From Cameras to Concentrations: Estimating Black Carbon From Vehicles with Machine Learning and Computer Vision Techniques

by

Aryan Sadeghi

A thesis submitted in conformity with the requirements
for the degree of Master of Science

Institute of Health Policy, Management and Evaluation
University of Toronto

From Cameras to Concentrations: Estimating Black Carbon From Vehicles with Machine Learning and Computer Vision Techniques

Aryan Sadeghi
Master of Science
Institute of Health Policy, Management and Evaluation
University of Toronto
2025

Abstract

Urban traffic congestion contributes significantly to black carbon (BC) pollution, a harmful byproduct of fuel combustion with disproportionate impacts on residents living near busy roads. While understanding the impact of BC on our health is critical, accurate BC monitoring remains limited because the necessary instruments are costly and sparsely deployed. The inability to more frequently and equally get data on BC from vehicles, in turn, creates data deserts in under-monitored areas. In contrast, traffic surveillance cameras are widely deployed across cities worldwide and offer an abundant, high-resolution data source that captures vehicle types, density, and flow as factors known to influence BC levels. We hypothesize that features representing vehicle emission covariates derived from traffic video, when combined with environmental data, can serve as a proxy to estimate BC concentrations. We developed a computer vision-based system using YOLOv8 and XGBoost to extract traffic features and predict minute-level BC concentration, explaining up to 72% of variance in selected Toronto locations. We discuss the implications of vision-based sensing on privacy, feature limitations, and large-scale deployment. This thesis closes with plans to extend this study by considering other low-cost sensing modalities.

To Mom and Dad, whose devotion shaped who I am.

Acknowledgements

I thank my supervisor, Dr. Camellia Zakaria. She gave generous time for long meetings and offered careful and detailed feedback that shaped this thesis at every stage. She kept her door open and never rushed my questions. I am also thankful for the sweet chocolates she kept in her room. Many of the ideas in these pages began after a brief chocolate pause, which gave us a smile and a fresh start.

I am also grateful to Professor Marianne Hatzopoulou and her team in the Department of Civil & Mineral Engineering at the University of Toronto. They provided the equipment needed for data collection and took the time to teach us how the devices operate. Their practical guidance and steady support made our field work possible.

My friend, Amirhossein Rostami, deserves special thanks. Born to be a software engineer, he is capable of combining complex ideas and turning them into simple and strong solutions. Many parts of this system carry his mark. Without those long talks with him as we watched beautiful sunsets in Toronto, this work would not have reached its current quality.

And finally, this thesis reflects the unwavering support and love I received from my family far from here, on another continent across the ocean. I owe deep gratitude to all of them, especially my little sister, Parnian, whose steady support, even from far away, gave me a strong reason to push forward and not give up. She made hard days easier, and I am proud to share this work with her.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Black Carbon (BC) is a short-lived climate pollutant[1], produced during the incomplete combustion of fossil fuels, biomass, and biofuels. It has emerged as a particularly complex urban air pollutant; BC brings serious health and environmental consequences. Scientific evidence links exposure to BC and fine particulate matter ($PM_{2.5}$) with increased risks of asthma, cardiovascular disease, and premature death, with children, the elderly, and those with pre-existing conditions especially vulnerable [1]. Beyond immediate health threats, BC strongly absorbs sunlight and is widely regarded as *among the largest positive climate forcers after* CO_2 [2].

A substantial body of Canadian research supports the conclusion that "traffic is one of the dominant sources of urban black carbon". Specifically near busy roads and highways, BC concentrations can rise sharply, creating local "hotspots" that may reach five to ten times the levels observed at urban background sites [3, 4]. Prior environmental research in Toronto and Vancouver shows, near-road black carbon can be dominated by local traffic emissions, while the remaining share reflects regional background [3]. These hotspots matter not only because levels are high, but because they sit next to traffic corridors that frequently run through lower-income and racialized neighborhoods [4]. As a result, recent Canada-wide analysis finds higher traffic-related air pollution exposures in lower-income and racialized communities; the magnitude varies by city and pollutant (including BC where measured) [5]. This disparity is reflected in health indicators such as elevated asthma hospitalizations [1, 3]. As of 2025, Toronto's public network reports $PM_{2.5}$ at just several core stations (e.g., Downtown, East, North, West) with additional specialized/roadside sites; direct BC measurement is limited to a subset of roadside or study locations [6]. Meanwhile, low-cost sensor networks are disproportionately sited in higher-income, whiter neighborhoods, leaving gaps in underserved areas[7], and creating "data deserts"[2].

Cost is a key driver of these gaps. Gold-standard instruments like aethalometers can exceed $10,000 per unit[3], require expert maintenance, and thus remain limited to a few fixed sites [7]. Another example is mobile campaigns using Google Street View vehicles equipped with aethalometers. These campaigns have delivered unprecedented spatial detail [9]; Yet, these research campaigns are expensive. As a result, they rarely provide the timely and reliable data that, according to the World

[1] Short-lived climate pollutants (SLCPs), such as BC, have atmospheric lifetimes of days to weeks.

[2] "Data deserts" are neighborhoods with little or no air-quality monitoring, often those with higher proportions of marginalized populations.

[3] Aethalometers quantify BC via light absorption on a filter tape [8].

Health Organization (WHO), are essential for public health interventions [10]. Low-cost sensors (e.g., PurpleAir) improve coverage but still concentrate in wealthier districts [7]. Together, these exposures and data gaps place BC as a challenge in the domain of environmental justice: communities with the highest risk often have the least monitoring and the least agency to respond. Reflecting this, the WHO and United Nations Educational, Scientific, and Cultural Organization (UNESCO) have framed closing data and exposure gaps as an ethical imperative.[4] For cities intent on reducing inequities, better data is a necessary first step [7].

In the context of vehicular traffic, BC contributions vary by vehicle type. Heavy-duty diesel vehicles (HDVs), although often less than 5% of urban traffic, can emit BC at rates up to two orders of magnitude greater than gasoline-powered light-duty vehicles (LDVs) [11, 12]. Both the mix and proximity of vehicles, especially those stopping and accelerating near a receptor, are strong predictors of short-term BC spikes [3]. Traffic flow dynamics further modulate these patterns, with congestion and frequent acceleration events linked to higher emissions, and steady flow associated with lower emissions per vehicle [3]. Environmental factors compound this complexity. Wind speed and direction, temperature, humidity, and atmospheric stability all influence how BC is transported, diluted, or trapped near the ground [2]. Mobile monitoring in Canadian cities has shown that even minute-to-minute shifts in environmental features can cause BC concentrations to double or halve, regardless of steady traffic flows [3]. For example, low wind speeds (often below 2 m/s) create stagnant conditions in which BC accumulates and persists, resulting in acute local hotspots. Another example is humidity. Elevated humidity can alter the optical properties and deposition rates of BC particles, sometimes enhancing local concentrations or facilitating rapid removal [2]. Temperature and atmospheric mixing height also matter: cold, stable conditions (early mornings, winter) favor pollutant buildup, while warmer, more turbulent air in the afternoon enhances dispersion [2].

Given the growth in traffic and the scarcity of BC monitoring, there is a growing need for scalable and equitable approaches to exposure estimation. The rapid expansion of urban traffic-camera networks offers a promising proxy: in Toronto, hundreds of street and expressway cameras are accessible via the City's traffic camera program and Open Data portal, providing continuous information on vehicle activity across space and time [13]. We hypothesized that, with recent advances in computer vision, especially with object detection models like YOLO (You Only Look Once) [14], it is feasible to extract rich traffic features from video feeds automatically. Integrating these features with environmental data could provide a scalable, low-cost method for predicting BC concentrations at high temporal resolution, which is a critical need for responsive public health and environmental justice interventions. However, the potential for such an approach to accurately capture minute-level variation in BC remains underexplored.

The motivation for this thesis is grounded in addressing these gaps by combining advances in computer vision for traffic pattern recognition with machine learning to create a scalable tool for BC estimation.

[4] WHO Air Quality Guidelines and UNESCO's *Ethics of Artificial Intelligence* outline principles for fair, trustworthy, and equitable use of technology in public health.

1.1 Thesis Statement and Research Questions

This thesis is motivated by the need for a more accessible modality for estimating BC from vehicular sources and aims to develop a prediction technique for estimating BC using traffic video feeds, computer vision, environmental features, and machine learning. With more accessible estimates of BC, the technique can increase data availability in under-monitored locations and support closing data deserts. Our core hypothesis is that traffic video, combined with environmental data, can provide sufficient vehicle-emission covariate information to estimate minute-level BC concentrations. To address this, we pose three central research questions:

1. Can a machine learning model trained on video-derived traffic features and environmental data accurately predict minute-level BC concentrations?
2. What video-derived and environmental proxy features are most predictive of local BC concentrations at temporal scales?
3. When and why does the model underperform across environmental features, traffic situations, and sites?

1.1.1 Thesis Contributions

This thesis makes several key contributions to the field of vision-based environmental health research. We particularly focus on scalable, low-cost estimation of BC exposure from traffic sources.

1. **Integrated Data Management Platform:** We developed a custom web-based panel to manage data collection and processing workflows. The system combines traffic features with live environmental variables (wind speed, temperature, humidity) retrieved automatically from online Application Programming Interfaces (API), ensuring high temporal and spatial alignment between features and target measurements. This platform is developed to conveniently facilitate consolidating data from multiple modalities for the large-scale studies we plan for in this project and its extensions.
2. **In-situ Data Collection:** We conducted a field study in real-world urban settings, collecting synchronized minute-level BC concentrations alongside traffic video and environmental data across multiple locations in Toronto. Data were collected using a portable microAeth AE51, custom camera setups, and integrated microphones.
3. **System Evaluation:** We developed an integrated pipeline for extracting synchronized traffic and environmental features, applied signal processing techniques and machine learning algorithms to build a BC estimation model. The final XGBoost model explained up to 72% of test-set variance, with wind speed and vehicle features as key predictors. We discussed mitigation strategies drawn from WHO and UNESCO guidelines and proposed directions for equitable and privacy-conscious deployment.

Together, this thesis demonstrates the technical feasibility of producing video-based features and integrating them with standard environmental data for estimating BC. We close the thesis by discussing the key ethical challenges that must be addressed to ensure fair, trustworthy adoption in diverse urban contexts. All these resources are available publicly via our GitHub repo: https://github.com/sensAILabs/BlackPitch-dataCollection.

1.2 Related Work

1.2.1 Black Carbon Pollution and Health

Urban air pollution has been a major subject of scientific concern for over a century, beginning with the widespread impacts of coal combustion and "soot" in industrialized cities. The 1952 London Smog disaster, primarily caused by particulate-laden emissions, sparked international recognition of the health consequences of urban air pollution and led to landmark environmental policies in Europe and North America. As measurement technologies advanced in the late 20th century, scientists began distinguishing among particulate components, and the term "black carbon" emerged to describe the refractory, light-absorbing fraction of fine particulate matter ($PM_{2.5}$) produced by incomplete combustion of fossil fuels and biomass [2].

Scientific consensus on the significance of BC accelerated through the 1990s and early 2000s, as field and epidemiological studies linked BC to severe public health and climate effects [1, 2]. The inhalation of BC particles is associated with a wide range of adverse health outcomes, including both acute and chronic effects [1]. Short-term exposure to elevated BC levels is robustly linked to increases in asthma attacks, respiratory symptoms, emergency department visits, and hospital admissions for both respiratory and cardiovascular causes [1, 15]. For example, time-series analyses have found that daily BC increments are associated with increased risk of acute myocardial infarction and exacerbations of chronic obstructive pulmonary disease (COPD) [1]. Experimental studies also show that BC induces inflammation and oxidative stress, both in airway tissues and systemically, contributing to acute health events [1]. Long-term exposure to BC and related $PM_{2.5}$ has even more profound consequences. Major cohort studies in North America and Europe have demonstrated that chronic exposure to BC is associated with increased rates of mortality from cardiovascular disease, ischemic heart disease, stroke, and lung cancer, even at concentrations below current regulatory standards [1]. A pooled analysis of over 28 million adults in Europe found that each 1 $\mu g/m^3$ increase in long-term BC exposure increased the risk of all-cause mortality by up to 6%, with even stronger associations for cardiopulmonary mortality [1, 16].

Canadian analyses estimate that air pollution, with BC as a major component, accounts for over 15,300 deaths annually, with traffic corridors presenting particularly elevated risks [15, 17]. Globally, air pollution is linked to roughly 6.7 million premature deaths per year, with BC contributing disproportionately due to its toxicity and ability to penetrate deep into the lungs [1]. The underlying mechanisms of BC toxicity are well-studied. BC particles are especially hazardous because of their small aerodynamic diameter (typically <1 μm), which enables them to reach the alveolar regions of the lung, enter the bloodstream, and promote systemic inflammation and oxidative stress [1]. This increases the risk of atherosclerosis, arrhythmia, and thrombotic events. Children, elderly adults, and people with pre-existing cardiovascular or respiratory conditions are especially vulnerable [1, 18]. In addition to direct effects, BC acts as a carrier for other toxic substances, such as polycyclic aromatic hydrocarbons (PAHs), heavy metals, and secondary organic aerosols, further increasing its toxicity [1]. Epidemiological evidence also points to heightened asthma incidence and impaired lung development in children living near busy roads with elevated BC levels [1, 18].

In summary, BC has evolved from a largely invisible component of "soot" to a focal point of urban environmental health and climate policy. Its health effects, ranging from acute respiratory distress to increased long-term risk of heart disease, stroke, and death, underscore the urgent need

for high-resolution, equitable exposure assessment and targeted intervention in urban environments.

1.2.2 Monitoring and Prediction of Black Carbon

Efforts to quantify urban air pollution began in the late 19th and early 20th centuries, driven by the visible impacts of soot and smoke in rapidly industrializing cities. Early measurement focused on bulk "smoke" using rudimentary methods, such as the Ringelmann chart (a visual scale of darkness) and basic filter-based gravimetric analysis [2]. These methods offered only crude estimates of total particulate pollution and could not distinguish between the various sources or chemical constituents of airborne carbonaceous particles. The push for more precise measurements accelerated in the 1970s and 1980s, as the health and climate effects of fine particulates became apparent. Filter-based reflectometry methods, which quantified the "blackness" of filters after air sampling, enabled more specific estimation of BC and related species (e.g., "black smoke," "elemental carbon") [2]. However, these approaches suffered from low temporal resolution (typically daily), labor-intensive sample handling, and challenges in calibration and source attribution. Thermal-optical analysis, introduced in the 1980s, provided further advances by separating organic carbon (OC) from elemental carbon (EC, a proxy for BC) based on their volatility and optical properties [2]. This method remains the "gold standard" for EC/OC fractionation but requires complex instrumentation, expert operation, and is not suitable for high-frequency or spatially distributed monitoring [2].

A major leap occurred with the development of the aethalometer [8], an automated instrument capable of real-time, continuous measurement of BC based on light absorption at specific wavelengths. Successive generations, such as the AE16, AE33, and the portable AE51 "microaethalometer," enabled minute-level and even second-level BC monitoring in the field. Aethalometers are widely regarded as the reference standard for mobile and stationary real-time BC sensing in both research and regulatory applications [3]. However, their high cost (often exceeding $10,000 per unit), the need for frequent calibration, and susceptibility to filter loading and humidity effects have limited their deployment, especially in lower-resource settings [7]. Other advanced instruments include the Multi-Angle Absorption Photometer (MAAP), which corrects for scattering artifacts in optical BC measurement, and the Single Particle Soot Photometer (SP2), which directly measures refractory BC mass by incandescence [2]. While highly accurate, these devices are generally restricted to laboratory or specialized mobile campaigns due to their complexity and cost.

The limitations of high-cost, low-density sensor networks have spurred interest in alternative approaches for mapping urban BC at fine scales. Mobile monitoring campaigns, using instrumented vehicles or bicycles equipped with portable aethalometers, have revealed steep BC gradients over small spatial scales, including within individual school zones and neighborhoods [9]. Yet, such campaigns are resource-intensive, episodic, and rarely sustained over time, leaving persistent gaps in temporal coverage. Recent years have seen the rise of low-cost optical particle sensors (e.g., PurpleAir, Plantower), which have enabled more air quality networks to be established [19]. However, these sensors primarily detect $PM_{2.5}$ mass, not specifically BC, and their readings can be biased by humidity, particle composition, and environmental conditions [7]. As a result, while useful for general pollution mapping, they are poorly suited for accurate BC quantification or for source apportionment. Despite technological progress, BC monitoring networks remain sparse and unevenly distributed, particularly in the neighborhoods most impacted by traffic-related emissions and environmental injustice [7, 4].

