Predicting Near-Road Air Quality Using Artificial Neural Networks: Exploring Traffic-Related Influences

Yihong Ning

Abstract

This study explores the use of Artificial Neural Networks (ANN) to predict near-road air pollutant concentrations, including PM1.0, PM2.5, PM10, and NO_2, based on traffic and environmental variables. Data collected at a high-traffic site in downtown Columbia, SC, were used to train and evaluate an ANN model. The model significantly outperformed traditional methods like Multiple Linear Regression (MLR) and Bayesian Model Averaging (BMA), achieving higher predictive accuracy and capturing complex nonlinear interactions. This project demonstrates the ANN's potential for advancing urban air quality management but also highlights its limitations, such as overfitting risks and dependency on high-quality data.

1. Introduction

Air quality is a critical determinant of public health and urban livability. According to the World Health Organization, outdoor air pollution contributed to approximately 4.2 million deaths globally in 2019, underscoring the urgent need to understand pollutant dynamics in high-exposure urban environments. Among various pollution sources, vehicular emissions are of particular concern, as road traffic significantly elevates ambient pollutant concentrations, impacting not only motorists but also pedestrians, cyclists, and nearby residents (Matthaios et al. 2022; Rangel et al. 2022; Rossi, Ceccato, and Gastaldi 2020). The rapid increase in vehicle numbers, coupled with intensified congestion and fuel consumption, leads to elevated levels of harmful pollutants—including PM1.0, PM2.5, PM10, NO2, SO2, CO, HC, and O3—near traffic intersections and along congested corridors (Askariyeh, Zietsman, and Autenrieth 2020; Hystad et al. 2013).

This project seeks to address the complexity inherent in near-road air quality prediction. Traditional methods often overlook the intricate, dynamic interplay of localized traffic conditions, variable meteorological factors, and complex atmospheric chemistry. By leveraging Artificial Neural Networks (ANNs), we aim to achieve more accurate and nuanced predictions of near-road pollutant concentrations. Our approach integrates high-resolution traffic data, meteorological variables, and pollutant measurements to capture nonlinear relationships and adapt to evolving, "noisy" datasets. Ultimately, we demonstrate how advanced machine learning techniques like ANNs can provide actionable insights for urban planners and policymakers, guiding interventions that safeguard public health and enhance environmental quality.

2. Relevant Work

Conventional air quality prediction efforts frequently rely on frequentist statistical methods using long-term averages from fixed monitoring stations. While these methods can identify general pollution trends, they often fail to represent near-road microenvironments where local traffic patterns, meteorological conditions, and chemical transformations interact in complex ways. For instance, Li, Hsu, and Tsay (2011) employed linear regression models to relate aerosol optical thickness to PM10 concentrations, but their approach struggled to capture the nonlinear and context-dependent relationships at street level.

In contrast, a variety of machine learning and advanced modeling techniques have emerged to address these limitations. Sophisticated statistical learning algorithms, deep learning frameworks, and hybrid models have shown improved predictive performance under diverse conditions (Du et al. 2023; Kadiyala, Kaur, and Kumar 2013; Lin et al. 2025). Among these, ANNs stand out due to their capacity to handle nonlinearities and integrate multiple data sources, such as pollutant concentrations, meteorological factors, and traffic variables (Fontes et al. 2013).

Despite the success of these methods, challenges remain. Some models excel at capturing complex patterns but perform poorly when data are irregular or limited—conditions common in near-road environments. ANNs, however, have demonstrated robustness to data gaps and unpredictability (Song and Han 2020). Still, a critical concern is their "black box" nature, which limits interpretability and may reduce their utility in policy contexts. Recent efforts have begun to address this challenge by incorporating interpretability techniques, such as Local Interpretable Model-Agnostic Explanations (LIME), enabling stakeholders to better understand the underlying factors driving model predictions. This study builds on these advances, aiming to leverage the strengths of ANNs while providing the interpretability and robustness needed to support evidence-based policy-making in urban air quality management.

