Title

The Effects of Truck Idling and Searching for Parking on Disadvantaged Communities

Authors

Jaller, Miguel, PhD

Xiao, Runhua (Ivan)

Publication Date

2023-06-01

DOI

10.7922/G2SX6BJB

RESEARCH REPORT

Institute of Transportation Studies | UNIVERSITY OF CALIFORNIA

The Effects of Truck Idling and Searching for Parking on Disadvantaged Communities

Miguel Jaller, Ph.D., Associate Professor, Department of Civil and Environmental Engineering, University of California, Davis
Runhua (Ivan) Xiao, M.S., Researcher, Department of Civil and Environmental Engineering, University of California, Davis

June 2023

ITS UCDAVIS Institute of Transportation Studies

Report No.: UC-ITS-2022-02 | DOI: 10.25338/B8QD38

Technical Report Documentation Page

1. Report No. UC-ITS-2022-02	**2. Government Accession No.** N/A	**3. Recipient's Catalog No.** N/A
4. Title and Subtitle The Effects of Truck Idling and Searching for Parking on Disadvantaged Communities		**5. Report Date** June 2023 **6. Performing Organization Code** ITS-Davis
7. Author(s) Miguel Jaller, PhD, https://orcid.org/0000-0003-4053-750X Runhua (Ivan) Xiao, https://orcid.org/0000-0002-9676-8334		**8. Performing Organization Report No.** UCD-ITS-RR-23-65
9. Performing Organization Name and Address Institute of Transportation Studies, Davis 1605 Tilia Street Davis, CA 95616		**10. Work Unit No.** N/A **11. Contract or Grant No.** UC-ITS-2022-02
12. Sponsoring Agency Name and Address The University of California Institute of Transportation Studies www.ucits.org		**13. Type of Report and Period Covered** Research Report (January 2022 – June 2023) **14. Sponsoring Agency Code** UC ITS
15. Supplementary Notes DOI:10.7922/G2SX6BJB Supporting Dataset DOI:10.25338/B8QD38		

16. Abstract

This project identifies factors that affect three truck-related parameters: idling, searching for parking, and parking demand. These parameters are examined in communities in Kern County California that have high air pollution levels and are located near transportation corridors, industrial facilities, and logistics centers. Daytime truck idling is concentrated in and around commercial and industrial hubs, and nighttime idling is concentrated around major roads and highway entrances and exits. Truck idling, searching for parking, and parking demand correlate with shorter distances from freight-related points-of-interest such as warehouses, increased size of nearby industrial or commercial land use, and proximity to areas of dense population or income inequality. Based on these findings, policy recommendations include targeted anti-idling interventions, improved truck parking facilities, parking systems that provide real-time availability information to drivers, provision of alternate power sources in parking facilities to allow trucks to turn off, cleaner fuels and technologies, enhanced routing efficiency, stricter emission standards, and stronger land-use planning with buffer zones around residential areas.

17. Key Words Trucking, trucks, parking demand, engine idling, air pollution, industrial areas, underserved communities, environmental justice		**18. Distribution Statement** No restrictions.	
19. Security Classification (of this report) Unclassified	**20. Security Classification (of this page)** Unclassified	**21. No. of Pages** 77	**22. Price** N/A

Form Dot F 1700.7 (8-72) Reproduction of completed page authorized

About the UC Institute of Transportation Studies

The University of California Institute of Transportation Studies (UC ITS) is a network of faculty, research and administrative staff, and students dedicated to advancing the state of the art in transportation engineering, planning, and policy for the people of California. Established by the Legislature in 1947, ITS has branches at UC Berkeley, UC Davis, UC Irvine, and UCLA.

Acknowledgments

This study was made possible with funding received by the University of California Institute of Transportation Studies from the State of California through the Road Repair and Accountability Act of 2017 (Senate Bill 1). The authors would like to thank the State of California for its support of university-based research, and especially for the funding received for this project. The authors are grateful for the invaluable data provided by various sources, including Geotab, CalEnviroScreen, the US Census Bureau, and OpenStreetMap. We appreciate the organizations and individuals responsible for maintaining and sharing these datasets, which have significantly contributed to the depth and quality of our analysis.