As of 2025, Toronto's regulatory network includes a limited number of $PM_{2.5}$ stations, with direct BC measurement available only at a subset of sites [6]. Satellite-based proxies and land-use regression models have further expanded spatial coverage, but currently lack the resolution and specificity needed to monitor minute-to-minute BC exposure at the street or household level [9]. Altogether, each new generation of BC measurement technology has addressed specific gaps, but fundamental challenges remain in scaling monitoring to the urban block level, reducing costs, and ensuring equitable data access. This motivates the search for alternative, scalable proxies for BC, such as those explored in this thesis: leveraging traffic video and environmental data to estimate minute-level BC concentrations in real time.

Machine Learning for Air Pollution Prediction

The prediction of urban air pollution levels, including BC, has long been a central challenge in environmental science. Early efforts relied on classical statistical approaches such as multiple linear regression and time-series analysis [20], typically using easily measured covariates like traffic counts, environmental variables, land use, and proximity to emission sources, to explain spatial and temporal variation in $PM_{2.5}$ and BC concentrations. However, these linear models struggled to capture the nonlinear interactions and complex dependencies that characterize real-world air quality data, especially at high spatial and temporal resolution [20].

As data volumes and sensor networks expanded, machine learning (ML) methods emerged as a powerful alternative. Classical algorithms, including decision trees, random forests, gradient boosting machines, and support vector machines, have demonstrated strong predictive performance for pollution forecasting [21]. These models excel at handling high-dimensional data and automatically capturing nonlinear relationships between predictors and pollutant levels. Ensemble tree methods, for example, are commonly used to integrate heterogeneous features—such as traffic volume, weather, satellite-based aerosol indices, and land-use variables—for both spatial and temporal prediction of $PM_{2.5}$ and BC [21]. Among these, XGBoost and random forest models are widely favored for their robustness, computational efficiency, and interpretability via feature importance metrics.

In recent years, a rapidly growing body of research has demonstrated that fusing high-resolution, vision-based traffic features with environmental data substantially improves model accuracy for air pollution prediction. Studies integrating vehicle counts and classifications extracted from traffic camera feeds have explained a significant portion of the variance in urban particulate matter concentrations. For example, Fan et al. reported that combining traffic-camera vehicle counts with weather variables explained up to 67% of the hourly variance in $PM_{2.5}$ within dense urban corridors [22]. Other works confirm the scalability and cost-effectiveness of video-derived features for air pollution modeling, particularly where direct BC sensing is infeasible due to sensor cost or coverage constraints [23].

The evolution of machine learning approaches has also paralleled advances in deep learning. The past decade has seen the emergence of convolutional neural networks (CNNs) and recurrent neural networks (RNNs) for pollution modeling [24]. These models can learn complex spatial and temporal dependencies from large, heterogeneous datasets, enabling finer-grained exposure mapping and short-term forecasting. Hybrid models that combine spatiotemporal deep learning architectures with traffic video-derived features and environmental data are now being actively explored [22]. Despite their promise, deep learning models often require large labeled datasets and may suffer from

reduced transparency compared to tree-based approaches.

Feature engineering and data fusion remain critical for successful ML-based pollution prediction. Integrating real-time traffic activity (including video or audio-derived counts), environmental data, and low-cost sensor measurements can improve model accuracy and generalizability [22]. Model interpretability is also essential, with tools such as SHAP (Shapley Additive exPlanations) and permutation feature importance offering insight into which variables drive predictions [25].

Evaluation metrics commonly include the coefficient of determination (R^2), root mean square error (RMSE), and mean absolute error (MAE), typically assessed via cross-validation or independent test sets [20]. State-of-the-art models often achieve R^2 values above 0.6 for $PM_{2.5}$ or BC at hourly or daily levels in urban environments, with accuracy typically decreasing at higher temporal resolution due to stochastic influences and data noise [22].

Despite these advances, several technical challenges remain. Accurate synchronization of video features with air quality data is required for meaningful integration. The under-detection of certain vehicle types (notably HDVs in mixed fleets), occlusion, variable lighting, and camera perspective can affect data quality. The difficulty of inferring emission rates without direct knowledge of engine characteristics remains a further limitation [11, 12]. Finally, data sparsity, especially in marginalized or sensor-scarce neighborhoods, can limit the generalizability of trained models, while persistent biases in feature selection or sensor deployment risk reinforcing inequities in exposure assessment [7]. The need for interpretable, robust, and transferable machine learning systems thus remains central to advancing equitable urban air quality management.

1.2.3 Local Contributors to Vehicular Black Carbon Concentrations

Urban BC concentrations are shaped by a complex interplay of local and regional factors. Near busy roads, local traffic (especially stop-and-go HDVs in the closest lanes) can dominate short-term concentrations; however, a regional background component remains and can vary over time independent of local traffic activity. Two categories of contributors, traffic-related and environmental, jointly determine the intensity, timing, and spatial distribution of BC exposure at the local level [2, 3]. Understanding their roles is critical for both accurate exposure assessment and the design of mitigation strategies.

Traffic-Related Contributors

Vehicular emissions remain the dominant proximate source of BC in nearly all urban settings [3, 4]. Extensive literature now details the specific vehicle- and traffic-level factors that modulate BC concentrations at fine spatial and temporal scales. In particular, HDVs, while typically are less than 5% of the total fleet, emit BC at rates that can exceed gasoline-powered vehicles by one to two orders of magnitude [11, 12]. Empirical studies from Toronto, Montreal, and San Francisco consistently find that street segments or time periods with elevated HDV presence display disproportionately high BC readings, even after adjusting for overall traffic volume [3, 9]. For example, Hilker et al. [3] observed that short bursts of HDV activity in Toronto arterial corridors led to transient BC spikes up to five times above background.

The proximity of vehicles to monitoring sites also plays a decisive role in observed BC levels. Both fixed-site and mobile monitoring campaigns have established that BC concentrations can vary sharply,

sometimes by a factor of ten, over distances of less than 20 meters, especially in "street canyon" geometries where air recirculation is limited [9]. Traffic dynamics further modulate emission rates. Stop-and-go traffic, frequent acceleration and deceleration, and congestion are repeatedly linked to incomplete combustion and sharply elevated per-vehicle BC emissions [11, 3, 12]. By contrast, free-flowing or highway traffic generally produces lower emissions per kilometer travelled, a pattern replicated in real-world emission factor studies across North America and Europe [26]. Vehicles that are stopped or accelerating near the receptor contribute most strongly to local peaks, with emission plumes often detected seconds after the passage of a diesel truck or bus [11, 3]. Time-resolved analysis of high-frequency BC measurements in Toronto and San Francisco has shown that the "impact window" [5] explains a large fraction of acute exposure events [9, 3].

Recent advances in mobile monitoring and computer vision have enabled detailed quantification of these covariates. Apte et al. deployed instrumented vehicles to map block-by-block pollution gradients in Oakland, California, showing BC "hotspots" tightly correlated with truck routes, loading zones, and areas of chronic congestion [9]. Chambliss et al. leveraged video-based traffic counts to attribute up to 70% of near-road BC variation to differences in vehicle mix and activity [4]. These and similar studies have repeatedly found that the combination of HDV frequency, stop-and-go events, and near-sensor proximity is the most important predictor of acute BC exposure. Analytical techniques used in these studies include high-resolution aethalometry, mobile platform deployment (cars, bikes, buses), computer vision-based vehicle counting, and, increasingly, data fusion with citywide traffic sensor feeds [9, 3, 26]. The consensus from this body of work is that simply knowing the total vehicle count is insufficient. Detailed features such as HDV fraction, stop frequency, queue length, and lane distribution are all required for accurate BC exposure assessment.

In summary, the current scientific consensus, grounded in both Canadian and international literature, is that traffic composition (especially HDVs and buses), vehicle activity (stop-and-go dynamics), and proximity to the receptor are the dominant and best-quantified predictors of local BC concentrations in cities [3, 4, 9, 11, 12, 26]. These findings underpin the traffic feature engineering in the present study and motivate our exploration for vision-based proxies to estimate BC concentration as a scalable estimation technique.

Environmental Contributors

In urban environments, BC concentrations are not determined by emissions alone. A range of environmental factors critically influence how pollutants disperse, accumulate, or are trapped close to street level [2, 3]. Understanding these dynamics is essential, both for accurate BC exposure assessment and for designing fair, effective monitoring systems.

One of the most important environmental determinants is wind speed and direction. Studies in Canadian cities such as Toronto and Montreal have demonstrated that even modest shifts in wind conditions can double or halve BC concentrations at a fixed location [3, 27]. High wind speeds promote pollutant dispersion, while low wind speeds can lead to the rapid build-up of BC near busy roads. This is especially relevant for minute-level modeling, where fluctuations in environmental features have immediate impacts on measured air quality. Atmospheric stability, influenced by factors like temperature and the vertical "mixing height" of the boundary layer, is another key driver. Stable

[5] Impact window is defined as the period immediately following HDV/LDV passage within a critical radius of the sensor

conditions, often found during nighttime or cold mornings, suppress vertical mixing, trapping BC and other pollutants at ground level [2, 3]. By contrast, warmer and more turbulent air during afternoons disperses pollutants more widely, reducing acute local exposures. These relationships have been systematically mapped in Toronto and other Canadian cities, showing clear temporal patterns in BC variability across the day and seasons [3, 27].

Urban geometry introduces additional complexity. Street canyons of buildings closely lining both sides of a street create a semi-enclosed environment that restricts air movement. In these settings, pollutants emitted by vehicles are less likely to be dispersed by the wind and can persist at high concentrations near the sidewalk. Hilker et al., Chambliss et al., and Hatzopoulou's earlier simulation studies have shown that BC levels can be higher inside street canyons than in open areas, and that these hotspots frequently coincide with vulnerable neighborhoods [3, 4, 28]. Another important environmental variable is humidity, which can alter both the optical properties and deposition rates of BC particles. Under high-humidity conditions, the behavior of airborne particles changes, sometimes increasing local BC concentrations [2].

Research at the University of Toronto has consistently shown that integrating real-time environmental data (wind, temperature, humidity) with traffic features substantially improves the explanatory power of BC models [27, 3]. For instance, recent work in Toronto demonstrated that wind speed alone could explain over a third of the short-term variance in measured BC, sometimes exceeding the predictive power of total vehicle counts [3]. This finding underscores the technical necessity of synchronizing environmental and traffic data streams in modern exposure assessment systems. In summary, environmental contributors are not just background noise in BC monitoring, but essential, model-defining variables. Incorporating these factors, as emphasized in the Canadian and international literature, is fundamental for any robust, fair, and technically sound approach to urban BC estimation.

1.2.4 Environmental Justice, Ethics, and Fairness in Urban Sensing

BC pollution is not just a technical or scientific challenge; it is also an urgent issue of environmental justice [7, 3]. A growing body of evidence shows that BC "hotspots" are especially common in neighborhoods adjacent to major highways, busy arterials, and industrial zones [9, 4]. In Toronto, for example, areas along Highway 401, the airport corridor, and in parts of Scarborough routinely experience BC levels up to five or ten times the city background [3, 4]. In Montreal, census tracts near the Turcot Interchange and industrial districts show similar patterns [4]. These zones are often home to larger proportions of low-income, immigrant, or racialized residents, many of whom have limited mobility or choice in where they live [7]. Despite the elevated exposure in these neighborhoods, monitoring resources remain scarce.

As of 2025, Toronto operates fewer than 20 public air quality stations for $PM_{2.5}$, and only a handful directly measure BC [6]. Montreal's fixed-site network is similarly sparse. Recent analysis indicates that over 70% of community air quality sensors are located in affluent districts, leaving "data deserts" in under-resourced, high-risk areas [7]. As a result, residents facing the greatest health risks from BC pollution often lack timely, locally relevant air quality information [4].

National and international policy organizations now emphasize that ensuring equitable access to high-quality environmental data is both an ethical imperative and a practical necessity for public health action [10, 29, 30, 31]. Hatzopoulou et al. and team highlight that actionable urban sensing

requires not only technical innovation but also engagement with the social context, transparency, and community involvement [27, 4]. At the same time, as sensor networks and computer vision become more widespread, new privacy and data governance challenges arise, especially in residential or pedestrian-heavy areas [29]. In response to all of these challenges, mentioned above, we propose a low-cost, scalable approach to approximate BC exposure and to deliver timely information where monitoring is sparse.

1.2.5 Traffic Surveillance using Computer Vision and Machine Learning

Given that traffic activity is the dominant and most dynamic source of urban BC [3, 4], and considering that the high cost and sparse deployment of BC sensors, especially in the communities most affected by vehicle emissions, a scalable, and low-cost approach to sensing vehicle patterns is essential to closing the exposure data gap. Reliable measurement of traffic dynamics is therefore fundamental to understanding and modeling near-road BC exposure.

Historically, urban traffic measurement has relied on a variety of physical and electronic sensors, each with distinct strengths and limitations. Inductive loop detectors, pneumatic road tubes, and infrared beam counters have been mainstays of fixed-location traffic sensing, offering granular vehicle counts and speed data [32]. However, these systems require invasive installation, are costly to maintain, and often result in sparse citywide coverage. Non-intrusive radar and microwave sensors have expanded capabilities for speed and volume estimation, yet are limited by issues such as signal interference and the inability to differentiate between vehicle types. Manual traffic surveys remain in use but are labor-intensive, expensive, and unsuited for continuous, long-term deployment across large urban areas [32].

The past two decades have witnessed a major shift from traditional traffic measurement methods toward vision-based sensing powered by advances in computer vision and machine learning. Efforts to extract information from traffic camera feeds focused on simple frame differencing and background subtraction to count vehicles or detect movement. Early lane detection relied on classical computer vision algorithms such as the Hough Transform, Canny edge detection, and color segmentation to extract lane boundaries and infer vehicle trajectories [33]. However, these approaches were limited by variable lighting, occlusions, shadows, and the complexity of urban scenes. With the rise of deep learning, particularly CNNs, object detection and classification accuracy improved dramatically. The introduction of algorithms such as R-CNN, Fast R-CNN, and later YOLO revolutionized real-time multi-class detection in video streams [14]. Recent YOLO variants, including YOLOv8 used in this thesis, can accurately localize and classify multiple vehicle types, including cars, buses, heavy trucks, and motorcycles within a single frame, even under challenging conditions [14]. The deployment of these models on urban camera networks enables the extraction of per-frame vehicle counts, tracking, and categorization, which are critical features for BC exposure modeling. Temporal aggregation of detected features is equally important. Frame-by-frame detections are combined to yield minute-level traffic metrics (counts, stops, queue lengths), and synchronized with environmental and BC sensor data using advanced alignment techniques, such as the cosine similarity method detailed in this thesis. The resulting dataset allows for comprehensive analysis of how dynamic traffic activity, lane-level behavior, and environmental conditions jointly contribute to local air pollution. Finally, vision-based sensing enables the generation of new features such as queue length, stopping events, lane occupancy, and vehicle proximity to the sensor not previously accessible at scale.

These have been shown in recent literature to explain a significant share of minute-to-minute BC variability in urban corridors [22, 4]. The integration of these techniques forms the technical backbone of our approach, offering both scalability and high spatial-temporal resolution, and addressing key limitations of prior studies that relied solely on aggregate traffic counts or costly in-situ sensing. Given the widespread availability of urban traffic video feed and accessible application of ML models, proving the The feasibility of this approach may help with the broader goals of environmental justice and equitable urban monitoring, supporting the need for low-cost and city-wide BC exposure estimation.

1.2.6 Summary and Research Gaps

The literature reviewed above demonstrates substantial progress in understanding the sources, health impacts, and spatial patterns of BC pollution in urban environments. The health consequences of BC are now well established, with robust evidence connecting both short and long-term exposure to increased risks of asthma, heart attack, stroke, and premature mortality [1, 15]. BC monitoring remains sparse, costly, and unevenly distributed, which leaves data deserts and limits routine street-level evidence. Mobile campaigns provide detailed data collection, but are episodic; low-cost BC sensors expand coverage, but are not evenly placed. The prior work also clarifies two findings. First, traffic is a major source of urban BC, with the highest levels next to busy roads and during stop–and–go conditions. HDVs have a disproportionate impact, and short spikes are shaped by very-near activity (idling, queuing, acceleration). Second, environmental features matter at short time scales, wind speed and direction control ventilation, while local geometry can trap or disperse plumes. Together, these two findings point directly to the features we use in this thesis, lane-level counts and stop fractions (including HDV share), distance/proximity to the camera, and basic weather (wind, humidity, temperature) captured at minute resolution.