3. Data

3.1 Study Site and Data Collection

Data were collected at a single sampling site near the intersection of Taylor Street and Pine Street in downtown Columbia, South Carolina (Figure 1). This location was chosen for two primary reasons. First, Taylor Street experiences a high Average Annual Daily Traffic (AADT) volume of approximately 22,200 vehicles (Anon n.d.), ensuring the capture of substantial variation in pollutant concentrations attributable to vehicular emissions. Second, the proximity of the site to the university campus allowed for secure placement of monitoring equipment, facilitating regular maintenance and minimizing the risk of data loss or sensor damage.

Figure 1. Geographic and Instrumental Setup for Traffic-Related Air Quality Monitoring

The instrumentation deployed at this site recorded fine-scale measurements to enable detailed modeling with Artificial Neural Networks (ANNs). Air quality sensors measured concentrations of PM1.0, PM2.5, PM10, and NO2 at two-minute intervals, along with environmental parameters such as temperature, humidity, and atmospheric pressure. These environmental variables were included because they influence pollutant behavior and dispersion, thus providing critical contextual information for ANN-based predictive modeling. Unfortunately, data for Sulphur Dioxide (SO2) and Carbon Monoxide (CO) were not captured due to equipment issues.

In parallel, traffic flow data were collected starting March 6, 2020. These records included one-minute timestamped measurements of traffic volumes, vehicle lengths, speeds (in mph), and gap times (in seconds) between consecutive vehicles. Each observation was linked to a unique timestamp, enabling high-resolution synchronization of traffic activities with pollutant and environmental measurements. This integrated dataset provides the foundation for constructing and training ANN models that can learn complex, nonlinear interactions between traffic behavior and pollutant dynamics.

3.2 Impact of COVID-19 Lockdowns on Data Availability

The data collection period coincided with the onset of the COVID-19 pandemic and the subsequent lockdown on March 17, 2020. This event introduced unexpected challenges in data continuity. Due to restrictions on access to the sensor site, the traffic sensors eventually ceased operation as their batteries were not replaced in time, while the PM sensors continued recording until April 13. NO_2 sensors failed the same day, resulting in datasets covering 135 hours of PM measurements and 84 hours of NO_2 concentrations, both overlapping with the available traffic data.

Despite these constraints, the synchronization of traffic and air quality measurements during the pre-lockdown and early lockdown periods remains invaluable. Although limited in duration, this dataset captures a rare scenario of abrupt changes in urban activity, potentially offering unique insights into how dramatic reductions in traffic volume can influence near-road pollutant levels. The constraints imposed by the pandemic underscore the challenges of continuous environmental monitoring and highlight the importance of flexible analytical approaches like ANNs that can gather useful patterns even from smaller, irregular datasets.

3.3 Variable Interpretations and Transformations

This study's primary objective is to use ANNs to predict near-road pollutant concentrations, specifically PM1.0, PM2.5, PM10 (log-transformed), and NO2, using a combination of traffic-related and environmental predictors. The variables fall into three categories:

Response Variables [1–4]:

[1] PM1.0: Log-transformed particulate matter (1.0 µm) concentration (µg/m³)

[2] PM2.5: Log-transformed particulate matter (2.5 µm) concentration (µg/m³)

[3] PM10: Log-transformed particulate matter (10 µm) concentration (µg/m³)

[4] NO2: Nitrogen dioxide concentration (ppm) (untransformed)

Traffic-Related Predictors [5–9]:

[5] Commuting-Vehicle: Number of passenger cars (sedans, SUVs) per 15-minute interval

[6] Truck: Number of trucks per 15-minute interval

[7] Multi-Trailer: Number of multi-trailer vehicles per 15-minute interval

[8] Speed: Average vehicle speed (mph) over the interval

[9] Gap: Average time gap (seconds) between vehicles

Environmental Predictors [10–12]:

[10] Temperature (°F)

[11] Humidity (%)

[12] Pressure (hPa)

Preliminary exploratory analyses revealed that raw PM distributions were right-skewed, potentially hindering the training stability and performance of ANNs. To address this, PM1.0, PM2.5, and PM10 concentrations were log-transformed. This transformation helps the ANN learn from data more efficiently by reducing the influence of extreme outliers and improving distributional properties. NO_2 did not exhibit severe skewness or variability that warranted transformation, and maintaining NO_2 in its original scale preserves interpretability for policy and planning applications.