The Effects of Truck Idling and Searching for Parking on Disadvantaged Communities

Miguel Jaller, Ph.D., Associate Professor, Department of Civil and Environmental Engineering, University of California, Davis
Runhua (Ivan) Xiao, M.S., Researcher, Department of Civil and Environmental Engineering, University of California, Davis

June 2023

UC DAVIS Institute of Transportation Studies

Report No.: UC-ITS-2022-02 | DOI: 10.25338/B8QD38

Table of Contents

Table of Contents

List of Tables

List of Figures

Glossary

AADT	annual average daily traffic
Adj R2	adjusted R-squared
AIC	Akaike information criterion
CNN	convolutional neural network
CV	cross-validation
DT/RF	decision trees/random forests
DV	dependent variable
HDT	heavy-duty truck
IV	independent variable
KDE	kernel density estimation
KNN	K-nearest neighbor
LL	log likelihood
MAE	mean absolute error
MSE	mean squared error
NN	neural network
PR2	pseudo R squared
R2	R-squared
RF	Random Forest
RMSE	root mean squared error
SEM	spatial error model
TGT	total geohashes traveled
UGT	unique geohashes traveled

Executive Summary

Executive Summary

Several communities in California experience a disproportionate impact from air pollution due to factors such as their proximity to transportation corridors, industrial installations, and logistics centers. Compared to other communities, these typically have lower socioeconomic status, higher population density, and a higher occurrence of health conditions linked to poor air quality. Assembly Bill 617 (AB 617) was implemented in 2017 with the goal of decreasing air pollution and improving public health in targeted areas by conducting thorough monitoring, community engagement, and the execution of effective air quality management strategies.

This study investigates truck-related activity in a sample of AB 617 communities, particularly concerning truck idling, searching for parking, and truck parking demand. By utilizing advanced modeling techniques and comprehensive datasets, this research seeks to identify the critical factors influencing these activities and their consequences for air quality and environmental health in the AB 617 communities. In doing so, the study gathered various datasets that encompassed data on the communities' socio-economic status, road networks, land use, and points-of-interest, which in this context are logistics centers and warehouses. In addition, the research team collected data on truck idling, parking search times, and parking demand. These extensive and diverse datasets provided a firm base for creating reliable models to predict and comprehend the factors influencing truck-related activities and their impact on air quality in AB 617 communities.

The study used and compared a range of modeling techniques, including random forest (RF), Convolutional Neural Network, Bayesian Ridge Regression, and Spatial Error Model (SEM). These models were chosen based on their ability to accurately predict and analyze the factors influencing truck idling, parking search times, and parking demand. For instance, the RF and CNN models were used to predict average idle time, while the RF and Bayesian Ridge Regression models were used to predict truck parking search times. The RF model was also applied to analyze truck parking demand, and the SEM model was used to examine the relationships between environmental factors and air pollution. The models' performance was assessed using several metrics, such as mean squared error, root mean squared error, mean absolute error, and model accuracy. The SEM model also provided valuable insights into the connections between truck-related activities, socio-economic variables, and air pollution.

The study revealed that various factors significantly influenced average idle time, parking search times, and parking demand in AB 617 communities. Idle time in an area correlated positively with the percentage of vehicles present that are heavy-duty trucks. It also correlated positively with land use, distance from points-of-interest, and markers of income inequality. Moreover, heavy-duty trucks had the highest average idle time, followed by medium- and light-duty trucks. Disparities in idling activities were also noted across truck types and geographical areas. Longer parking search times correlated with the distance from main roads and the number of employees in various industries, while vehicle type also played a notable role in these variations. The study also identified parking search hotspots in several areas, highlighting the need for targeted interventions to improve parking efficiency. Higher truck parking demand had a significant positive correlation with amount of commercial land area nearby (within 800 m), distance to main roads, income inequality, and population age and lower education levels. In addition, inferred truck parking demand strongly correlated with higher PM 2.5 concentration, while socio-economic variables were found to be significant in explaining variations in environmental indicators. These findings have important implications for policymakers and stakeholders in developing effective strategies to address air pollution and its associated health impacts in AB 617 communities.