Chapter 2

Proposed Solution

This chapter presents the proposed solution and the evaluation of our model. We begin with a system overview (Figure 2.1) that defines the inputs used in our pipeline, the video-based feature extractor (object detection and lane logic on image frames), and black carbon (BC) estimation models. We then describe the instrument configuration, data-collection dashboard, and construction/synchronization of the BC ground truth. Finally, we report results, answering the research questions on what the system learns and where it struggles.

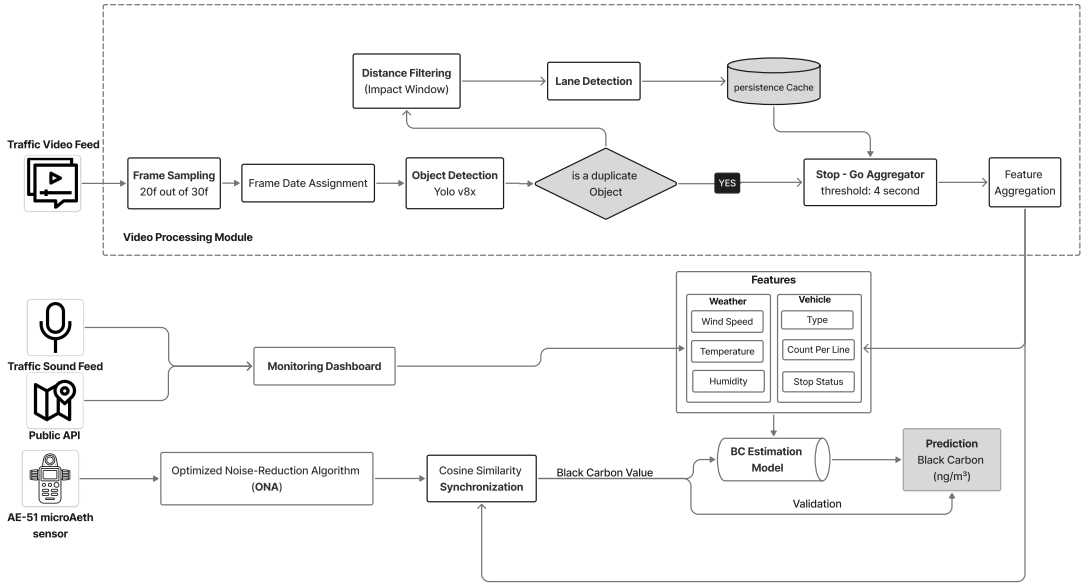


Figure 2.1: End-to-end system overview for BC estimation with computer vision and machine learning techniques.

Table 2.1: List of vehicular-based and environmental features for black carbon prediction.

Feature	Description	Feature	Description
$TotalVehicle_t$	Count of all vehicles at time t	LDV_{lane}	Count of LDV in lanes: lane 1 within 3 m, lane 2 between 3–6 m, lane 3 beyond 6 m
$StopLDV_{lane}$	Count of stopped LDV in each lane (vehicle stationary for at least 4 s)	HDV_{lane}	Count of HDV in the same lane distance bands
$StopHDV_{lane}$	Count of stopped HDV in each lane (vehicle stationary for at least 4 s)	*his_humid*	Historical humidity (2 min prior, %)
his_temp	Historical temperature (2 min prior, °C)	*his_wind*	Historical wind speed (10 m above ground, m/s)

2.1 System Overview

This section defines the components of the system (Figure 2.1) we use throughout the chapter. We begin with the input traffic video and basic environmental features. We then describe the feature extractor that operates on video frames: object detection with YOLO, lane geometry and masks, and the aggregation of lane-level counts, stop fractions, and distance-weighted activity for LDVs and HDVs[1]. Finally, we outline the BC estimation model, which contains the regression families evaluated in this work.

2.1.1 Input Data: Traffic Video and Environmental Data

Our model uses two inputs, traffic video and environmental data. For the traffic video, we collected our own traffic video using a fixed smartphone camera co-located with the microaethalometer (see Section 2.2.1). The same pipeline can be extended to use public traffic-camera images operated by the City of Toronto. These cameras publish still images that can be fetched at the provider's refresh rate, and the dashboard can be extended to ingest those images for feature extraction [13]. For environmental data, wind speed, temperature, and humidity were retrieved automatically from online Application Programming Interfaces (API) ([34] and logged into the dashboard at one-minute cadence. Units were standardized before modeling (wind in km/s at ∼10 m Above Ground Level (AGL), temperature in °C, relative humidity in %). By combining environmental data variables from the monitoring dashboard with lane-specific traffic metrics from the video pipeline (as summarized in Table 2.1), the module produces a unified set of explanatory variables for BC prediction.

2.1.2 Video Processing Module

The video processing module illustrated in Figure 2.1 ingests traffic video and outputs lane-specific traffic and vehicle features such as counts by class (LDV/HDV), activity distance by proximity, and stop or idle events[2]. The pipeline includes five steps: (i) frame sampling and timestamping, (ii) object detection and tracking with confidence filtering and duplicate removal, (iii) distance based impact window filtering, (iv) lane detection with vehicle to lane assignment, and (v) a stop and go

[1] Light-duty vehicles and Heavy-duty vehicles as described in chapter 1
[2] As described in chapter 1 prior studies identify these features as key covariates for near-road BC.

aggregator that flags vehicles stationary for at least 4 seconds. The module then writes lane-specific counts and metadata for downstream modeling and also produces annotated frames for visual quality checks.

Frame Sampling To reduce computational load while maintaining sufficient temporal resolution for vehicle detection, we implemented a frame sampling module. Rather than processing every frame of the input video, frames are selected at fixed intervals based on the target sampling rate, expressed in seconds. This ensures that each processed frame corresponds to a consistent real-world time gap, enabling efficient analysis without significantly compromising object detection accuracy. By discarding intermediate frames, the module optimizes inference time for YOLO-based detection and subsequent lane assignment while preserving key traffic dynamics necessary for BC prediction.

Frame Date Assignment Accurate temporal alignment between detected vehicles and BC measurements requires precise timestamping of each processed video frame. To achieve this, frame date assignment is performed using the original recording's metadata, which provides the absolute start time of the video. As each frame is processed in sequence, its timestamp is computed by incrementing the start time by the elapsed frame count divided by the video's frame rate, measured in frames per second (FPS). This ensures that every detection event can be mapped to a precise real-world time, allowing synchronized integration with high-frequency BC data and other environmental measurements.

Object Detection Vehicle detection in our pipeline is performed using the YOLOv8x model [14], selected for its high accuracy and real-time performance in video analysis tasks. The model was initialized with pretrained weights on the COCO dataset [35] and executed on a CUDA-enabled GPU to process each frame of the traffic video. We targeted only vehicle-related classes (e.g., `car`, `truck`, `bus`, `motorbike`) as detection objects, since these are the primary contributors to BC emissions in our study context. The model was configured to track objects persistently across frames, enabling consistent identification of individual vehicles via unique IDs.

Stop–Go Aggregator To distinguish between moving and stationary vehicles, we extended the YOLO-based tracking module with a temporal motion analysis routine. Each detected vehicle is persistently tracked across frames using a unique object ID. The module records the last seen time for each vehicle and monitors its positional changes between consecutive frames. A vehicle is classified as *stopped* if its centroid position remains effectively unchanged for a continuous period of at least four seconds. This threshold was selected to differentiate true stop or idle conditions (associated with elevated BC emissions) from transient deceleration events where vehicles slow but do not come to a complete stop.

Distance Filtering (Impact Window) To focus on vehicles influencing BC measurements, we implemented a distance filtering step based on bounding box size. The bounding box area serves as a proxy for the vehicle's distance from the camera, with larger areas indicating closer proximity. Only vehicles exceeding a predefined area threshold, corresponding to an estimated distance of approximately three meters from the monitoring setup, are retained for analysis. This "impact

window" filtering ensures that the detection pipeline prioritizes vehicles with the greatest potential to affect local BC concentrations, reducing noise from distant traffic with minimal direct influence.

Lane Detection and Vehicle-to-Lane Assignment Differentiating vehicles by traffic lane is essential in our study, as vehicles traveling in the lane closest to the monitoring station are expected to exert the most direct influence on BC concentrations. While YOLO provides accurate object detection, it does not natively classify detected vehicles by lane. To address this, we implemented a lane detection stage using the Hough Transform[36], a classical computer vision technique for detecting parametric shapes such as lines in images. The process begins by applying the Canny Edge Detection [37] algorithm to each video frame, extracting high-gradient regions corresponding to lane markings. The resulting binary edge map is then processed using the Probabilistic Hough Line Transform, which identifies straight line segments by mapping edge points (x, y) into a parameter space (ρ, θ), where ρ represents the perpendicular distance from the origin to the line and θ its orientation. Peaks in this multi-dimensional array correspond to prominent lines in the image. This approach is robust to noise and partial occlusions, making it suitable for outdoor traffic scenes where lane markings may be worn or partially obscured. Once lane boundaries are detected (Figure 2.2), we define region-based segmentation manually using detected lines for vehicle-to-lane assignment. Each YOLO-detected vehicle is represented by its bounding box, from which the centroid (x_c, y_c) is computed. The vehicle is then assigned to a lane by determining its position relative to the detected lane boundary lines. This enables independent counting of vehicles per lane, providing lane-specific traffic features for subsequent BC prediction modeling.

Figure 2.2: Result of applying Hough Transform with YOLO to determine the type of vehicle and distance from the microaethalometer.

2.1.3 Black Carbon Estimation Model

We developed multiple statistical machine learning models to predict BC concentrations from video-derived traffic and environmental features. The task was formulated as a supervised learning problem using synchronized BC measurements as ground truth. This list of models enables consistent benchmarking and selection of the most suitable regression model for BC prediction under real-world traffic conditions. The implemented regression algorithms are summarized in Table 2.2.

Table 2.2: Regression algorithms implemented for BC modeling.

Model	Description
Linear Regression (LR)	Baseline for interpretability and simplicity.
Support Vector Regression (SVM)	Kernel-based method for non-linear patterns.
Random Forest Regressor (RF)	Ensemble of decision trees with bootstrap aggregation.
Gradient Boosting Regressor (XBR)	Sequentially built trees minimizing residuals.
XGBoost Regressor (XGBoost)	Optimized gradient boosting with regularization.

2.2 Study Design

2.2.1 Field Setup and Sensor Deployment

To enable high-fidelity collection of BC, traffic, and environmental data in real-world urban settings, we established a custom sensing infrastructure at multiple Toronto sites. The field setup was designed for maximal data integrity, reproducibility, and technical alignment with recent work on vehicle detection studies [38]. This section details the physical setup, sensor modalities, dashboard, and data management tools, and technical innovations introduced in our fieldwork. Field sites were selected to capture a representative cross-section of Toronto street types and traffic conditions, focusing on locations with a mix of traffic flow and minimal occlusion from vegetation or parked vehicles. Selection of monitoring distance, orientation, and device height drew directly from protocols in the vehicle sensing literature, particularly the study by Billy Dawton et al. [38], which recommends a 3–6 meter offset from curb center for maximal audio and video signal quality. Figure 2.3 displays the spatial distribution of study sites across Toronto. Each site's proximity to high-traffic corridors and variety of traffic composition enabled robust generalization of BC–traffic relationships. At each site, a suite of sensors was deployed, with all devices co-located at a standardized height of 150 cm above the sidewalk, facing perpendicular to the nearest traffic lanes. Figure 2.4 illustrates a real-world example of sensor placement mimicking the setup by Dawton et al. [38].

- **Black Carbon Monitor:** We utilized the microAeth AE51 (AethLabs), capable of real-time measurement at 10-second and 30-second intervals. The device was calibrated pre- and post-data collection, and all flow and filter maintenance followed manufacturer best practices.
- **Camera:** Video was captured using an Apple iPhone SE 2022, configured for continuous recording in full HD (1920x1080) at 30 frames per second. The smartphone was mounted stably, with the lens axis at 160-170 cm above ground, offset approximately 3 meters from the nearest lane center (mirroring [38]).
- **Audio System:** For select experiments, we deployed a dual-microphone array (laser stereo microphones with 0 dB/20 dB filters) connected via a RELACART Mikit Dual Channel

Figure 2.3: Map of field sites in Toronto, annotated with sensor locations and percentage of usable data collected per site.

Figure 2.4: Setup of video, microphones, and microaethalometer device.

Compact Audio Interface. The microphones were positioned in accordance with Billy Dawton et al. protocol, at 3–6 meters from the traffic lanes, enabling potential future work in audio-based vehicle identification [38].

- **Environmental Data:** Environmental conditions were logged from real-time API feeds

(Open-Meteo [34]), with one-minute temporal granularity.

Panel and Data Management System

A central element of our workflow was a custom, web-based dashboard for data management, monitoring, and quality control (see Figure 2.5). Built using a Python backend (Flask [39]), Bootstrap and JS frontend [40], SQLlite database [41], and Docker for easy deployment [42], the dashboard enabled seamless control of data streams, manual annotation of field notes, and visual review of traffic. The dashboard also included an automated data ingestion module for weather, along with synchronization tools. Screenshots of the dashboard's different modules are shown in Figures 2.5a, 2.5b. More notes on dashboard implementation and run instructions are in Appendix A2.

2.2.2 Data Collection

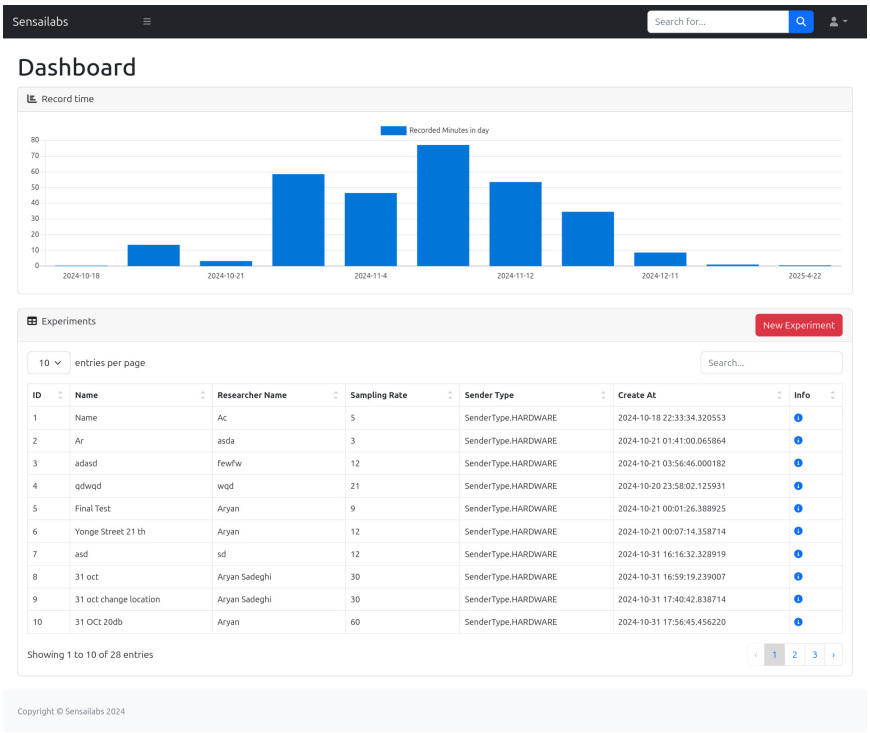
Our data collection spanned several months and involved multiple rounds of in situ field measurement in downtown Toronto. Each round had a distinct objective, like initial proof-of-concept, real-world stress testing of the full sensing pipeline, and final high-fidelity dataset acquisition for model development. Table 2.3 provides a detailed summary of all data collection rounds, including their temporal duration, used sensors, and whether the resulting data contributed to model training.

Table 2.3: Summary of data collection rounds: date, duration, sensors, technical notes, and use in analysis.

