/NBK138707

[9] Y. Kashtan, M. Nicholson, C. J. Finnegan, Z. Ouyang, A. Garg, E. D. Lebel, S. T. Rowland, D. R. Michanowicz, J. Herrera, K. C. Nadeau, and R. B. Jackson, "Nitrogen dioxide exposure, health outcomes, and associated demographic disparities due to gas and propane combustion by u.s. stoves," *Science Advances*, vol. 10, no. 18, p. eadm8680, 2024.

[10] European Environment Agency, "Air quality download service api: Air quality measurements time series," Dataset, published 14 Apr 2025, 2025. [Online]. Available: https://www.eea.europa.eu/en/datahub/datahubitem-view/778ef9f5-6293-4846-badd-56a29c70880d

[11] ——, "Europe's air quality status 2023," EEA Briefing, 2023. [Online]. Available: https://www.eea.europa.eu/publications/europes-air-quality-status-2023

[12] Comune di Milano, "Area b – perimetrazione e varchi (geoportal dataset id a17c3edd56164bc793615c298babd286)," Open GIS layer, accessed 11 Jul 2025, 2025. [Online]. Available: https://geoportale.comune.milano.it/portal/home/item.html?id=a17c3edd56164bc793615c298babd286

[13] H. Hersbach, B. Bell, P. Berrisford, and et al., "The era5 global reanalysis," *Quarterly Journal of the Royal Meteorological Society*, vol. 146, no. 730, pp. 1999–2049, 2020, dataset DOI: 10.24381/cds.adbb2d47 (Copernicus Climate Change Service).

[14] V. S. Carvalho, E. Pisoni, P. Thunis, D. Pernigotti, and M. Gerboles, "How wind speed and direction affect no_2 levels in european urban areas," *Atmospheric Environment*, vol. 189, pp. 385–399, 2018.

[15] H. Cheng, Y. Zhang, J. Duan, and K. He, "Relative humidity and multiphase chemistry: Implications for secondary nitrate formation," *Science of the Total Environment*, vol. 755, p. 142523, 2021.

[16] Q. Liu, M. Zhang, W. Gao, and B. Hu, "Impacts of thermal inversion on air quality in the po valley: Observation and modeling," *Environmental Pollution*, vol. 254, p. 113093, 2019.