Based on the study's findings, we have proposed preliminary policy recommendations to address the issues in these communities. These include implementing targeted anti-idling campaigns and regulations, improving truck parking facilities and management, promoting cleaner fuels and technologies, enhancing truck routing efficiency through intelligent transportation systems, implementing stricter emission standards for vehicles, and strengthening land-use planning and zoning regulations to create buffer zones between industrial, commercial, and residential areas while promoting mixed-use developments that encourage sustainable transportation options. These recommendations aim to improve air quality and public health in the affected communities.

In summary, this study sheds light on the relationship between truck-related activities and air quality in AB 617 communities, providing valuable insights into factors correlating with truck idling, parking search behavior, and truck parking demand. Our key findings inform preliminary policy recommendations, which aim to improve air quality and public health in these communities. By implementing parking and idling strategies, we can significantly improve air quality and public health in AB 617 communities while laying the groundwork for more comprehensive policy interventions.

Contents

Introduction

Truck idling and searching for parking are significant contributors to air pollution, disproportionately affecting environmental justice communities. These communities—often located near highways, ports, and freight distribution centers—experience the harmful effects of emissions from heavy-duty diesel trucks, such as nitrogen oxides (NOx) and carbon dioxide (CO_2). Major highway exposure studies in the past have estimated that residents living within a hundred yards of a major highway are exposed to higher than the background level of particulate matter (PM) emissions typically measured in ambient air (Harris, 2014). Furthermore, the World Health Organization (WHO) has identified diesel exhaust as a Group 1 carcinogen, linking it to an increased risk of lung cancer (Turner et al., 2020).

In California, over 1.8 million residents, predominantly low-income and minority populations, live within 500 feet of a busy road where they are exposed to high levels of pollution from truck traffic and other vehicles (Hopkins, 2017). These communities suffer from higher asthma rates, cardiovascular diseases, and other respiratory illnesses due to increased truck emissions exposure (Bailey and Solomon, 2004). In addition, truck idling and searching for parking contribute to traffic congestion and noise pollution, further increasing the cost of freight transportation and worsening the quality of life in these areas (Farahani, 2011; Giuliano and O'Brien, 2007; Jaller et al., 2020; Porter et al., 2022; Weinberger et al., 2020).

In the United States, truck idling alone wastes over 6 billion gallons of fuel annually, costing an estimated $20 billion (Gaines and Levinson, 2011). Typically, a long-haul truck idles for about 1,800 hours per year, using about 1,500 gallons of diesel (Profits, B.U., 2010). Diesel heavy-duty trucks emit approximately 20-30 times more NOx than passenger cars, contributing significantly to air pollution (Jonson et al., 2017). In sum, truck idling contributes to over 11 million tons of CO_2 and 180,000+ tons of NOx emissions per year within 42 US states (Frey and Kuo, 2009). Furthermore, a study conducted in Wilmington, California, found that trucks idling were responsible for up to 30% of the total NOx emissions (Park et al., 2011). These emissions have severe health impacts on communities living close to areas with high truck activity, such as highways and ports.

Efforts to reduce truck idling and searching for parking have been explored through various approaches. For instance, California's toxic air contaminant (TAC) identification and control program, which included regulations to reduce emissions from various sources, led to significant reductions in TAC emissions from 1990 to 2012, including a 68% decline in diesel particulate matter (DPM) despite a population increase (Propper et al., 2015). Additionally, deploying alternative fuel and electric trucks can significantly reduce NOx and CO2 emissions by up to 80% compared to traditional diesel-powered trucks (Jaller et al., 2020). Intelligent parking systems represent another approach to addressing the issue. Such systems rely on real-time data on parking availability, enabling drivers to locate suitable parking spots more efficiently. By providing real-time parking information to help drivers find available parking spaces more quickly, these systems could, for example, help London save £183.6 million worth of fuel per year, reduce CO2 emissions by 642,978 tons, and reduce traffic congestion and urban air pollution (Gao et al., 2020).