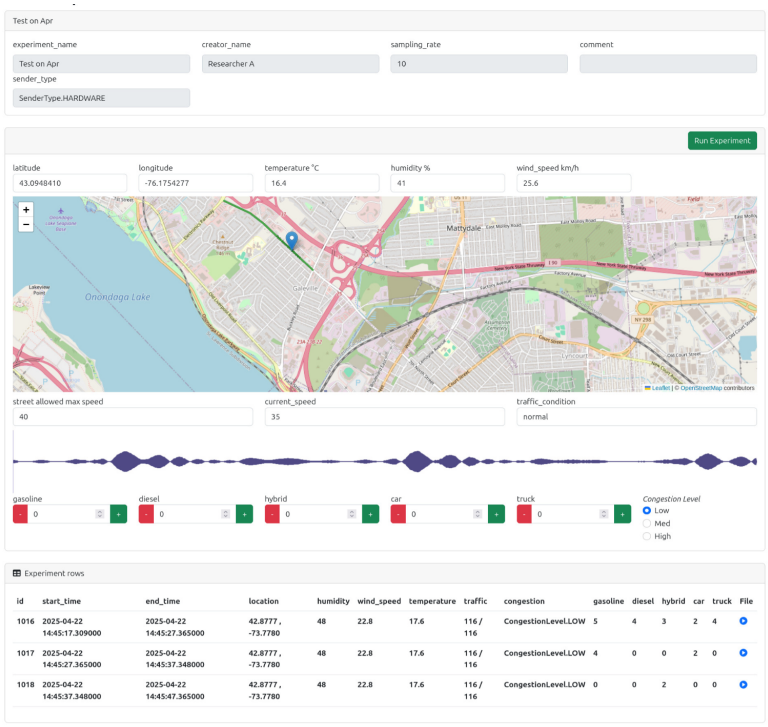
Date	Dur. (min)	Video	Sound	Tech. notes	Used
17 Oct 2024	35	0	1	Bus station deployment	1
31 Oct 2024	57	1	1	Street, system shakedown	0
4 Nov 2024	50	1	1	Street deployment	1
5 Nov 2024	85	1	1	Street deployment	1
12 Nov 2024	59	1	1	Street deployment	1
18 Nov 2024	33	1	1	Street deployment	1
19 Jan 2025	127	1	1	Device error, excessive noise	0

Timeline and data collection Rounds

The first field deployment occurred on **October 17, 2024**, collecting 35 minutes of data at a bus station in Toronto. This initial round served primarily as a technical shakedown, testing sensor connectivity, dashboard functionality, and ensuring time synchronization between modalities. While no final model training was based on this segment, insights gained here informed subsequent rounds and protocol refinements. On **October 31, 2024**, the system was deployed at a major Toronto arterial, yielding 57 minutes of continuous recording. This round marked the first real-world test of the full sensor suite and data management pipeline under typical street conditions. Only partial data from this round were included in the final model training due to incomplete ground-truth and suboptimal sensor alignment, but it provided critical stress testing and allowed for the calibration of both equipment and field procedure. Subsequent rounds (**November 4, 5, 12, 18, 2024**) focused on maximizing dataset quality and diversity. Sites were selected to include variations in traffic density, HDV prevalence, and environmental conditions. Over these four sessions, a total of approximately



(a) Dashboard main interface: log history visualization and data collection control.



(b) Record and synchronization interface: manual/automatic review of alignment, traffic context, and environmental data.

Figure 2.5: Dashboard system interfaces: (a) main status panel, (b) record/synchronization panel.

120 minutes of synchronized video, BC, audio, and weather data were acquired. Sensor placement, sampling rates, and dashboard logging procedures were fully standardized following lessons from earlier rounds. Only these rounds provided the curated, high-fidelity data ultimately used for model training and analysis. A final data collection round was conducted on **January 19, 2025**, lasting 127 minutes. However, technical issues, like, hardware malfunction in the BC monitor and operator error, introduced significant noise and data corruption. After quality control and anomaly detection, none of the data from this round was included in model development. This experience highlighted the importance of rigorous device pre-testing for future field deployments. All BC monitoring was performed at 10-second or 30-second intervals, aligned with the temporal granularity of environmental data (one-minute intervals) and subsampled video frames (1 frame per second). This design ensured dense, temporally aligned, multi-modal data streams amenable to downstream synchronization and modeling. Across all included field sessions, the final raw dataset comprised:

- 124 minutes of usable, synchronized recording, vehicle detection events.
- 378 BC readings (10/30 second intervals, post-QC).

Field Challenges and Mitigations

Despite careful planning, several operational challenges were encountered:

- **Occlusion:** Parked vehicles or transient buses sometimes obscured the field of view, necessitating post-hoc filtering of affected video frames.
- **Sensor Downtime:** A minor battery failure in the AE51 BC monitor on November 12 led to a short data gap, which was later interpolated for time-matching purposes but excluded from model training.
- **Ambient Noise:** Unusual construction noise and emergency vehicle sirens occasionally reduced audio quality; These periods were flagged in metadata for potential future audio analysis.
- **Device Error:** The January 19 malfunction (described above) reinforced the need for more aggressive device health monitoring.

Each challenge informed improvements in both hardware setup and field protocol. The use of a real-time dashboard with sensor health indicators, synchronized clocks, and frame-level annotation capacity proved essential for rapid troubleshooting and maximizing usable data.

In total, the multi-round data collection produced a dataset across several streets within close proximities of the Toronto region. Our setup was designed with the capacity to support more extensive data collection across additional regions. We acknowledge that, unlike many standard environmental studies, our dataset is limited in its ability to capture variations across seasons and times of day. Nonetheless, the data collected here provides a sufficiently robust foundation for model development. Only sessions with verified sensor alignment, full data completeness, and successful cross-modal synchronization were retained for the modeling pipeline.

2.2.3 Data Correction: Black Carbon Ground Truth

This part defines the BC ground truth used for training and evaluation. We start from the raw microaethalometer time series and apply ONA to suppress short-timescale noise while keeping

minute-level variation. The result is a cleaned BC target for modeling.
Even with Network Time Protocol (NTP) to reduce clock drift [43, 44], field deployments can still see small offsets after power cycles or operator errors. Therefore, we align time across devices. We run *Similarity Analysis & Synchronization* to estimate and correct any residual offset between the sensor and the camera timeline, check for minor drift, and lock both streams to a common minute index. Together, these steps yield a single, consistent BC target ready for the modeling pipelines introduced before in the chapter.

Optimized Noise-reduction Averaging (ONA)

Immediately after collecting the raw BC signal from the Aethalometer, we applied the Optimized Noise-reduction Averaging (ONA) algorithm provided by public tools [45] as a pre-processing step, before any similarity analysis or synchronization. The ONA method was developed to improve high time-resolution BC measurements by adaptively adjusting the averaging window based on the incremental light attenuation (ΔATN) within the instrument's filter tape. When ΔATN is small, indicating low particle loading and consequently low signal-to-noise ratio, ONA applies a longer averaging window to suppress high-frequency noise. Conversely, when ΔATN is large, the algorithm uses a shorter averaging window to preserve temporal resolution during periods of rapid concentration change. This adaptive averaging significantly reduces noise, virtually eliminates negative BC values, and retains meaningful short-term fluctuations in the signal [45]. In our workflow, the ONA-processed BC signal served as the cleaned input for the subsequent cosine similarity–based synchronization with traffic activity data. This ensured that the alignment procedure operated on a high-fidelity signal, minimizing the influence of instrument noise on the determination of the optimal time shift.

Similarity Analysis & Synchronization

Temporal misalignment between the BC ground truth signal and the traffic activity signal can obscure their true relationship. To address this, we evaluated and compared two frequency-domain synchronization approaches: (1) spectrum similarity, and (2) cosine similarity over complex-valued Fourier coefficients. Let $x[n]$ denote the traffic signal and $y[n]$ the BC signal, each with N samples. The Discrete Fourier Transform (DFT) of each signal is:

$$X[k] = \sum_{n=0}^{N-1} x[n]e^{-j2\pi kn/N}, \quad Y[k] = \sum_{n=0}^{N-1} y[n]e^{-j2\pi kn/N},$$

where k indexes discrete frequencies and j denotes the imaginary unit.

Spectrum Similarity. In the first approach, we computed the *power spectrum* of each signal:

$$P_X[k] = |X[k]|^2, \quad P_Y[k] = |Y[k]|^2,$$

and measured similarity as the Pearson correlation between $P_X[k]$ and $P_Y[k]$ across a range of candidate time shifts $\Delta t \in \mathcal{S}$. While this captures the overall frequency energy distribution, it discards phase information [46], limiting its sensitivity to pure temporal offsets.

Cosine Similarity between signals To preserve both magnitude and phase, we computed the cosine similarity directly on the complex DFT coefficients. For each candidate shift Δt, we decomposed $X[k]$ and $Y[k]$ into their real and imaginary parts:

$$X_{\cos}[k] = \Re(X[k]), \quad X_{\sin}[k] = \Im(X[k]), \quad Y_{\cos}[k] = \Re(Y[k]), \quad Y_{\sin}[k] = \Im(Y[k]),$$

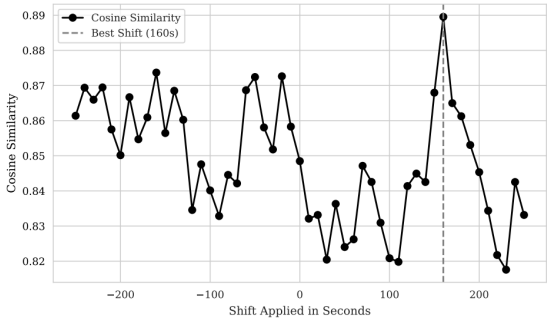
and defined the cosine similarity as:

$$\text{Cosine}(X, Y) = \frac{\sum_{k=0}^{N-1} \left(X_{\cos}[k]Y_{\cos}[k] + X_{\sin}[k]Y_{\sin}[k]\right)}{\sqrt{\sum_{k=0}^{N-1} \left(X_{\cos}[k]^2 + X_{\sin}[k]^2\right) \cdot \sum_{k=0}^{N-1} \left(Y_{\cos}[k]^2 + Y_{\sin}[k]^2\right)}}.$$

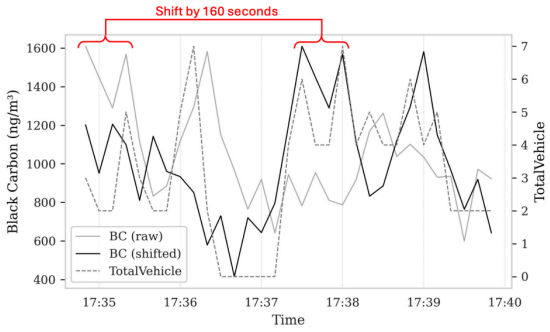
The optimal shift was selected as:

$$\Delta t_{\text{optimal}} = \arg\max_{\Delta t \in \mathcal{S}} \text{Cosine}\left(X[k],\, \mathcal{F}\{y[n - \Delta t]\}\right),$$

where $\mathcal{F}\{\cdot\}$ denotes the DFT. Signals were trimmed or zero-padded to equal lengths before the DFT to avoid bias from length mismatches. Cosine similarity consistently outperformed spectrum similarity in identifying the correct temporal offset, as expected due to its preservation of phase information [47].



(a) Cosine similarity between Fourier transformed BC vector and Fourier transformed *TotalVehicle* vector. The corrected BC signal is determined by the shift that maximizes cosine similarity, in this case, 160 seconds.



(b) Original BC signal (gray) plotted alongside vehicle counts (dotted line) reveals a temporal misalignment. The corrected BC signal (black), shifted by 160 seconds, shows alignment with vehicle activity.

Figure 2.6: Cosine similarity analysis and BC signal correction using optimal time shift.

Importantly, the $\Delta t_{\text{optimal}}$ determined via cosine similarity also corresponded to a local maximum in time-domain Pearson correlation between BC and traffic signals (as denoted in Figure 2.7). This is both expected and contextually consistent; in our study setting, higher traffic volumes generally correspond to elevated BC concentrations. The fact that spectral alignment maximized local correlation in the time domain reinforces the validity of the synchronization and its alignment with the underlying causal relationship between vehicular activity and pollution levels. Conclusively, the integration of robust synchronization, validated feature extraction, and dashboard-driven data integrity checks directly addresses key limitations of prior BC exposure studies and lays the foundation for the subsequent modeling and analysis presented in this thesis.

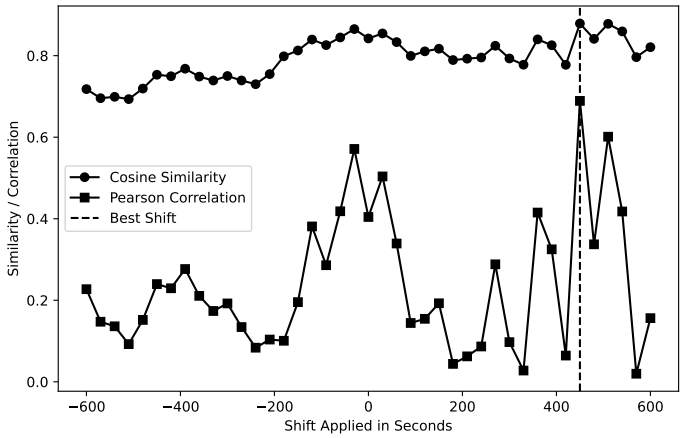


Figure 2.7: Cosine similarity and Pearson correlation versus applied shift. The optimal shift from cosine similarity aligns with a local Pearson correlation peak, consistent with the expected traffic–pollution relationship.

2.2.4 Final Dataset

The processes above led us to produce a final dataset for model development. The dataset includes a feature set (Table 2.1) and two BC targets, the raw microaethalometer signal and the ONA-processed signal. Both BC targets are aligned to the feature timestamps using the synchronization procedure described above, so every row represents the same one-minute interval across sources. We keep only minutes with complete covariates and valid BC readings (before and after ONA); trimming variants, when used, are applied later during modeling. Table 2.4 summarizes the resulting dataset collected under in-the-wild conditions (sites, minutes, and splits).

Table 2.4: Summary of data collected.

	Location	**Duration**	**LDV**	**HDV**	**Wind**	**Humidity**	**BC (in ng/m^3)**
Set$_A$	*[redacted]*	49.0 mins	97.4%, n=736	2.6%, n=20	μ=15.59, σ=0.10, min=15.5, max=15.7	μ=68.86, σ=0.99, min=68.0, max=70.0	μ=828.03, σ=219.51, min=427, max=1508
Set$_B$	*[redacted]*	32.5 mins	95.1%, n=442	4.9%, n=23	μ=27.03, σ=1.57, min=24.5, max=28.0	μ=66.28, σ=0.45, min=66.0, max=67.0	μ=729.82, σ=260, min=-242, max=1623
Set$_C$	*[redacted]*	32.5 mins	99.5%, n=551	0.5%, n=3	μ=5.70, σ=0.69, min=4.9, max=6.3	μ=67.29, σ=1.49, min=66.0, max=69.0	μ=391, σ=191, min=93, max=973
Set$_D$	*[redacted]*	10.0 mins	100%, n=90	0%, n=0	μ=9.20, σ=0.00, min=9.2, max=9.2	μ=78.00, σ=0.00, min=78.0, max=78.0	μ=502, σ=269, min=167, max=1391

To address potential multicollinearity among predictors, we computed the Pearson correlation matrix for the raw feature set (Figure 2.8a). Pairs of features exhibiting an absolute correlation coefficient greater than 0.70 were considered highly collinear and were excluded from further modeling. This threshold ensures that redundant variables, which may inflate variance and destabilize model coefficients, are removed while retaining distinct and informative predictors. For instance, `his_temp` and `his_wind` displayed a correlation of 0.96, leading to the removal of one variable from the model input set. The remaining features, after filtering for multicollinearity, were subsequently passed on to the regression model (Section 2.1.3) for training and evaluation.