All response and predictor variables were synchronized to 15-minute intervals. Aggregating the original high-frequency measurements (two-minute for pollutants and one-minute for traffic data) into 15-minute averages provides a balanced temporal resolution that captures traffic-induced fluctuations while ensuring robust training samples for the ANN. This approach smooths out transient anomalies and helps the ANN focus on stable relationships between predictors and target variables.

3.4 Data Preprocessing for ANN Training

Before training the ANN models, additional preprocessing steps were undertaken. Missing values were identified and, where appropriate, imputed using simple interpolation or removed if insufficient contextual data were available. This ensures that the ANN model is trained on a coherent dataset without breaks that could impede learning. Input features (traffic and environmental variables) were standardized or normalized to improve model training stability. This scaling ensures that all variables contribute more equally to the learning process, preventing the ANN from disproportionately focusing on features with larger numeric ranges.

By transforming response variables where necessary, standardizing predictors, and aligning data on a consistent temporal scale, we have curated a dataset suitable for robust ANN modeling. These efforts help ensure that the models can more effectively learn the complex, nonlinear relationships between traffic behavior, environmental conditions, and near-road air quality, actionable insights for policymakers and urban planners.

4. Methodology

4.1 Artificial Neural Network (ANN) Architecture

The core objective of this study is to leverage Artificial Neural Networks (ANNs) to model and predict near-road air pollutant concentrations (PM1.0, PM2.5, PM10, and NO_2) using traffic-related and environmental inputs. Given the complexity and nonlinearity inherent in near-road pollutant dynamics, a feedforward ANN was selected due to its flexibility in capturing complex patterns, as well as its ability to integrate diverse sets of input variables.

Network Design:

Input Layer: The number of input neurons corresponds to the total number of predictor variables (traffic counts for cars, trucks, multi-trailers; average speed, average gap; and environmental factors such as temperature, humidity, and pressure).

Hidden Layers:

The initial hidden layer consists of 128 neurons, followed by a second layer with 64 neurons, and a third layer with 32 neurons. Each hidden layer employs Rectified Linear Unit (ReLU) activation functions to maintain computational efficiency and help mitigate the vanishing gradient problem.

To reduce overfitting, dropout layers are introduced with a dropout rate of 0.3 after the first two hidden layers. Dropout stochastically removes a fraction of neurons during training, improving the model's generalization capability.

Output Layer: For predicting continuous pollutant concentrations (or their transformed values), a single linear neuron is utilized, providing a continuous output that can be directly compared to target pollutant levels.

This architecture strikes a balance between complexity and interpretability. The chosen number of layers and neurons was determined through preliminary experimentation and cross-validation, ensuring that the network is both expressive enough to capture complex relationships and efficient enough to converge within the available computational resources.

4.2 Training and Evaluation Setup

The dataset was split into training and testing subsets, with 80% of the data allocated for training and 20% reserved for independent testing. All predictor variables were scaled or standardized to improve numerical stability and accelerate training convergence. Additionally, missing values were handled via imputation or removed if necessary, and the response variables PM1.0, PM2.5, and PM10 were log-transformed to address skewed distributions, as described in the Data and Preprocessing section.

The ANN was implemented in Python with TensorFlow/Keras. The model was compiled using the Adam optimizer and a Mean Squared Error (MSE) loss function. Adam's adaptive learning rate capabilities facilitate efficient convergence, while MSE provides a straightforward measure of prediction error. Models were generally trained for up to 100 epochs with a batch size of 32. These default values strike a balance between computational efficiency and the model's ability to learn complex relationships. Adjustments to these parameters were explored as needed during preliminary experimentation. During training, 20% of the training data was used as a validation set (validation_split=0.2), enabling the monitoring of validation loss to prevent overfitting. Early stopping criteria with patience=10 and restore_best_weights=True halted training once the validation loss stopped improving. This safeguard ensures that the final model reflects the best validation performance, rather than a potentially overfit state near the end of training. In addition to dropout layers embedded in the ANN architecture, early stopping effectively served as a form of regularization. Together, these measures reduced the risk of overfitting, improving the model's ability to generalize beyond the training data.