However, despite those measures and advances, more targeted actions are needed to address disadvantaged communities' specific needs and mitigate the adverse impacts of truck idling and parking emissions on their health and well-being. For example, it is possible that intelligent parking systems, while reducing emissions overall, may not significantly decrease exposure to harmful pollutants for low-income and minority communities, necessitating more targeted interventions to address their specific needs. Thus, understanding the relationship between truck idling, searching for parking, and the resulting air pollution is critical to developing effective interventions to address the unique challenges disadvantaged communities face. By conducting a comprehensive analysis of available data on truck parking and idling patterns and incorporating information on disadvantaged communities' physical infrastructure, socio-demographic, and economic variables, this project aims to provide valuable insights into the underlying factors contributing to these issues. Furthermore, by examining the association between high-emission locations and increased idling or searching for parking behaviors, we can better assess the potential contribution of these activities to the overall air pollution in the affected communities. This in-depth understanding can inform targeted planning and policy recommendations that can directly address the specific needs of environmental justice communities, ultimately helping to mitigate the adverse health impacts of truck emissions on their residents.

Therefore, this project aims to address these issues by identifying factors contributing to truck idling and searching for parking, with a particular emphasis on disadvantaged communities in Kern County, California. The project intends to provide general targeted planning and policy recommendations for local and state agencies to reduce the negative impacts of truck idling and parking search on these vulnerable communities. By employing a comprehensive analysis of data from various sources, such as GeoTAB's Intelligence Data tool and the San Joaquin Valley Air Pollution Control District's Community Air Monitoring portal, the research team identifies truck parking and idling hotspots, and estimates emissions from these activities. Additionally, the study estimates econometric exploratory models to understand the underlying factors that may explain idling and searching for parking behaviors. The expected impact of this research is to provide actionable recommendations that can help reduce air pollution from truck idling and searching for parking in disadvantaged communities. The process developed in this project has the potential to be replicated in other communities confronting similar challenges.

The structure of this report is as follows: First, a literature review provides an overview of existing research in the field. Next, the data description and methodology are explained, followed by a detailed presentation of the spatial modeling and analyses conducted. The results and findings are then presented. The report ends with discussion and policy recommendations, and the conclusions.

Literature Review

Environmental and Safety Impacts of Truck Idling and Parking

Truck idling and parking is an important issue widely studied in the literature. The environmental and safety impacts of truck idling and parking are a key concern, as these activities can significantly affect air quality, public health, and the environment.

One of the primary impacts of truck idling and parking is air quality. Studies have found that truck idling is a significant source of air pollution, emitting various pollutants, including particulate matter, NOx, and greenhouse gases (Divekar et al., 2023; Jaller et al., 2013; Thaker and Gokhale, 2016). These pollutants can seriously affect individuals exposed to them, particularly those with pre-existing respiratory conditions. Trucks' engines usually idle to provide power for onboard climate control and accessories, accommodating the federally mandated truck driver rest periods (Lutsey et al., 2004). Truck idling can be an issue for rural and urban areas.

Truck parking mostly generates negative externalities in urban areas due to the limited parking supply of facilities and regulations. Commercial trucks frequently search for parking and even park illegally (Kawamura et al., 2014; Kim and Wang, 2021, p. 202) when parking is unavailable or inappropriately located. In an aggregate view, the two truck activities of idling and searching for parking can contribute to environmental impacts that generate unexpected costs for various stakeholders.

In addition to damaging air quality, truck idling and parking can have high fuel consumption, which also negatively impacts the environment due to the imperfect working condition of engines. When the engine runs in idle conditions, it takes a rich mixture of air and fuel, so the fuel consumption rate is high (Rahman et al., 2013). Studies have estimated that the fuel consumed by trucks during 5 miles of driving is equivalent to just 10 min idling, and 10 min of idling per day will consume more than 27 gallons of fuel per year. Also, fuel consumption from idling is estimated at more than 2 billion gallons per year (Gaines et al., 2006). The restriction of capacity within the trucking industry and extra fuel consumption lead to increased transportation costs for shippers, which are likely to be passed on to consumers through higher prices for consumer goods (Dhariwal et al., 2022). Furthermore, truck parking can take up valuable land, negatively impacting local communities. For example, a study found that proximity to truck parking facilities increases commercial and industrial land values in Little Rock, Arkansas. For a commercial parcel valued at the mean land value in the city, a 1% increase in distance to parking corresponded to a $2,465/acre reduction in land value (Mahmud et al., 2022).