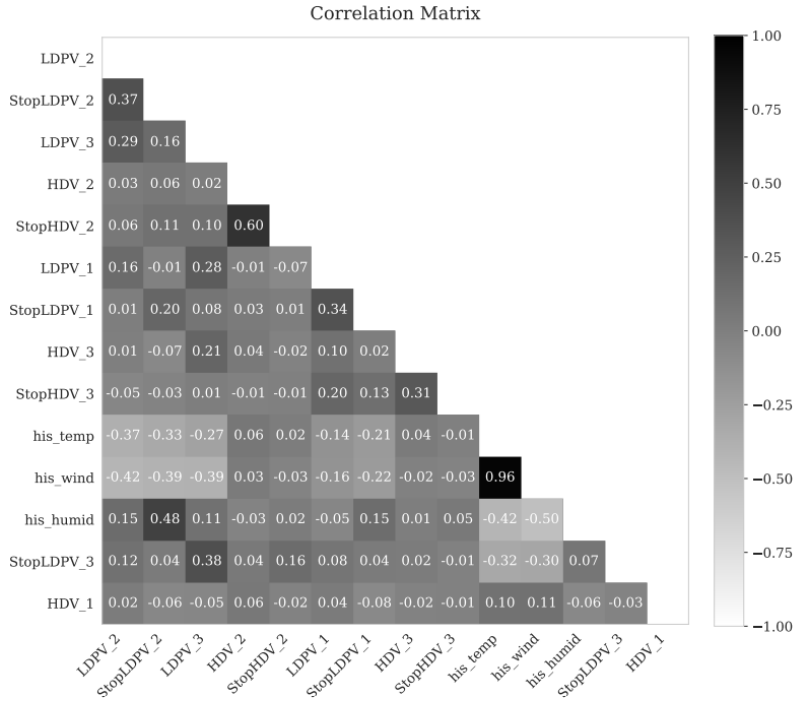


Figure 2.8: Correlation matrix. We excluded highly correlated features above 70%.

(a)

2.3 Results and Analysis

We present the evaluation setup and the way we handle noise in the BC target, and then report predictive performance (RQ1), model interpretation (RQ2), worst-case sensitivity analysis, and why and when model underperforms (RQ3).

2.3.1 System Evaluation

We use a fixed train/test split. All model selection and hyperparameter tuning occur strictly on the training portion. The held-out test portion is evaluated *once*. The same split is used for every pipeline (algorithm × noise setting) so that differences reflect methods rather than data partitioning. For descriptive plots and subgroup summaries, we use a 25% test, 75% training *stratified sampling*[48].

Metrics We report the metrics in Table 2.5, all on the test portion. These metrics are widely used in sensor/model evaluations because they summarize variance explained and penalize large deviations (important during high-concentration excursions) [49]:

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2, \qquad \text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}, \qquad R^2 = 1 - \frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}. \quad (2.1)$$

.

Table 2.5: Evaluation metrics for model performance assessment.

Metric	Description
R^2	Proportion of BC variance explained (closer to 1 is better).
MSE	Average squared deviation between predictions and ground truth.
RMSE	Square root of MSE, interpretable in ng/m^3.

BC target preparation: Minute-level aethalometer series are noisy at short time bases. We therefore evaluate eight post-processing options for the BC target, combining ONA (Subsection 2.2.3) and different Confidence Interval (CI) timing techniques: *None*, *ONA*, *CI-trim (local)*, *CI-trim (global)*, *ONA+CI (local)*, *ONA+CI (global)*, *None+CI (local)*, *None+CI (global)*. Trimming removes outliers according to confidence bounds (applied to the target only; features are unchanged). To show how each noise option alters the raw BC series, Figure 2.9 summarizes the percentage of data points modified by each method.

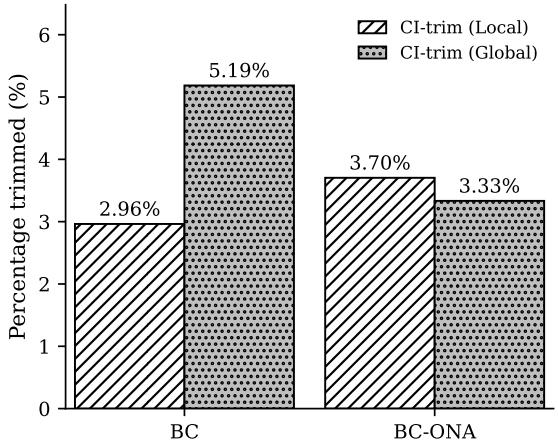


Figure 2.9: Percentage of minutes trimmed under local vs. global CIs for raw BC and BC–ONA.

Local vs. global (general) CI trimming. We use "CI-trim (local)" to mean that confidence limits are computed *within each dataset/site* (i.e., distributional bounds are estimated separately per dataset before concatenation), so trimming respects local variability. We use "CI-trim (global)" to mean that confidence limits are computed *once on the pooled data* and then applied uniformly across datasets. Local trimming is therefore adaptive to site-specific distributions; global trimming imposes a single set of bounds.

Pipelines, cross-validation, and model selection. Each *algorithm* combined with each noise reduction strategy defines one pipeline. Within the training portion only, we conduct k-fold cross-validation to select hyperparameters by minimizing mean CV RMSE (tie-breaker mean CV R^2). This avoids using the test set during tuning and mitigates selection bias [50]. Across different models and noise reduction settings, we selected the algorithm model that achieves the best performance across all measures.

For full reproducibility, the exact hyperparameter grids and selected values are reported in Appendix A1.4.

2.3.2 RQ1: Prediction Performance

We now quantify predictive accuracy on the held-out test portion. Following the selection rule in Section 2.3.1, we focus on the best noise reduction strategy (*ONA + CI-trim (local)*), then compare algorithms trained under that strategy. Table 2.6 summarizes performance across the five algorithms under the chosen noise reduction strategy. Nonlinear tree ensembles consistently generalize better than linear or kernel baselines. XGBoost yields the lowest out-of-sample error (Test RMSE = 129.42 ng/m^3, $R^2 = 0.72$), followed closely by Random Forest and Gradient Boosting; SVR and the linear baseline underperform. This ranking is expected when the outcome reflects interactions among traffic composition (e.g., HDV presence and queuing), source–receptor proximity (lane-resolved distance), and wind, all of which introduce non-additive effects that tree ensembles capture naturally [51]. Importantly, the train–test gaps for the top three models are modest, indicating that regularization and early stopping (for boosting) controlled overfit.

Table 2.6: Model evaluation using different ML algorithms with XGBoost regression yielding best performance (test split, noise reduction strategy = ONA + CI-trim local).

	Train RMSE	**Train MSE**	**Test RMSE**	**Test MSE**	R^2	**p-value**
Linear Regression (baseline)	184.70	34115.48	211.66	44799.06	0.25	-
Support Vector Regression	216.20	46743.69	238.49	56876.81	0.05	$p > .01$
Random Forest Regressor	82.71	6840.17	131.86	17387.02	0.71	$p < .001$
Gradient Boosting	83.61	6990.81	139.46	19447.89	0.68	$p < .001$
XGBoost	**95.68**	**9155.43**	**129.42**	**16748.60**	**0.72**	$p < .001$

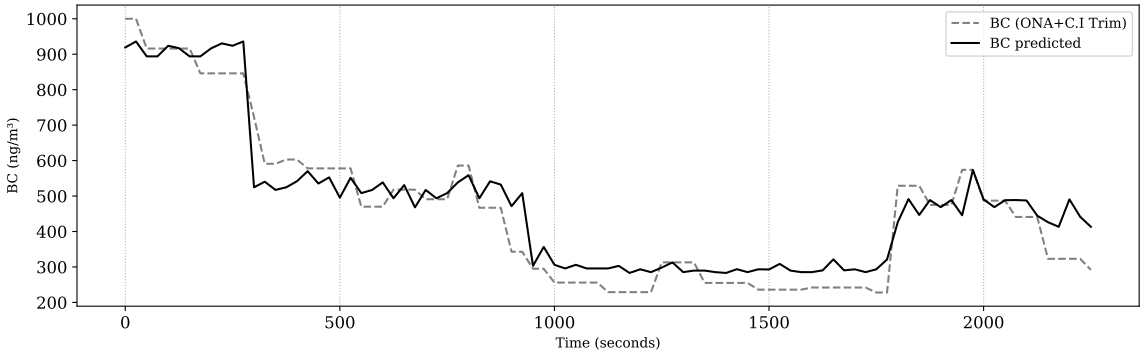
Table 2.7 shows systematic error reductions when (i) ONA is applied and (ii) trimming is performed with *local* confidence bounds. The chosen cell (ONA+CI local + XGBoost) is outlined. Same information for all of the other models is in Appendix A1.5 Together, these results indicate that both the *noise reduction strategy* and the *model class* contribute meaningfully to deployment-relevant accuracy.

Table 2.7: XGBoost performance from noise reduction and trimming.

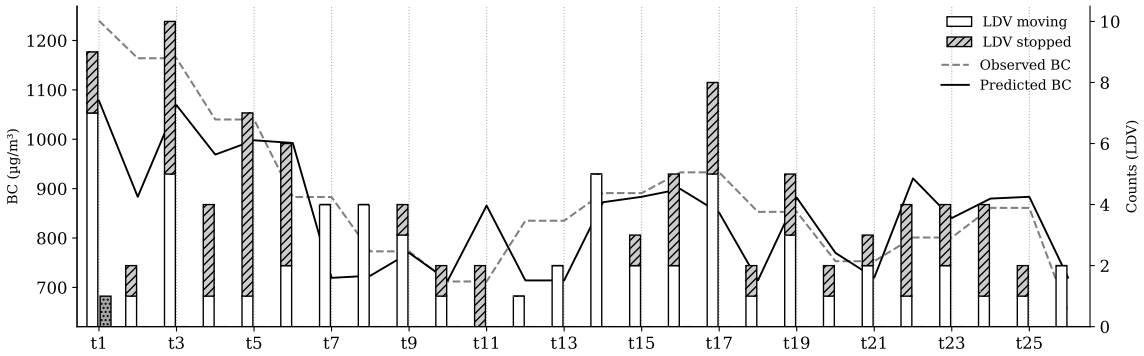
Strategy	**Train RMSE**	**Train MSE**	**Test RMSE**	**Test MSE**	R^2
None (baseline)	158.93	25259.00	212.47	45143.92	0.49
ONA	121.83	14841.89	140.10	19627.79	0.67
CI Trim — global	137.27	18842.19	210.78	44428.09	0.50
CI Trim — local	111.02	12324.98	207.22	42939.41	0.52
ONA + CI Trim — global	**75.25**	**5662.92**	135.57	18378.99	0.69
ONA + CI Trim — local	95.68	9155.43	**129.42**	**16748.60**	**0.72**

Figure 2.10 pairs a long-interval series and a selected short-interval (13 minutes) to show how predictions move with traffic activity. In the short interval, predicted BC rises during short bursts when stopped and moving counts increase near the setup, then relaxes as flow clears. Figure 2.10 shows a long window of predicted vs. observed BC and a 13-minute zoom with counts of moving and stopped vehicles by type. Over the long window, the model follows the main swings in the measured

signal. In the short interval, short peaks in BC line up with stop-and-go near the camera and with HDVs in the closest lane.



(a)



(b)

Figure 2.10: Predicted vs. observed BC over a long interval (2.10a) and a short interval (13-minute) (2.10b): BC peaks align with stop-and-go right next to the sensor and with HDVs in the closest lane.

2.3.3 RQ2: Feature Importance and Interpretation

With the test split fixed and model selection complete (Section 2.3.2), we now ask what drives the predictions and whether those drivers match near-road evidence. We use SHAP values on the final tree ensemble [25]. The global summary (Figure 2.11) points to two main groups of features. First, higher recent wind speeds push predictions down. This is consistent with near-highway studies showing weaker pollutant gradients when wind and boundary-layer mixing are stronger [52]. Second, very-near traffic, distance-weighted counts in the closest lane, and lane-level stop fractions push predictions up. Prior Canadian work reports that heavy-duty diesel vehicles and stop-and-go flow raise minute-level BC and, especially right next to traffic [27, 53]. These patterns are also shaped by geometry. Street canyons and wind speed can hold exhaust near the road and increase near-field build-up, which matches urban dispersion studies [54, 55]. Taken together, all of the internal logic of the model matches the determinants repeatedly highlighted in previous near-road studies [27, 53, 56].

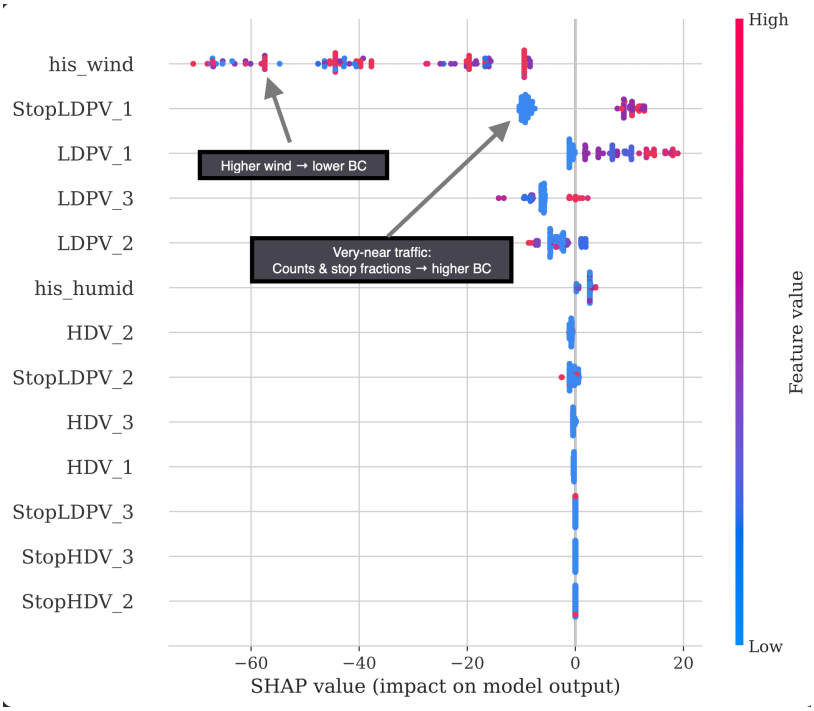


Figure 2.11: Global SHAP summary (beeswarm). Wind speed has a negative effect (dilution), while very-near traffic activity (distance-weighted counts; stop fractions) raises predicted BC, in line with near-road field studies and dispersion theory [27, 53, 52].

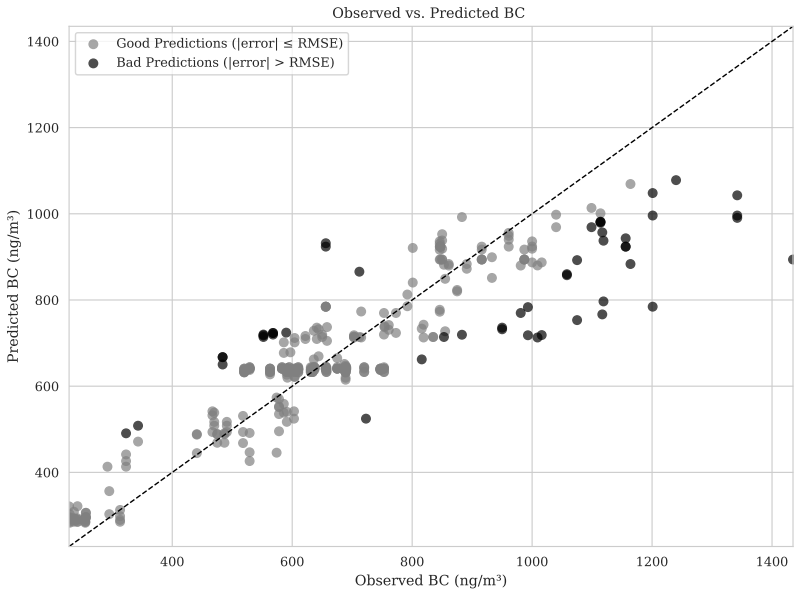


Figure 2.12: Parity with worst-case overlay: predicted vs. observed BC on the test split, with minutes above the RMSE threshold highlighted in black. High-BC regions contain a disproportionate share of worst-case points.

2.3.4 RQ3: Sensitivity Analysis

We next look at where the model misses and why. We mark data points with absolute error larger than the test RMSE as "worst-case" points. As shown in Figure 2.12, these worst cases sit mostly at high BC. This is where short, local changes are common. Grouped summaries make the pattern clearer. Using the same stratified sampling as in Section 2.3.1, minutes with low wind speed and high humidity account for a larger share of bad predictions than other situations (Figure 2.13). Prior field and dispersion studies also report that low wind speed reduces dilution, and moisture conditions can change how plumes build and clear [54, 55, 52].