Evaluation Metrics:
The following metrics were used to evaluate the predictive performance of the ANN:

- **Root Mean Squared Error (RMSE):** Provides a measure of the average prediction error magnitude. Lower RMSE values indicate better predictive accuracy.
- **Coefficient of Determination (R^2):** Measures how well the model explains the variance in the observed data. Higher R^2 values indicate better explanatory power.
- **Mean Absolute Error (MAE):** Used as a supplementary metric to provide an absolute measure of average prediction error.

Cross-Validation:

In addition to a single train-test split evaluation, k-fold cross-validation was employed in a separate routine to further assess model robustness. By systematically rotating the train-test partitions, cross-validation provided more stable estimates of model performance and helped confirm that the ANN could generalize across different subsets of the data.

4.3 Comparative Framework

While the ANN serves as the primary predictive model, this study also includes two additional modeling approaches to provide context and benchmarks, Multiple Linear Regression (MLR) and Bayesian Model Averaging (BMA). MLR was implemented as a baseline model. It assumes linear relationships between predictor variables and pollutant concentrations. Although MLR is straightforward and interpretable, it often struggles to capture nonlinearities and interactions present in near-road pollution datasets. The Bayesian approach incorporates prior knowledge, updates parameter distributions as new data become available, and quantifies uncertainties in model predictions. Bayesian Model Averaging (BMA) extends this principle by combining multiple candidate models into a weighted ensemble, effectively managing model uncertainty and improving predictive robustness. BMA can handle data-limited scenarios and multicollinearity among predictors, which is particularly relevant in near-road air quality modeling.

All three models (ANN, MLR, and BMA) were trained and tested on the same dataset, using identical response variables, predictor sets, and evaluation metrics. This uniformity ensures a fair comparison of their predictive capabilities. RMSE, R^2, and MAE served as the primary metrics for evaluating and contrasting model performances. Additionally, cross-validation procedures were applied to each method, providing a consistent measure of model stability and generalizability.

While MLR offers a clear, interpretable coefficient-based framework, and BMA provides a probabilistic view of model uncertainty, the ANN may initially appear as a "black box." To address this concern, interpretability methods, such as LIME (Local Interpretable Model-Agnostic Explanations), were applied to the ANN. LIME provides insights into how specific predictors influence individual predictions, thereby improving the ANN's transparency. By comparing results and interpretability features across these three modeling approaches, we gain a comprehensive understanding of their relative strengths and practical applicability.

5. Experiment and Results

5.1 Model Performance

Table 1 presents a summary of the ANN's performance metrics (R^2, RMSE) for NO_2, PM1.0, PM2.5, and PM10. The ANN outperformed conventional approaches, achieving higher R^2 and lower RMSE values, reflecting the model's ability to capture the complex, nonlinear relationships between traffic variables, environmental conditions, and near-road air pollutant concentrations.

Table 1. ANN Performance on Near-Road Air Pollutants

Air Pollutant	R^2	RMSE
NO_2	0.8676	15.0871
PM1.0	0.8814	0.1643
PM2.5	0.8890	0.1570
PM10	0.8836	0.1597

Note: PM values are log-transformed. NO_2 is in ppb, PM in µg/m³ (before transformation).

Figure 2 shows the ANN's predicted vs. actual values for each pollutant. The plots illustrate a close alignment around the 1:1 line, particularly for NO_2 and PM2.5, indicating that the ANN predictions closely approximate observed concentrations. Although some dispersion is observed in the PM1.0 and PM10 datasets, overall performance remains robust.

The strong performance of the ANN can be attributed to its capacity to model nonlinearities and interactions that traditional linear methods often fail to capture. Irregular traffic patterns, varying meteorological conditions, and complex atmospheric chemistry near roads introduce significant uncertainties. By leveraging multiple hidden layers and nonlinear activation functions, the ANN effectively learns these complexities, resulting in improved predictive accuracy and explanatory power.

While ANNs are often criticized as "black boxes," interpretability techniques such as LIME (Local Interpretable Model-Agnostic Explanations) offer insights into which features most strongly influence predictions for individual instances. LIME explanations for select predictions are shown in Figure 3, illustrating feature contributions toward increasing or decreasing predicted pollutant concentrations.