Regulations aimed at reducing idling can improve the energy efficiency of heavy-duty vehicles, which can lead to cost savings for trucking companies and reduced emissions. A recent study found that replacing gasoline fuels with natural gas can improve energy efficiency and reduce CO_2 emissions from medium- and heavy-duty spark-ignition engines. Natural gas showed reduced pumping loss, improved idle stability, and greater resistance to knock, which helped to increase fuel efficiency and reduce CO_2 emissions by 22.5% (Divekar et al., 2023).

Truck idling and parking can also have an impact on noise pollution. The constant idling of engines can generate excessive noise, which can be disruptive to communities and individuals living or working near truck parking facilities or busy trucking routes. Studies have shown that exposure to excessive noise can cause various adverse health effects, including sleep disturbances, cardiovascular disease, and cognitive impairment (Passchier and Passchier, 2000). Also, a study points out that noise exposure from heavy truck traffic disproportionately affects communities of color and can adversely affect sleep, health, and cognitive performance. The study evaluates conventional and supplemental noise metrics in a community noise survey in Southwest Detroit and shows the benefits of using these metrics and community engagement in noise assessment (Peng et al., 2019). Another study found that the increasing number of vehicles in the Dhanbad Township area in India has led to a worsening of the sound environment and rapid escalation of local noise levels. The noise levels obtained at different locations in the city exceed the limits prescribed by the Central Pollution Control Board's (CPCB's) Noise Pollution (Regulation and Control) Rules, 2000, and, even in designated "silence zones," the noise levels exceeded the permissible norms of 50 dB (Debnath and Singh, 2018). Therefore, communities, governments, and the freight transportation industry must work together to address the problem of noise pollution from truck idling and parking.

Several studies have examined the safety impacts of truck idling and parking. A study measuring air pollutant emissions from idling heavy-duty diesel trucks found significant increases in PM and NOx (Rahman et al., 2013). Other studies found that parked trucks can obstruct visibility and create hazards for pedestrians and motorists by blocking sidewalks and pedestrian walkways (Boggs et al., 2019; Conway et al., 2016; Nevland et al., 2020; Wali et al., 2018). A study evaluated the impact of commercial vehicle parking shortage on crash frequency involving parked commercial vehicles along ramps in Tennessee, United States. The study found that higher frequency crashes occurred on entrance ramps that were shorter in length and often adjacent to public parking facilities, with one-third of the collisions occurring on ramps with a parking facility utilization rate of 90% or higher. The results showed that several ramp characteristics, including the presence of a parking facility, ramps with illegally parked vehicles, diamond-shaped ramp configurations, larger ramp shoulder widths, and exit ramps, were significantly associated with an increase in the occurrence of an illegally parked commercial vehicle-involved crash (Boggs et al., 2019). These safety impacts highlight the importance of regulations that address truck idling and parking, both to protect public health and safety and to maintain a healthy and livable urban environment. In California, the state government has implemented various regulations to address these safety impacts, including restrictions on truck idling and requirements for truck parking facilities to be located in designated areas away from residential and commercial buildings (see the next section). These regulations serve as a model for other states and localities looking to address the safety impacts of truck idling and parking.

In summary, studies have found that truck idling and parking can negatively impact air quality, public health, and the environment, as well as traffic safety. There are various regulations and best practices aimed at reducing truck idling and parking, which can help mitigate these impacts.