Figure 2.13c shows the fraction of bad predictions at each site. To understand why certain locations accumulate more "worst-case" minutes, we inspected minutes with $|e| >$ RMSE and cross-checked raw logs and sample video (see also Figure 2.13). Two key findings stand out. First, HDVs are underrepresented in our corpus. Only $n = 46$ HDVs appear in the full dataset (Table 2.4). Consistent with that sparsity, HDV-related features are ranked lower by SHAP in our global summary even though prior work shows HDVs emit much more BC than light-duty passenger vehicles [12, 57]. In other words, the model has a limited opportunity to learn HDV effects at one-minute resolution. Second, one lane includes a streetcar that, while electric, temporarily reorganizes nearby traffic during boarding/alighting. Figure 2.14 compares the predicted BC value with the observed value in the same lane. Events like passing streetcars interrupt the flow and, more importantly for a single ground-level camera, cause short occlusions that undercount stopped vehicles. After excluding minutes affected by these occlusions, test RMSE improved to 114. Generally speaking, object occlusion is a known issue in ground-level video sensing, but it does not undermine our approach, especially when cameras are placed higher or from multiple angles to reduce blocking [58].

BC has both local and regional sources [3], and our modeling explicitly targets near-road, local contributions captured by lane-level traffic features. As a result, regional background can still drive some of the worst-case data points. Wind speed and direction can intermittently transport BC and other combustion aerosols, yielding high observed values during periods when video-derived features indicate low local emissions. These dynamics likely help explain residual variance and several of the "missed peaks" highlighted in our sensitivity analysis. Consistent with our literature framing, near-road BC is often dominated by local traffic, while the remaining share reflects regional background [3].

Key Takeaway

Our model combines traffic-video features (distance-weighted counts, stop fractions) with environmental features (wind, humidity) and, with XGBoost, reached $R^2 = 0.72$ and RMSE $= 129.42\,\mathrm{ng/m^3}$ on the test split. SHAP aligns with this pattern from prior studies. Wind speed and near-lane counts are the most important contributors in predicting BC values. Errors were larger when the wind was low and humidity was high, consistent with the case-based review, and regional background contributions also influenced several worst-case data points. Accordingly, we interpret the model's accuracy as strongest for local, traffic-driven variation at the minute scale, with higher uncertainty when background loads or long-range transport contribute a larger fraction of the signal.

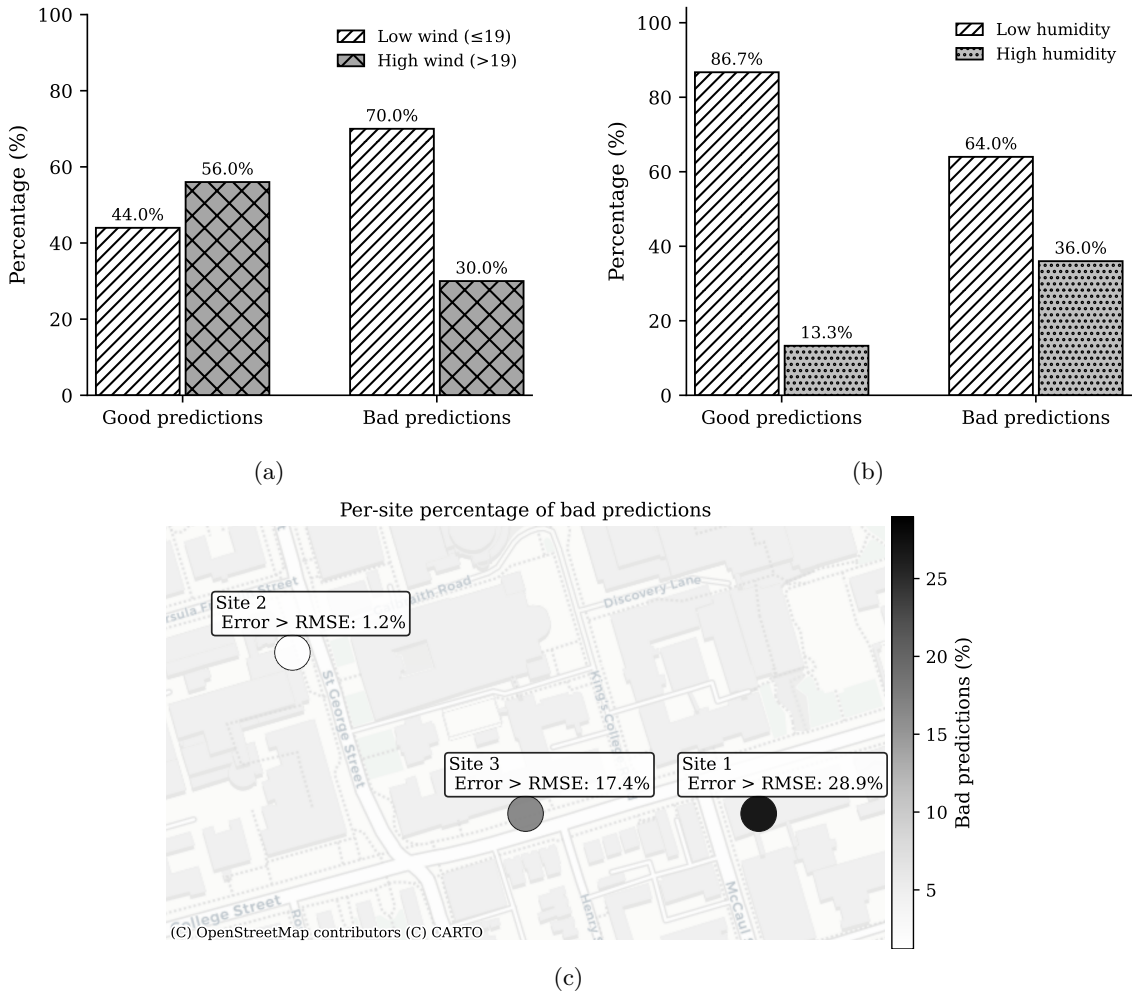


Figure 2.13: Prediction error by characteristics: (a) wind-speed groups; (b) humidity groups. Error shares are higher under low wind speed and high humidity. (c) Per-site error percentage: fraction of minutes with absolute error > RMSE at each site

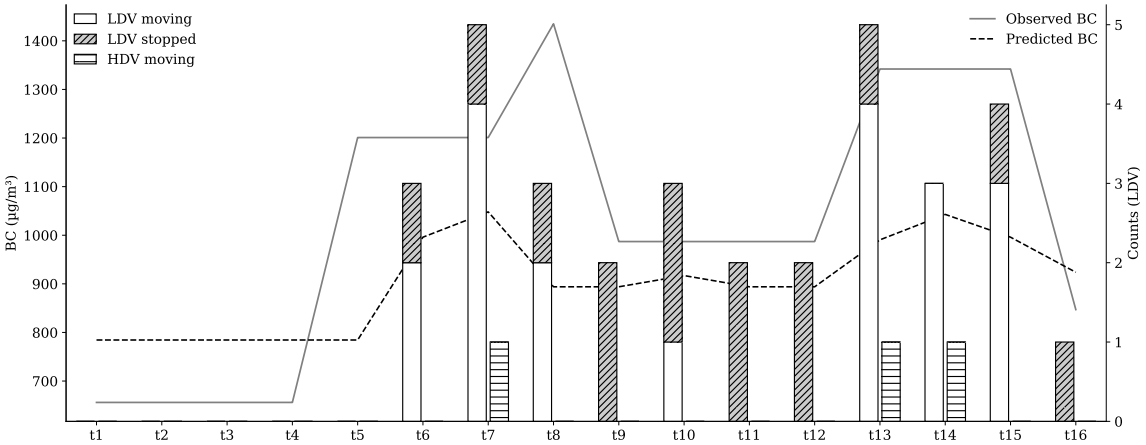


Figure 2.14: Predicted and observed BC concentrations, combined with counts of moving and stopped LDVs and HDVs over 3 minutes, charted at 10-second intervals. Predicted values do not follow the general trend of observed BC emissions, but align with vehicle presence.

Chapter 3

Discussion

This chapter reflects on what our system can and cannot do. We discuss the implications of our work concerning practical considerations in large-scale deployment, whereby not every place will have traffic surveillance infrastructure, and therefore requires us to consider an alternative sensing modality. We demonstrate the feasibility of extending our work with other sensors and close with notes on environmental justice and next steps.

3.1 Modeling and Practical Deployment Challenges

This section summarizes the main limits we face in practice and in modeling. We first show that BC has time dependence, which a model with only same-minute features may miss, and then we bring more discussions on deployment constraints and data limits.

First, autocorrelation function analysis (ACF) [59] in BC target shows time dependency and autocorrelation (Figure 3.1), which means current BC values depend partly on recent past values. A model that only uses contemporaneous traffic and weather features may miss this short-lag carryover unless explicit lag terms or a time-series component are included. Exploring hybrid ensembles and time series models can improve robustness by capturing short-lag persistence while retaining nonlinear covariate effects.

Second, using video for BC estimates has clear strengths. It is interpretable (we can see the traffic) and well matched to the features that matter (very-near activity and stop-and-go). But it has practical limitations when it comes to larger deployments. First, cameras cost money to install. They need power and network access, and they need stable mounting angles to keep the lane view consistent. For small pilots, this is manageable, but for a city-scale, it can be a barrier. Second, video raises privacy risk if frames are stored or transmitted. While on-device blurring/redaction can be added as a temporary solution, we are looking for a privacy-by-design solution.

Finally, as discussed earlier in error analysis (see Section 2.3.4), the biggest errors appear when the video does not capture the short events that drive BC spikes. For example, when a streetcar stops in-lane, vehicles behind it may be blocked from view just as they idle and then accelerate, leading to missed counts and weaker features. A related limit is that video alone cannot reliably tell engine or fuel type (e.g., diesel vs. gasoline or electric), even when it can label body class (car, bus, truck). This matters for BC because emission rates differ by engine type, heavy-duty diesel

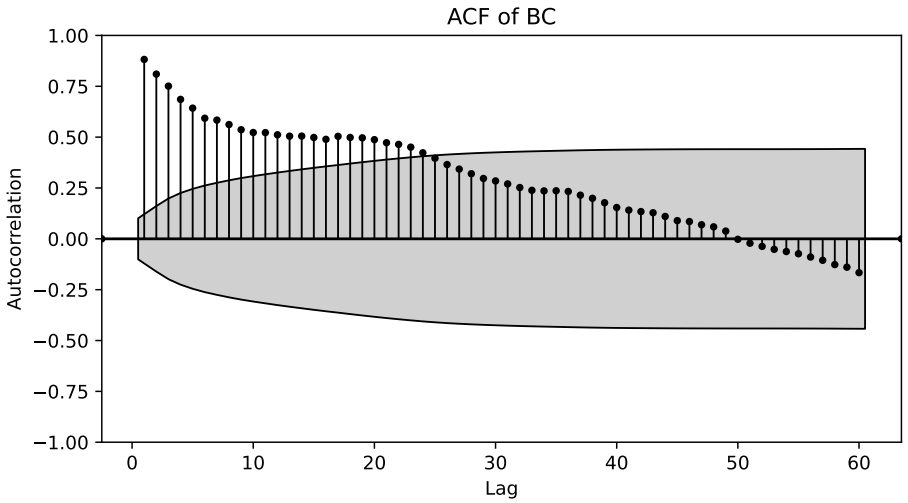


Figure 3.1: ACF of minute-level BC with 95% confidence bands. Significant short-lag correlations indicate temporal dependence/persistence in BC.

trucks and buses emit much more BC than light-duty cars, and can dominate minute-level BC when present [11, 12, 3]. When the model only observes total vehicle counts without any engine or fuel information, minutes containing a small number of diesel vehicles can resemble minutes with only light-duty cars, which can bias predictions in those episodes.

3.2 Integrating Audio Data as a Complementary Modality

We hypothesize that adding an audio sensor input can help the model to consider information that video sometimes misses. For instance, in cases where prediction fails, as per Section 2.3.4, these were specifically due to the camera occluding incoming traffic or traffic activity. However, if audio was included in this data collection, we would have been able to tell or determine the sounds of the engine, the accelerations of passing vehicles. Further, as per Figure 2.14, HDVs were sparse. We were able to verify that this could be due to the occlusion of the streetcar. Adding a microphone can add cues related to heavy engines (e.g., low-frequency energy, specific harmonics), increasing the chance that the model sees the event even if only one or two frames show the vehicle clearly. Simple acoustic features (band power, spectral peaks, event rates) can help distinguish engine load, hard acceleration, and the presence of heavy vehicles. Prior traffic studies also have used roadside microphones to identify vehicle type or motion states in mixed traffic, typically with short-time spectral features and lightweight classifiers [38]. In short, audio has technical value, but real-world use must be guided by governance, transparency, and community engagement so we do not trade one risk for another. Our goal is not to replace video; a combined video+audio design that can fill the gap for video corner cases. Keep using video for geometry and counts, and use audio to flag engine-driven spikes when the camera view is blocked or cannot distinguish engine type. [38].

3.2.1 Audio-based traffic sensing

Field prototypes with stereo microphones report that acoustic cues can separate large vehicles (e.g., trucks, buses) from light-duty cars under real roadside conditions, using modest sampling rates and short windows [38]. Recent research further shows that simple time–frequency features (e.g., band energies, MFCCs) with compact models can classify broad vehicle classes and estimate flow without cameras, though background noise and overlapping sources remain practical limits.

We are getting design takeaways from these prior studies. Mount microphones close to the roadway but away from strong reflections. Use moderate sampling rates (e.g., ∼16 kHz); compute robust spectral features (band energies, spectral centroid, MFCCs) [38, 32].

3.2.2 A Proof of Concept

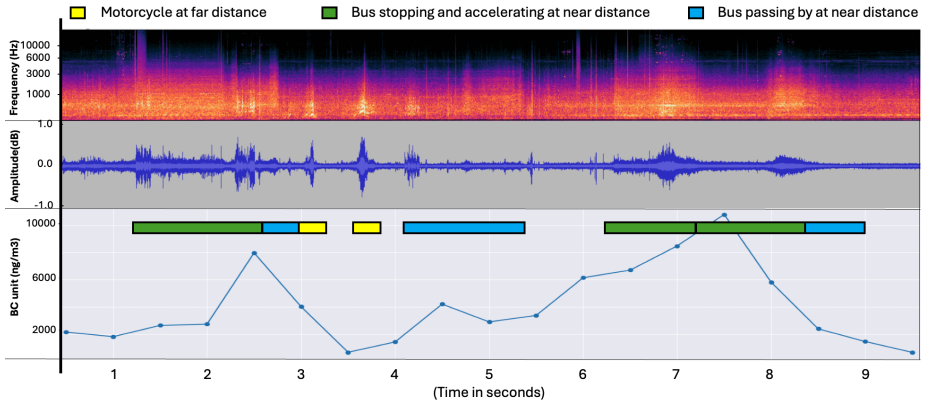


Figure 3.2: Audio signals in spectrogram (top) and waveform (mid) vary with different states of buses. Bottom shows the corresponding BC measured using a microaethalometer.

Figure 3.2, from our bus station data, shows that BC often rises shortly after a bus passes the microphone. We see clear bursts in the audio band that align with bus presence, followed by a brief increase in BC. This relation does not establish causality between sound behavior and BC changes, but it is consistent with studies that use audio features to characterize vehicle type and state. Building on this, in the analysis below, we use a compact feature set to test similarity between changes in BC and changes in sound features, including band-limited energy in 50 to 500 Hz, transient rates during acceleration, and a per-minute count of sound events. Table 3.1 summarizes selected preliminary features, and a complete model comparison is left for future work.

Similarity in changes in BC and sound features

We build on three studies that extract and evaluate sound features for traffic analysis, with the goal of comparing changes in BC with changes in audio features [60, 61, 38]. Prior work proposes a hierarchical pipeline that sorts vehicles into broad classes using low-level audio descriptors such as zero crossing rate, root mean square energy, spectral roll off, and spectral centroid, together with support vector machines and rule-based stages [60]. Follow-up research targets engine-level characteristics and evaluates a wider set of descriptors, including MFCCs, spectral bandwidth, spectral contrast, and harmonicity [61]. A separate study uses stereo microphone arrays to extract spatial features

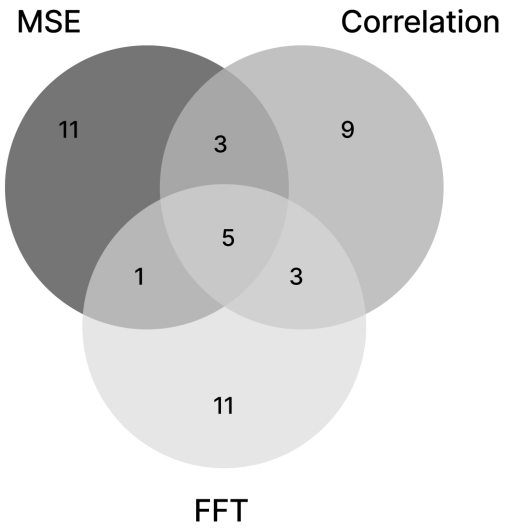


Figure 3.3: Overlap of Top 20 sound Features

Table 3.1: Sound features present in at least two lists. Cells in light gray mark the top five features.