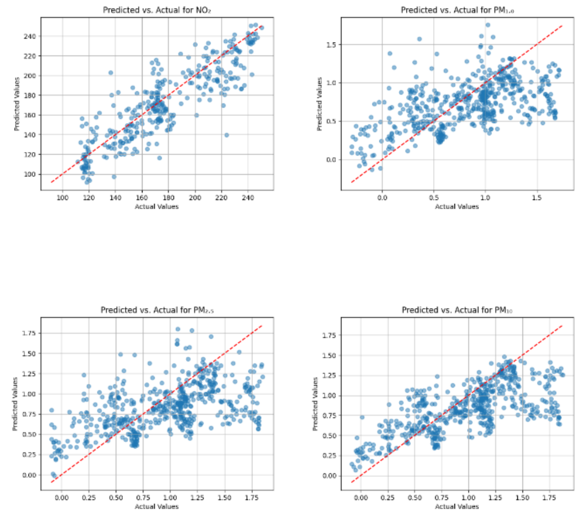

Figure 2. Predicted vs. Actual Plots for NO_2, PM1.0, PM2.5, and PM10

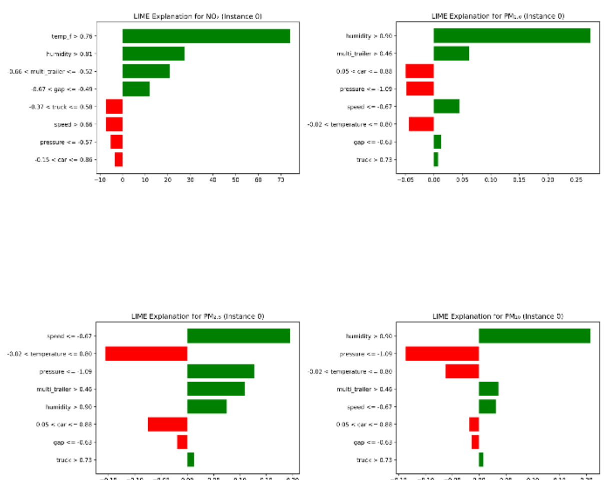

Figure 3. LIME Explanation for Predictions

Multi-trailer and truck counts consistently emerged as strong positive contributors to NO_2 and PM levels. Higher volumes of these vehicles are associated with increased emissions, reinforcing findings by previous research that heavy-duty vehicles significantly degrade near-road air quality (Mahesh, Ramadurai, and Shiva Nagendra 2018). Temperature and humidity often played substantial roles in increasing PM concentrations. Warmer, more humid conditions can facilitate particle formation and reduce pollutant dispersion. Lower speed and reduced gaps between vehicles (indicative of congestion) tended to increase pollutant levels, especially for NO_2. This aligns with the notion that stop-and-go traffic, rather than free-flowing movement, intensifies emissions due to frequent acceleration and idling events.

5.2 Comparison with Alternatives

To provide context, the ANN results were compared against two other modeling frameworks: Multiple Linear Regression (MLR) and Bayesian Model Averaging (BMA) (Table 2). MLR assumes linear relationships and was used as a simple baseline. Compared to the ANN, MLR exhibited substantially lower R^2 values and higher RMSE, indicating inadequate capture of nonlinearities and complex interactions. For instance, MLR models for PM often yielded R^2 values below 0.5, contrasted with ANN's R^2 exceeding 0.88. This gap underscores the limitations of linear models in dynamic near-road environments. BMA aggregates predictions from multiple models based on their posterior probabilities, explicitly quantifying uncertainty. While BMA improved over MLR in capturing complex influences and offered uncertainty estimates, its predictive accuracy still trailed the ANN. BMA's strength lies in its capacity to handle limited or highly variable data and provide probabilistic insights, but it may not model nonlinearities as flexibly as ANNs.