Existing Regulations for Truck Idling and Parking

State and local regulations regarding truck idling and parking have been established to mitigate the impacts of excessive noise and air pollution from commercial vehicles. At the state level, California has enacted regulations limiting how much time a truck can idle. The California Air Resources Board (CARB) has established idling restrictions for trucks that operate within the state, which include a limit of 5 minutes for diesel-fueled trucks and 3 minutes for other types of trucks (California Air Resources Board, 2004). This regulation applies to truck idling at rest areas, truck stops, and other locations where trucks are parked. At the local level, various cities and counties in California have established their own regulations to further limit truck idling. For example, Los Angeles's city ordinance prohibits commercial vehicles from idling for more than 5 minutes (Anti-Idling Laws Gain Momentum Around the Country). San Francisco also has similar regulations in place ("Vehicle Idle Reduction," 2011).

Regarding truck parking, in California, several state and local regulations are in place. The California Vehicle Code (CVC) section 22500 establishes the requirement for trucks to park in designated areas (California Code). Furthermore, the California Code of Regulations, Title 13, Section 2462 requires that truck stops provide adequate parking for commercial vehicles (California Vehicle Code Section 2462). Additionally, some local jurisdictions in California have further restrictions on truck parking. For example, Los Angeles has regulations regarding truck parking, such as requiring truck drivers to obtain a permit to park their vehicles on the street (SEC. 80.69.2).

As for the examination and assessment of policies, many studies have tried to examine current regulations and best practices. Alho et al. propose an agent-based simulation approach for studying the challenges of managing freight vehicle parking demand at large urban traffic generators, such as shopping malls and commercial buildings with offices and retail. Through a case study in Singapore, the study finds that demand management strategies, such as changes to parking capacity, availability, and pricing, as well as services and technology-based solutions, can improve travel time, parking costs, emission levels, and reduce queuing (Alho et al., 2022). Another study develops an overnight parking choice model for heavy commercial vehicles and integrates it with SimMobility, an agent-based urban simulation platform, to evaluate the potential of this tool for policy evaluation. Using simulations applied to a case study in Singapore, it compares two parking supply scenarios in terms of vehicle kilometers traveled due to changes in the first and last trips of vehicle tours and resulting impacts in traffic flows (Gopalakrishnan et al., 2020). James finds that a partial switch to CNG or propane fuel source is environmentally and economically feasible for the University of South Carolina shuttle fleet if federal or state funding assistance is available and suggests ways to reduce idling practices to reduce the environmental impacts (James, 2020). Also, a study found that the California emissions reduction plan had a more significant impact on reducing nitrogen dioxide (NO_2) concentrations in areas within 500 m of major highways that serve as truck routes (goods movement corridors) compared to areas farther away or adjacent to routes that prohibit trucks. The study also found that the subjects living in goods movement corridors experienced statistically significantly larger reductions in NO_2 exposure than those living in faraway areas (Su et al., 2020). In short, studies have shown that various strategies and solutions, including demand management, changes to parking capacity and pricing, technology-based solutions, and alternative fuel sources can effectively improve travel time, parking costs, emission levels, and reduce queuing. Studies have also demonstrated that regulatory and policy interventions, such as California's emissions reduction plan, significantly reduce air pollution concentrations in areas near major truck routes. These findings can inform the development of policies and regulations to improve air quality and reduce the negative impacts of truck idling and parking.

Overall, California's state and local regulations aim to mitigate the impacts of truck idling and parking on noise and air pollution. These regulations provide a framework for reducing excessive exposure to noise from truck engines and for promoting designated truck parking areas to reduce air pollution. The regulations also play a crucial role in protecting California residents' health and quality of life.