Features in at least two lists	
MAX_SE_IND	ZCR
abs_max	ptp_range
std	A72
BW_20dB	BW_30dB
EnAbove4kHz	SE18
SpectralCentroid	iqr

in addition to basic frequency features, aiming at vehicle classification under real-world noise with minimal roadside infrastructure [38]. Considering the setups described in these works, we extracted sound features from selected parts of our own dataset (Table 2.3). From our audio recordings, we extracted over 30 time-series features. These include ZCR, spectral flux, RMS, roll-off, MFCCs 1–13, spectral centroid, bandwidth, chroma energy distribution, harmonic ratio, ITD, and ILD. All features were downsampled or averaged to match our 1 Hz BC measurements. To find which of these features are more closely related to BC, we compared them to the measured BC concentration using three similarity methods:

1. **Pearson Correlation**: Computes the linear correlation between each audio feature and the BC signal.
2. **FFT Cosine Similarity**: Compares frequency-domain patterns between feature and BC vectors.
3. **Mean Squared Error (MSE)**: Measures the average squared distance between the feature and BC time series.

Each method returned a ranked list of audio features based on how well they match BC changes. Pearson emphasized global correlation. FFT similarity captured alignment in frequency patterns, and MSE reflected absolute time-series shape similarity. The top-ranked features varied by method, but we found overlap across the lists. Table 3.1 shows features those present in at least two lists.

To visualize agreement across metrics, we compared the top 20 features from each similarity method. Figure 3.3 shows the overlap between these feature sets. Five features appeared in all three top-20 lists: `MAX_SE_IND`, `ZCR`, `abs_max`, `ptp_range`, and `std`. These shared features capture signal energy, zero-crossing rate, and amplitude spread. These are properties that reflect engine type and load. In addition, several features appeared in at least two lists. Overlap supports the idea that certain spectral and statistical features consistently co-vary with BC and should be prioritized for modeling.

When we place results side-by-side with our top video features (near-lane counts, stop fractions) and environmental features (e.g., wind speed, humidity), the overlap is easy to read, both sets respond to the same physical events. Video features describe what is happening in the lane; audio features describe how loudly and how often it happens. This overlap adds robustness when one modality is weak (e.g., occlusion).

3.3 Impact on Environmental Justice

Our goal from the start has been to make data on BC concentration more available and to close the data deserts. This thesis shows that traffic video combined with modern machine learning can provide useful local estimates where dedicated instruments are scarce, and therefore can complement sparse monitoring in communities with the greatest need. Environmental justice research on BC and ultrafine particles remains limited overall, but work in Canadian and United States cities highlights clear concerns [5, 4]. In Toronto, for example, Elford et al. found that areas with a higher share of immigrants face greater ultrafine particle exposure during walks to school [62]. These findings are consistent with broader results from United States cities that show higher exposures for low-income and visible minority groups, particularly Black and Hispanic or Latino populations, often in neighborhoods close to major roads and industrial sites [4]. In this context, our video-based traffic approach offers a practical way to approximate BC concentration levels linked to local traffic activity and to increase data availability where routine monitoring is limited. From a policy perspective, tools like ours can strengthen the evidence base for environmental justice action. Fine-scale, time-resolved estimates tied directly to traffic dynamics can help identify when and where exposure is highest and can guide specific interventions such as idling control near schools and intersections, rerouting heavy-duty vehicles away from homes during sensitive hours, and adjusting signal timing where stopped vehicle activity is frequent [63]. Integrating affordable, high-resolution, traffic-derived pollution estimates into planning and public health work can help reduce localized burdens and promote greater equity.

In addition to environmental-justice aims, we believe our technique is also policy-relevant. Specifically, the ability to estimate BC concentrations using publicly accessible traffic cameras and basic environmental data offers a low-cost alternative for identifying unmonitored hotspots. This method links minute-level, lane-level traffic activity to BC right at the curb. This is the scale where cities act with idling control near schools and intersections, targeted truck routing during sensitive hours, and small signal-timing changes where queues form. At the federal level, the CEPA Right to a Healthy Environment Implementation Framework [31] foregrounds equity, access to information, and targeted exposure reduction; our approach can complement these frameworks by identifying when and where local, traffic-driven exposures are highest without claiming readiness for operational deployment. Consistent with the WHO Global Air Quality Guidelines [30], which call for lowering population exposure to combustion-related particles, the method offers a low-cost way to surface unmonitored hotspots that merit further measurement or mitigation. We note limitations, the present system does not quantify non-traffic sources or regional background and requires additional validation before integration into routine policy workflows.

3.4 Conclusion and Next Steps

Traffic video is widely used for transport planning and safety, but its value for studying situations about environmental effects is less explored. In this thesis, we used traffic video to build features that describe vehicle activity (counts by class, stopping, and acceleration) and paired them with environmental features to estimate minute-level BC. Using real-world data and an XGBoost regressor, the model reached $R^2 = 0.72$ and RMSE = 129.42 (ng/m^3). The most influential predictors were wind speed (lower BC when wind is stronger), acceleration after stopping (higher BC during start-ups), and distance to the sensing point (higher BC closer to traffic). These patterns are consistent with how traffic and near-road conditions shape short-term BC levels. Uncaptured regional contributors where identified as a source of error in modeling. We also observed the worst cases when short events were missed in video (e.g., brief queuing behind a streetcar), or when the engine type could not be inferred from body class alone; Yet the overall approach demonstrates the feasibility to support useful, time-resolved BC estimates.

For next steps, we can extend the system along five tracks. First, we can add an audio module to capture cues that video misses (e.g., heavy-engine signatures and hard acceleration), and fuse audio, video, and environmental features in a single model. Second, we may deploy at more sites in Toronto, with priority to areas that lack routine monitoring, and use elevated or multi-angle views to reduce occlusion. Third, we can enhance operations and evaluation by utilizing larger and longer datasets, conducting stress tests under low wind and high humidity conditions, and establishing clear privacy and consent practices for both video and audio. Incorporating a time series model can also be studied to solve the autocorrelation challenge with the BC signal.

Finally, future work could also explore collaborations with local planning bodies or community groups to assess the practical utility of such tools in real-world decision-making contexts. These documents signal a federal focus on environmental justice and equitable participation.

The goal is a practical, lower-cost pipeline that delivers local BC estimates where they are most needed, while keeping an understanding of limits and uncertainty.

Funding Statement

This research initiative was supported by the Natural Sciences and Engineering Research Council of Canada, RGPIN-2025-06914.

Appendices

A1 Data Samples and Detailed Results

This appendix complements the raw and processed dataset described in Chapter 2. It shows brief, representative snippets from the raw data files and notes the expected fields.

A1.1 Raw data example from microAeth AE51 (CSV excerpt)

Listing 1: Sample rows from microAeth AE51

```
Date,Time,Ref,Sen,ATN,Flow,Pcb temp,Status,Battery,BC,Ona_#_pts_avg
2024/11/04,18:49:00,890665,921559,-3.40984263756,100,19,0,98,,NULL
2024/11/04,18:49:30,890783,921490,-3.3891073947402,99,19,0,98,2379,3
2024/11/04,18:50:00,890907,921527,-3.3792031816136,99,19,0,98,1136,3
2024/11/04,18:50:30,890941,921473,-3.369526908093,100,19,0,98,1099,3
2024/11/04,18:51:00,891037,921486,-3.3601631390269,100,19,0,98,1064,2
```

A1.2 Dataset Example (JSON excerpt)

This is an example of the final dataset row before passing to the model training pipeline. The field "BC post" indicates BC after applying ONA (see Section 2.2.3).

```
{
  "Time": "2024-11-05 17:34:50",
  "BC": 1202,
  "BC post": 1114,
  "car_line1": 1,
  "car_line3": 3,
  "truck_line2": 1,
  "car_line2": 2,
  "car_line1_stop": 0,
  "truck_line2_stop": 0,
  "truck_line3": 0,
  "truck_line1": 0,
  "traffic": 0.65,
  "history_temperature": 19.1,
  "history_wind_speed": 24.5,
  "history_humidity": 67,
  "forecast_temperature": 18.7,
  "forecast_wind_speed": 9.9,
  "forecast_humidity": 68
}
```

A1.3 Video Processing Event Log

This example from the video processing pipeline shows an event log of detections in the video. When an object is first seen, the system assigns a unique ID and records its location and timestamp (see Section 2.1.2).

Listing 2: Example event log of detected objects with IDs and timestamps

```
car_line1 : 2771 2024-11-05 17:23:02
car_line1 : 2775 2024-11-05 17:23:03
car_line2 : 2790 2024-11-05 17:23:05
truck_line2 : 2777 2024-11-05 17:23:07
person_line1 : 2785 2024-11-05 17:23:09
bicycle_line1 : 2795 2024-11-05 17:23:09
```

A1.4 Hyperparameter Tuning Protocol

We tuned model hyperparameters using cross-validated grid search in `scikit-learn`. Each model was wrapped in a common `Pipeline` so that parameters are referenced with the double-underscore convention, for example `model__max_depth`. The search spaces mirror the configuration used in our code.

Model	Parameter	Values
LinearRegression	(none)	No hyperparameters tuned
RandomForestRegressor	`model__min_samples_split`	[2, 5]
	`model__n_estimators`	[30, 50, 100, 200]
	`model__max_depth`	[None, 5, 6, 7, 9, 10]
GradientBoostingRegressor	`model__n_estimators`	[20, 50, 100, 200, 250]
	`model__learning_rate`	[0.01, 0.05, 0.1, 0.2]
	`model__max_depth`	[3, 5, 6, 7]
SVR	`model__C`	[0.1, 1, 10]
	`model__epsilon`	[0.01, 0.1, 1]
XGBRegressor	`model__n_estimators`	[20, 50, 100, 200, 250]
	`model__learning_rate`	[0.01, 0.05, 0.1, 0.2]
	`model__max_depth`	[3, 5, 6, 7]

Procedure. We ran a cross-validated grid search per model and selected the configuration with the best mean validation score. Unless noted otherwise, we used five-fold cross-validation, refit the best configuration on the full training split, and reported scores on the held-out test split. A fixed `random_state` ensured repeatability.

Listing 3: Skeleton used for cross-validated grid search

```
search = GridSearchCV(
    estimator=pipe,                 # e.g., Pipeline([('prep', ...), ('model', ...)])
    param_grid=models[name]['params'],
    scoring='neg_root_mean_squared_error',
    cv=5, n_jobs=-1, refit=True
)
search.fit(X_train, y_train)
best_params = search.best_params_
cv_table   = search.cv_results_
```

Notes by algorithm.

- **RandomForestRegressor** tunes tree count (`n_estimators`), depth (`max_depth`), and split size (`min_samples_split`) to control ensemble capacity and variance.
- **GradientBoostingRegressor** balances `n_estimators` and `learning_rate`; `max_depth` limits individual tree depth.
- **SVR** primarily depends on `C` (regularization) and `epsilon` (tube width). We used the default RBF kernel unless specified.
- **XGBRegressor** varies `n_estimators`, `learning_rate`, and `max_depth` under the squared-error objective.

Reproducibility. We preserve the full grid, cross-validation scores, and selected settings by saving `cv_results_` and `best_params_` artifacts for each run. These artifacts are archived alongside experiment configs.

A1.5 Model Performance Across Noise Reduction Techniques

This figure complements the results in Table 2.7, which reports R^2 for XGBoost with different noise reduction methods. Here we show the same comparison across all models evaluated in this thesis.

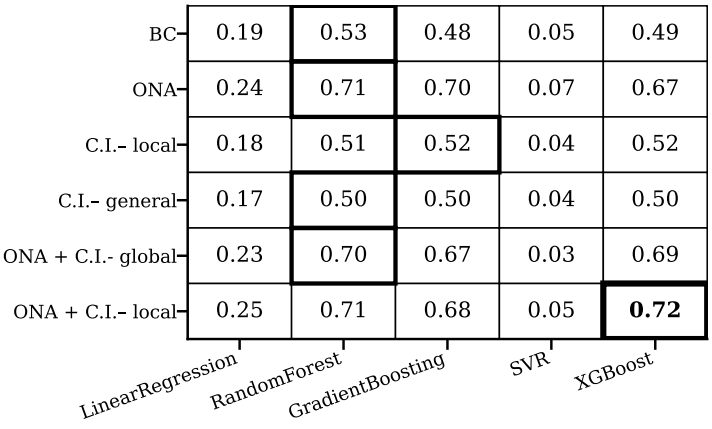


Figure A1: Performance comparison across noise reduction techniques for all models. Points show mean R^2. Models include Linear Regression, Random Forest, Gradient Boosting, SVR, and XGBoost.

A2 Dashboard Notes

A2.1 Early Database Schema

This figure shows an early iteration of the database schema used for the data collection dashboard (see Section 2.2.1). It captures the core entities and relationships considered during initial development.

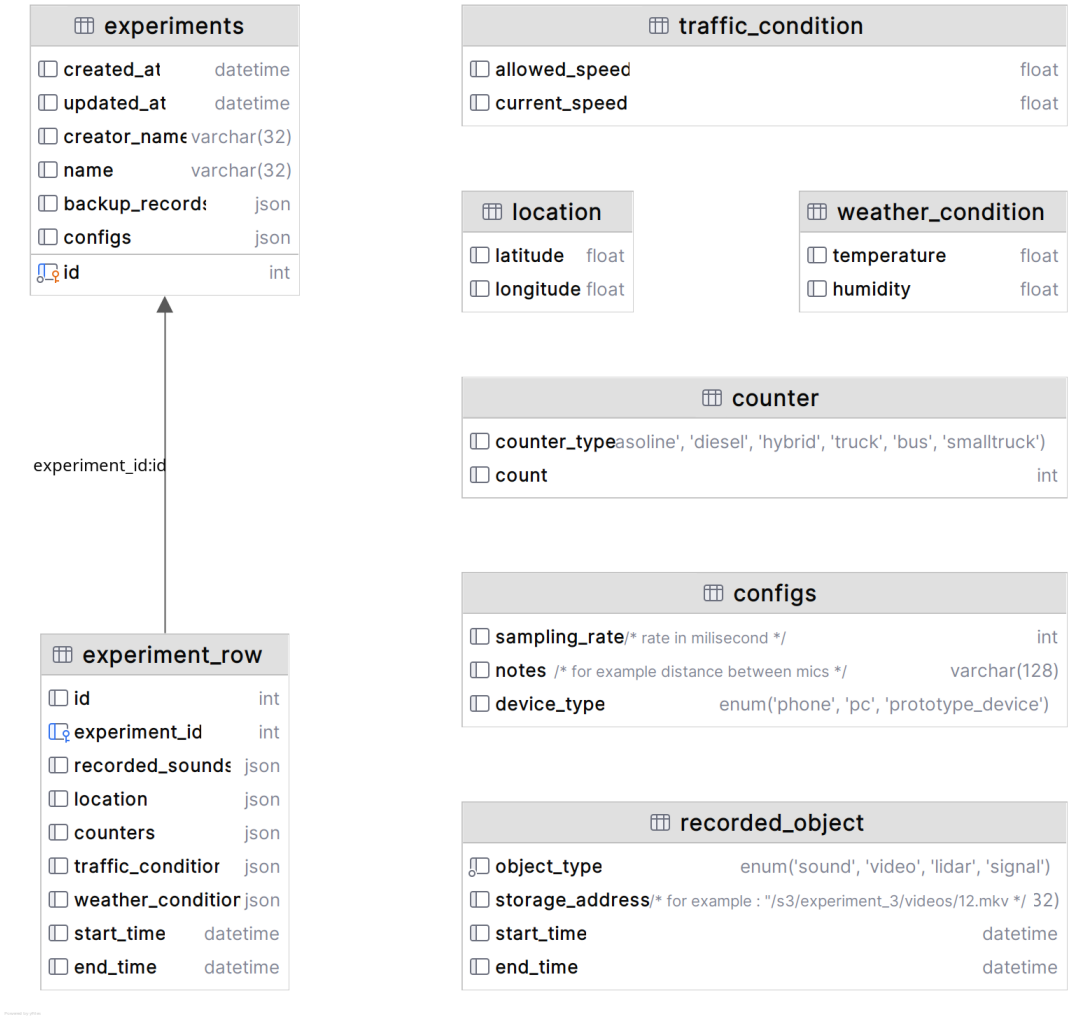


Figure A2: Early database schema for the data collection dashboard.