Table 2. Alternatives Performance on Near-Road Air Pollutants

Model	**MLR**		**BMA**	
Air Pollutant	**R^2**	**RMSE**	**R^2**	**RMSE**
NO_2	0.832	15.836	0.833	14.812
PM1.0	0.320	8.410	0.443	0.364
PM2.5	0.321	11.366	0.464	0.356
PM10	0.321	11.951	0.473	0.352

The ANN's superior performance suggests that nonlinear, data-driven methods can more effectively learn intricate patterns in near-road pollutant dynamics. Nevertheless, BMA's uncertainty quantification remains valuable. Decision-makers may benefit from both approaches: using ANNs for accurate predictions and BMA to understand the range of plausible outcomes and the probability distributions of parameter estimates. While BMA and MLR provide clearer interpretability at a parameter level, ANNs require model-agnostic explainers such as LIME and SHAP to improve interpretability. The trade-off is between the ANN's predictive prowess and BMA's probabilistic rigor. The choice may depend on project goals: if actionable accuracy is paramount, ANNs are suitable; if understanding uncertainty and parameter-level inference is essential, BMA complements ANN findings.

5.3 Implications and Contributions

These findings have practical importance for urban air quality management. The strong influence of heavy-duty vehicles on pollutants like PM and NO_2 highlights the potential impact of stricter emissions standards or route restrictions for these vehicles. Similarly, improving traffic flow to reduce congestion could lower NO_2 levels, as smoother traffic movement reduces acceleration-related emissions. By identifying the key drivers of pollution (e.g., truck counts, humidity, temperature), policymakers can design targeted interventions. For example, restricting heavy-duty vehicle access during peak travel times or optimizing traffic signal timing to minimize congestion can yield tangible improvements in air quality. Complementing ANN predictions with BMA's uncertainty quantifications provides a more comprehensive understanding of model confidence, which can be invaluable in decision-making processes.

The ANN model provided robust predictive performance for near-road pollutant concentrations, surpassing traditional MLR and Bayesian methods in explaining variance and reducing error. LIME interpretations revealed the dominant influence of heavy-duty vehicles, meteorological conditions, and traffic flow patterns on pollutant levels. While BMA offers an advantage in terms of probabilistic inference and uncertainty quantification, the ANN's ability to handle complex, nonlinear relationships and deliver accurate predictions makes it a powerful tool for informing policy and management strategies aimed at improving urban air quality.

6. Conclusion

This study explored the use of Artificial Neural Networks (ANN) to predict near-road air pollutant concentrations based on traffic-related and environmental variables. By leveraging high-resolution data collected from a high-traffic site in downtown Columbia, SC, we built and evaluated ANN models for predicting NO_2, PM1.0, PM2.5, and PM10 concentrations. The ANN architecture was specifically designed to capture the nonlinear, dynamic relationships inherent in near-road environments, incorporating features such as multi-trailer counts, traffic speed, gaps between vehicles, and meteorological factors like temperature and humidity.

The ANN demonstrated superior performance compared to traditional methods such as Multiple Linear Regression (MLR) and Bayesian Model Averaging (BMA). With significantly higher R^2 values and lower RMSE across all pollutants, the ANN effectively modeled complex interactions that linear and probabilistic methods struggled to capture. Furthermore, interpretability tools like LIME provided actionable insights into the contribution of individual predictors, addressing concerns about the “black box” nature of ANNs and enhancing their utility for policymakers and urban planners.

While ANNs proved highly effective in this study, certain limitations remain. The ANN’s complexity, particularly with multiple hidden layers, can lead to overfitting, especially when the training data is limited. Although regularization techniques such as dropout and early stopping were employed, overfitting remains a potential concern in data-constrained scenarios. ANNs require substantial and high-quality data to perform optimally. Missing or inconsistent data can lead to degraded performance or increased uncertainty in predictions.

Addressing limitations such as overfitting and data dependency will require further advancements in model design and preprocessing techniques. Future work may focus on integrating ANNs with probabilistic approaches like BMA to combine the strengths of both methods to better support evidence-based decision-making in urban air quality management.

Supplementary Materials

The presentation video is available at https://youtu.be/d_vfp8AU2iY.

The following supplemental materials are available on the GitHub repository at https://github.com/csce585-mlsystems/Transportation_ANN/tree/main/Final_Submission.