Idling Reduction Technologies (IRTs) and Mitigating Strategies

Idling Reduction Technologies (IRTs) are systems and devices that reduce the time a commercial vehicle's engine is left running when the vehicle is not in use (Lust et al., 2009). The primary purpose of IRTs is to save fuel, reduce emissions, and improve operational efficiency (Frey et al., 2009). Some examples of IRTs for trucks include:

- Auxiliary power units (APUs): APUs provide heating and cooling for the truck's cab without requiring the main engine to run. They typically run on diesel fuel and use less fuel compared to idling the main engine (Samsun et al., 2016).
- Battery-powered HVAC systems: These systems provide heating and cooling for the truck's cab using battery power without requiring the engine to run. They can be charged using the vehicle's alternator or a shore power system, which provides electrical power from an external source (MacDonald et al., 2013).
- Shore power systems: These systems allow a truck to plug into an external electrical source, providing power for heating and cooling, lighting, and other electrical needs. This eliminates the need to idle the engine to provide power (Karimi et al., 2020).
- Engine control systems: These systems monitor and control engine idling by automatically shutting down the engine when it is not in use and restarting it when needed (Lin et al., 2003; Truck Electronic Engine Controls). This helps to reduce fuel consumption and emissions from commercial vehicles.

By reducing idling time and fuel consumption, IRTs play an important role in promoting sustainability and reducing the environmental impact of the transportation industry. These are just a few examples of IRTs used in the commercial vehicle industry. Also, several strategies can be used to mitigate idling in commercial vehicles and encourage the adoption of IRTs. These strategies include:

- Promoting the use of anti-idling policies by trucking companies and governments. A study aimed to assist Southwest Detroit Environmental Vision (SDEV) in addressing diesel truck emissions and idling in Southwest Detroit. The study used a community-based participatory research approach, collecting data from trucking/logistics companies, policy stakeholders, and community members to identify opportunities to address the issue (Ziegler, 2015).
- Providing incentives for the purchase of IRTs, such as subsidies and tax credits. Skerlos and Winebrake evaluate the effectiveness of tax credits to encourage consumers to purchase plug-in hybrid electric vehicles (PHEVs) in the U.S. The authors argue that the current credits do not account for variability in social benefits and consumer income and suggest offering the credits for increased social benefits (Skerlos and Winebrake, 2010).
- Implementing anti-idling laws that prohibit the idling of heavy-duty diesel engines for extended periods. Hasani et al. focus on anti-idling law as a tool for reducing emissions from urban freight. The challenge lies in the mismatch between theoretical research and practical implementation, which is due to the lack of collaboration and resources from city governments and the industry, as well as the fragmented nature of the urban freight sector with different companies pursuing different decarbonization strategies (Hasani et al., 2019).

In addition, technology-based solutions such as telematics and engine control systems can monitor and control idling, helping to reduce fuel consumption and emissions. Combining these strategies can help create a more sustainable and environmentally friendly transportation industry. A study evaluated the potential of using GPS and portable activity measurement systems (PAMS) data to analyze air quality in the Houston-Galveston Area Council's Drayage Loan Program (Farzaneh et al., 2020). The results showed that 91% of truck activities occur during the day, with low average speeds mainly due to extensive idling. The integrated data analytics system can help identify, implement, and measure the performance of idling reduction strategies, such as using a battery-electric air conditioning system.

Data

Data Sources and Collection Methods

The data for this project has been obtained from various sources to ensure a comprehensive understanding of truck idling, searching for parking, and their impacts on AB 617 disadvantaged communities. The following data sources have been used in this project:

- Geotab data: Geotab's Intelligence Data tool provides geohash level aggregated data. We used data from a 6-month period from October 2020 to April 2021. The dataset includes information on truck idling, searching for parking (all vehicle types), and truck parking locations. Geotab is a telematics company that provides insights into vehicle activity and behavior. The Geotab datasets used in our project include:
    - Areas of idling (IDL): This dataset contains information about the duration and location of engine idling events for different vehicle types, classes, and fuel types. This information is crucial in understanding the idling patterns of trucks and helps us analyze the environmental impact of excessive idling. The IDL dataset covers the Bakersfield, Shafter, Arvin, and Lamont areas and consists of 9,368 observations.
    - Searching for parking (SFP): The SFP dataset provides insights into the parking search behaviors of light to heavy-duty vehicles. By analyzing this dataset, we can better understand the factors that contribute to parking search times and develop strategies to alleviate parking-related congestion. The SFP dataset does not cover the Shafter, Arvin, and Lamont areas but provides 332 observations in Bakersfield for analysis.
    - Truck parking locations (TPL): This dataset contains information about the parking locations of Class 7 and 8 trucks. The TPL dataset helps us assess the parking demand for heavy-duty trucks and identify areas with high or low parking availability. The dataset covers the Shafter, Bakersfield, Arvin, and Lamont areas and includes 132 observations.
- 2021 CalEnviroScreen (CES) 4.0 data: The CES is a comprehensive dataset that identifies communities in California disproportionately burdened by multiple sources of pollution. It includes various indicators such as air quality, water quality, and socioeconomic factors.
- 2018 ZIP Business Pattern data: This dataset provides the number of establishments, employment, and annual payroll for businesses in each ZIP code area for Kern County, CA. It is organized by industry according to the 2017 North American Industry Classification System (NAICS) codes. This dataset helps identify potential sources of truck activity and parking demand in the study area.
- 2018 Five-Year American Community Survey Census Tract data: This dataset contains demographic, social, economic, and housing data for census tracts in Kern County. This information offers insights