A2.2 Running the Stack with Docker Compose

We use Docker Compose to start the dashboard and its services from a Compose file named `infrastructure.yaml`. The commands below assume you run them at the repository root.

```
# Bring services up (detached).
docker compose -f infrastructure.yaml up -d
```

```
# Check status.
docker compose -f infrastructure.yaml ps
```

```
# Stop and remove the stack.
docker compose -f infrastructure.yaml down
```

Bibliography

[1] HEI Panel on the Health Effects of Traffic-Related Air Pollution. *Traffic-Related Air Pollution: A Critical Review of the Literature on Emissions, Exposure, and Health Effects*. Tech. rep. Health Effects Institute, 2010. URL: https://www.healtheffects.org/publication/traffic-related-air-pollution-critical-review-literature-emissions-exposure-and-health.

[2] T. C. Bond, S. J. Doherty, D. W. Fahey, et al. "Bounding the role of black carbon in the climate system: A scientific assessment". In: *Journal of Geophysical Research: Atmospheres* 118.11 (2013), pp. 5380–5552. DOI: 10.1002/jgrd.50171.

[3] Nathan Hilker et al. "Traffic-related air pollution near roadways: discerning local impacts from background". In: *Atmospheric Measurement Techniques* 12.10 (2019), pp. 5247–5261. DOI: 10.5194/amt-12-5247-2019. URL: https://amt.copernicus.org/articles/12/5247/2019/amt-12-5247-2019.html.

[4] S. E. Chambliss et al. "Local- and regional-scale racial and ethnic disparities in air pollution determined by long-term mobile monitoring". In: *Proceedings of the National Academy of Sciences* 118.37 (2021), e2109249118. DOI: 10.1073/pnas.2109249118.

[5] Emmanuelle Batisse et al. "Examining the social distributions in neighbourhood black carbon and ultrafine particles in Montreal and Toronto, Canada". In: *Environment International* 198 (2025), p. 109395. DOI: 10.1016/j.envint.2025.109395.

[6] Ontario Ministry of the Environment, Conservation and Parks. *Ontario's Air Monitoring Stations (AQHI locations: Toronto Downtown, Toronto East, Toronto North, Toronto West)*. 2025. URL: https://www.airqualityontario.com/aqhi/locations.php.

[7] P. de Souza and P. L. Kinney. "Socioeconomic disparities of low-cost air-quality sensors in California". In: *American Journal of Public Health* 111.10 (2021), pp. 1901–1904. DOI: 10.2105/AJPH.2021.306603.

[8] A. D. A. Hansen, H. Rosen, and T. Novakov. "The aethalometer—an instrument for the real-time measurement of optical absorption by aerosol particles". In: *Science of the Total Environment* 36 (1984), pp. 191–196. DOI: 10.1016/0048-9697(84)90265-1.

[9] J. S. Apte, K. P. Messier, S. Gani, et al. "High-resolution air pollution mapping with Google Street View cars: Exploiting big data". In: *Environmental Science & Technology* 51.12 (2017), pp. 6999–7008. DOI: 10.1021/acs.est.7b00891.

[10] World Health Organization. *Regulation and Governance of Artificial Intelligence for Health*. https://www.who.int/publications/i/item/9789240084759. 2023.

[11] G. A. Ban-Weiss et al. “Measurement of black carbon and particle number emission factors from individual heavy-duty trucks”. In: *Environmental Science & Technology* 43.4 (2009), pp. 1419–1424. DOI: 10.1021/es8021039.

[12] X. Zheng, Y. Wu, S. Zhang, et al. “Characteristics of black carbon emissions from in-use light-duty passenger vehicles”. In: *Environmental Pollution* 231 (2017), pp. 348–356. DOI: 10.1016/j.envpol.2017.08.002.

[13] City of Toronto. *Traffic Camera Locations—Open Data Portal*. https://open.toronto.ca/dataset/traffic-camera-locations. 2025.

[14] Glenn Jocher, Ayush Chaurasia, and Jing Qiu. *Ultralytics YOLO*. https://github.com/ultralytics/ultralytics. 2023. (Visited on 08/20/2025).

[15] Health Canada. *Health Impacts of Traffic-Related Air Pollution in Canada*. Tech. rep. 2022. URL: https://www.canada.ca/en/health-canada/services/publications/healthy-living/health-impacts-traffic-related-air-pollution.html.

[16] Gerard Hoek et al. “Long-term air pollution exposure and cardio-respiratory mortality: a review”. In: *Environmental Health* 12.1 (2013), p. 43. DOI: 10.1186/1476-069X-12-43. URL: https://ehjournal.biomedcentral.com/articles/10.1186/1476-069X-12-43.

[17] *The Health Impacts of Air Pollution in Canada*. Tech. rep. Estimated 15,300 premature deaths per year in Canada (2016 baseline). Ottawa, ON: Health Canada, 2021. URL: https://publications.gc.ca/collections/collection_2022/sc-hc/H2-3-2021-eng.pdf.

[18] C. Guo et al. “Effect of long-term exposure to fine particulate matter on lung function decline and risk of chronic obstructive pulmonary disease in Taiwan: a longitudinal, cohort study”. In: *The Lancet Planetary Health* 2.3 (2018), e114–e125. DOI: 10.1016/S2542-5196(18)30028-7.

[19] M. I. Mead, O. A. Popoola, G. B. Stewart, et al. “The use of electrochemical sensors for monitoring urban air quality in low-cost, high-density networks”. In: *Atmospheric Environment* 70 (2013), pp. 186–203. DOI: 10.1016/j.atmosenv.2012.11.060.

[20] Yan Zhang et al. “Application of artificial neural networks in atmospheric environmental prediction and assessment: A review”. In: *Atmospheric Environment* 51 (2012), pp. 1–11. DOI: 10.1016/j.atmosenv.2011.12.031.

[21] K. de Hoogh, A. Saucy, A. Shtein, et al. “Predicting fine-scale daily NO_2 for 2005–2016 incorporating OMI satellite data across Switzerland”. In: *Environmental Science & Technology* 53.17 (2019), pp. 10279–10287. DOI: 10.1021/acs.est.9b03107.

[22] Z. Fan et al. “Enhancing Urban Real-Time PM2.5 Monitoring in Street Canyons by Machine Learning and Computer Vision Technology”. In: *Sustainable Cities and Society* 105 (2025), p. 105009. DOI: 10.1016/j.scs.2023.105009.

[23] Y. Chen, J. Zhou, and J. Lu. “Vision-based vehicle detection and counting for urban road traffic monitoring”. In: *Sensors* 20.5 (2020), p. 1322. DOI: 10.3390/s20051322.

[24] Xinyu Han, Jing Wang, Wenlong Yang, et al. “Deep learning approaches for air pollution prediction: A review and bibliometric analysis”. In: *Atmospheric Environment* 277 (2022), p. 119087. DOI: 10.1016/j.atmosenv.2022.119087.

[25] Scott M. Lundberg and Su-In Lee. "A Unified Approach to Interpreting Model Predictions". In: *Advances in Neural Information Processing Systems 30 (NeurIPS 2017)*. 2017. URL: https://proceedings.neurips.cc/paper/7062-a-unified-approach-to-interpreting-model-predictions.pdf.

[26] Q. Zhang et al. "Spatiotemporal variations of black carbon concentrations in Canadian cities: Land use regression modeling with high resolution mobile monitoring data". In: *Environmental Science & Technology* 53.14 (2019), pp. 7740–7749. DOI: 10.1021/acs.est.9b01684.

[27] Scott Weichenthal et al. "Characterizing the Impact of Traffic and the Built Environment on Near-road Ultrafine Particle and Black Carbon Concentrations". In: *Environmental Research* 132 (2014), pp. 305–310. DOI: 10.1016/j.envres.2014.04.007.

[28] Marianne Hatzopoulou and Eric J. Miller. "Linking an activity-based travel demand model with traffic emission and dispersion models: Transport's contribution to air pollution in Toronto". In: *Transportation Research Part D: Transport and Environment* 15.6 (2010), pp. 315–325. DOI: 10.1016/j.trd.2010.03.007.

[29] UNESCO. *Recommendation on the Ethics of Artificial Intelligence*. https://unesdoc.unesco.org/ark:/48223/pf0000381137. 2021.

[30] World Health Organization. *WHO Global Air Quality Guidelines: Particulate Matter (PM2.5 and PM10), Ozone, Nitrogen Dioxide, Sulfur Dioxide and Carbon Monoxide*. Geneva: World Health Organization, 2021. ISBN: 9789240034228. URL: https://www.who.int/publications/i/item/9789240034228.

[31] Environment and Climate Change Canada and Health Canada. *Implementation Framework for the Right to a Healthy Environment under the Canadian Environmental Protection Act (CEPA)*. 2025. URL: https://www.canada.ca/en/environment-climate-change/services/canadian-environmental-protection-act-registry/publications/right-healthy-environment-implementation-framework.html.

[32] Michael R. Wilby and Carolina Quintana. "A review of traffic monitoring systems in the urban environment". In: *European Transport Research Review* 3.3 (2011), pp. 145–152. DOI: 10.1007/s12544-011-0051-7.

[33] Mohamed Aly. "Real time Detection of Lane Markers in Urban Streets". In: *Proceedings of the IEEE Intelligent Vehicles Symposium (IV)*. Eindhoven, The Netherlands, June 2008. URL: https://github.com/mohamedadaly/caltech-lane-detection.

[34] Patrick Zippenfenig. *Open-Meteo.com Weather API*. 2023. DOI: 10.5281/zenodo.7970649. URL: https://open-meteo.com/.

[35] Tsung-Yi Lin et al. "Microsoft COCO: Common Objects in Context". In: *European Conference on Computer Vision (ECCV)*. Springer, 2014, pp. 740–755.

[36] Richard O. Duda and Peter E. Hart. "Use of the Hough transformation to detect lines and curves in pictures". In: *Communications of the ACM* 15.1 (1972), pp. 11–15. DOI: 10.1145/361237.361242.

[37] John Canny. "A Computational Approach to Edge Detection". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* PAMI-8.6 (1986), pp. 679–698. DOI: 10.1109/TPAMI.1986.4767851.

[38] Billy Dawton et al. “Initial Evaluation of Vehicle Type Identification using Roadside Stereo Microphones”. In: *2020 IEEE Sensors Applications Symposium (SAS)*. Kuala Lumpur, Malaysia, 2020, pp. 1–6. DOI: 10.1109/SAS48726.2020.9220076.

[39] Pallets Projects. *Flask Web Framework*. Accessed 2025-08-18. 2010. URL: https://flask.palletsprojects.com/.

[40] Bootstrap Authors. *Bootstrap v5 Documentation*. Accessed 2025-08-18. 2024. URL: https://getbootstrap.com/.

[41] SQLite Consortium. *SQLite Documentation*. Accessed 2025-08-18. 2024. URL: https://www.sqlite.org/docs.html.

[42] Dirk Merkel. “Docker: Lightweight Linux Containers for Consistent Development and Deployment”. In: *Linux Journal* 239 (2014). URL: https://www.linuxjournal.com/content/docker-lightweight-linux-containers-consistent-development-and-deployment.

[43] David L. Mills et al. *RFC 5905: Network Time Protocol Version 4: Protocol and Algorithms Specification*. Internet Engineering Task Force (IETF) Request for Comments. June 2010. URL: https://www.rfc-editor.org/rfc/rfc5905.

[44] *IEEE Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems (IEEE 1588-2008)*. IEEE, 2008. DOI: 10.1109/IEEESTD.2008.4579760.

[45] AethLabs. *How is the ONA algorithm mathematically implemented on the AethLabs.com data Dashboard?* AethLabs Help Center. 2020. (Visited on 08/17/2025).

[46] Alan V. Oppenheim, Ronald W. Schafer, and John R. Buck. *Discrete-Time Signal Processing*. 2nd ed. Upper Saddle River, NJ: Prentice Hall, 1999.

[47] Ronald N. Bracewell. *The Fourier Transform and Its Applications*. 2nd ed. New York: McGraw-Hill, 1986. ISBN: 0070070156.

[48] William G. Cochran. *Sampling Techniques*. 3rd. New York: Wiley, 1977.

[49] U.S. Environmental Protection Agency. *Air Sensor Performance Targets and Testing Protocols*. Accessed July 2025. 2021. URL: https://www.epa.gov/air-sensor-toolbox/air-sensor-performance-targets-and-testing-protocols.

[50] scikit-learn developers. *Cross-validation: Evaluating Estimator Performance*. scikit-learn User Guide. Accessed 2025-08-18. 2024. URL: https://scikit-learn.org/stable/modules/cross_validation.html.

[51] Tianqi Chen and Carlos Guestrin. “XGBoost: A Scalable Tree Boosting System”. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 2016, pp. 785–794. DOI: 10.1145/2939672.2939785.

[52] John L. Durant et al. “Short-term Variation in Near-highway Air Pollutant Gradients on a Winter Morning”. In: *Atmospheric Chemistry and Physics* 10 (2010), pp. 8341–8352. URL: https://acp.copernicus.org/articles/10/8341/2010/acp-10-8341-2010.pdf.

[53] William Farrell et al. “Near Roadway Air Pollution Across a Spatially Extensive Road and Cycling Network”. In: *Environmental Pollution* 212 (2016), pp. 498–507. DOI: 10.1016/j.envpol.2016.01.088.

[54] Wei Lv, Yan Wu, and Jian Zang. "A Review on the Dispersion and Distribution Characteristics of Pollutants in Street Canyons and Improvement Measures". In: *Energies* 14.19 (2021), p. 6155. DOI: 10.3390/en14196155.

[55] Yuan Dong Huang et al. "Effects of Wind Direction on the Airflow and Pollutant Dispersion inside a Long Street Canyon". In: *Aerosol and Air Quality Research* 19.5 (2019), pp. 1152–1171. DOI: 10.4209/aaqr.2018.09.0344.

[56] Maryam Shekarrizfard et al. "Modelling the Spatio-Temporal Distribution of Ambient Nitrogen Dioxide and Investigating the Effects of Public Transit Policies on Population Exposure". In: *Environmental Modelling & Software* 90 (2017), pp. 186–198. DOI: 10.1016/j.envsoft.2017.02.007.

[57] Shaojun Zhang et al. "Black carbon pollution for a major road in Beijing: Implications for policy interventions of the heavy-duty truck fleet". In: *Transportation Research Part D: Transport and Environment* 68 (2019), pp. 110–121. DOI: 10.1016/j.trd.2017.07.013.

[58] Alper Yilmaz, Omar Javed, and Mubarak Shah. "Object Tracking: A Survey". In: *ACM Computing Surveys* 38.4 (2006), p. 13. DOI: 10.1145/1177352.1177355.

[59] George E. P. Box et al. *Time Series Analysis: Forecasting and Control*. 5th ed. Hoboken, NJ: John Wiley & Sons, 2015. URL: https://www.wiley.com/en-us/Time+Series+Analysis%3A+Forecasting+and+Control%2C+5th+Edition-p-9781118675021.

[60] Elżbieta Kubera, Alicja A. Wieczorkowska, and Krzysztof Skrzypiec. "Audio-Based Hierarchic Vehicle Classification for Intelligent Transportation Systems". In: *Foundations of Intelligent Systems*. Ed. by Floriana Esposito et al. Vol. 9384. Lecture Notes in Artificial Intelligence. Cham: Springer, 2015, pp. 343–352. DOI: 10.1007/978-3-319-25252-0_37.

[61] Alicja Wieczorkowska et al. "Spectral features for audio based vehicle and engine classification". In: *Journal of Intelligent Information Systems* 50.2 (2018), pp. 265–290. DOI: 10.1007/s10844-017-0459-2.

[62] Spencer Elford and Matthew D Adams. "Associations between socioeconomic status and ultrafine particulate exposure in the school commute: An environmental inequality study for Toronto, Canada". In: *Environmental Research* 192 (2021), p. 110224.

[63] J. Zalzal and M. Hatzopoulou. "Fifteen years of community exposure to heavy-duty emissions: capturing disparities over space and time". In: *Environmental Science & Technology* 56.23 (2022), pp. 17134–17142. DOI: 10.1021/acs.est.2c04320.