- ANN Code
- README.md
- Input Data
- Presentation Slides

References

Anon. n.d. "Traffic Counts in South Carolina." Retrieved September 15, 2024 (https://scdottrafficdata.drakewell.com/publicmultinodemap.asp).

Askariyeh, Mohammad Hashem, Joe Zietsman, and Robin Autenrieth. 2020. "Traffic Contribution to PM2.5 Increment in the near-Road Environment." *Atmospheric Environment* 224:117113. doi: 10.1016/j.atmosenv.2019.117113.

Du, Wenjie, Lianliang Chen, Haoran Wang, Ziyang Shan, Zhengyang Zhou, Wenwei Li, and Yang Wang. 2023. "Deciphering Urban Traffic Impacts on Air Quality by Deep Learning and Emission Inventory." *Journal of Environmental Sciences* 124:745–57. doi: 10.1016/j.jes.2021.12.035.

Fontes, Tânia, Luís M. Silva, Sérgio R. Pereira, and Margarida C. Coelho. 2013. "Application of Artificial Neural Networks to Predict the Impact of Traffic Emissions on Human Health." Pp. 21–29 in *Progress in Artificial Intelligence*, edited by L. Correia, L. P. Reis, and J. Cascalho. Berlin, Heidelberg: Springer Berlin Heidelberg.

Hystad, Perry, Paul A. Demers, Kenneth C. Johnson, Richard M. Carpiano, and Michael Brauer. 2013. "Long-Term Residential Exposure to Air Pollution and Lung Cancer Risk:" *Epidemiology* 24(5):762–72. doi: 10.1097/EDE.0b013e3182949ae7.

Kadiyala, Akhil, Devinder Kaur, and Ashok Kumar. 2013. "Development of Hybrid Genetic-Algorithm-Based Neural Networks Using Regression Trees for Modeling Air Quality inside a Public Transportation Bus." *Journal of the Air & Waste Management Association* 63(2):205–18. doi: 10.1080/10962247.2012.741054.

Li, Can, N. Christina Hsu, and Si-Chee Tsay. 2011. "A Study on the Potential Applications of Satellite Data in Air Quality Monitoring and Forecasting." *Atmospheric Environment* 45(22):3663–75. doi: 10.1016/j.atmosenv.2011.04.032.

Lin, Yuan-Chien, Yu-Ting Lin, Cai-Rou Chen, and Chun-Yeh Lai. 2025. "Meteorological and Traffic Effects on Air Pollutants Using Bayesian Networks and Deep Learning." *Journal of Environmental Sciences* 152:54–70. doi: 10.1016/j.jes.2024.01.057.

Mahesh, Srinath, Gitakrishnan Ramadurai, and S. M. Shiva Nagendra. 2018. "Real-World Emissions of Gaseous Pollutants from Diesel Passenger Cars Using Portable Emission Measurement Systems." *Sustainable Cities and Society* 41:104–13. doi: 10.1016/j.scs.2018.05.025.

Matthaios, Vasileios N., Joy Lawrence, Marco A. G. Martins, Stephen T. Ferguson, Jack M. Wolfson, Roy M. Harrison, and Petros Koutrakis. 2022. "Quantifying Factors Affecting Contributions of Roadway Exhaust and Non-Exhaust Emissions to Ambient PM10–2.5 and PM2.5–0.2 Particles." *Science of The Total Environment* 835:155368. doi: 10.1016/j.scitotenv.2022.155368.

Rangel, Adan, Amit U. Raysoni, Mayra C. Chavez, Soyoung Jeon, Juan Aguilera, Leah D. Whigham, and Wen-Whai Li. 2022. "Assessment of Traffic-Related Air Pollution (TRAP) at Two near-Road Schools and Residence in El Paso, Texas, USA." *Atmospheric Pollution Research* 13(2):101304. doi: 10.1016/j.apr.2021.101304.

Rossi, Riccardo, Riccardo Ceccato, and Massimiliano Gastaldi. 2020. "Effect of Road Traffic on Air Pollution. Experimental Evidence from COVID-19 Lockdown." *Sustainability* 12(21):8984. doi: 10.3390/su12218984.

Song, Jun, and Ke Han. 2020. "Deep-MAPS: Machine Learning Based Mobile Air Pollution Sensing."