into the study area's population characteristics and living conditions, allowing for a deeper understanding of the communities affected by truck idling and parking issues.

- OpenStreetMap (OSM) road network, land use, and point-of-interest (POI) data: OSM data provides information on road networks, land use patterns, and POIs within the study area. Land use data from OSM is valuable in this context, as it can help identify potential hotspots for truck activity and areas with limited parking availability. OSM data is widely used in transportation research and planning, and its accuracy has been validated in various studies, making it a reliable source of information for this project.
- Kern County traffic count data system (TCDS): This data system includes the 2021 historical traffic count data (specifically Annual Average Daily Traffic and the percentage of this that consists of trucks, K%).

Data gathering methods include downloading data directly from the Geotab Ignition data platform, accessing Census API, and downloading OSM network and land use data using the ‘osmnx’ package. These methods ensure that the most up-to-date and accurate information is available for the analysis in this project. By combining data from these various sources, the research team can thoroughly investigate the factors contributing to truck idling and searching for parking in the study area and assess the impacts of these issues on disadvantaged communities.

Overview of Variable Groups Employed in the Analysis

This subsection presents the socio-demographic and economic variables used in our project analysis. These variables are crucial to understanding the complex interplay between truck idling, parking search behavior, and their impacts on disadvantaged communities. Table 1 below summarizes the different variable groups used in the study, which include dependent variables, independent variables from CES data, census data, and spatial variables derived from OSM data.

Table 1. Summary of variable groups

Category	Variable Group	Example Variables
Dependent Variables		
Geotab Data	Y1: Average idling time	N/A
	Y2: Average time to park	
	Y3: Inferred truck parking demand (AADT_truck)	
Independent Variables		
CalEnviroScreen (CES) Data	Exposure Indicators	Ozone concentrations, PM2.5
	Environmental Effects Indicators	Groundwater threats, Hazardous waste
	Sensitive Population Indicators	Asthma emergency visits, Low birth weight
	Socioeconomic Factors	Poverty, Unemployment
Census Data	Total employment	
	Employment across Industry sectors North American Industry Classification System (NAICS) codes	Agriculture, Construction
	Other census variables	Population, Educational attainment
Spatial Variables (OSM)	Distance to major road, point-of-interest, and specific land use area within an 800-meter radius	N/A
	Amenities, shops, tourism, and leisure facilities within a 400-meter radius	N/A
	Road density, nearest road distance, nearest traffic facility distance, nearest public service distance	N/A

Study Area

The study area selected for this project comprises the cities of Bakersfield, Shafter, Arvin, and Lamont in Kern County, California. We selected this area, in this study of activities related to truck emissions, for the following reasons:

1. These communities have been recognized as disadvantaged and overburdened by emissions from various sources, including industrial operations, agricultural activities, oil production, dust, and heavy-duty truck traffic (California Air Resources Board, 2021).

2. The California Air Resources Board (CARB) approved the Community Emissions Reduction Program (CERP) for the Arvin/Lamont community in response to AB 617, which mandates community-focused action to reduce air pollution and improve public health in communities that experience disproportionate burdens from exposure to air pollutants (California Air Resources Board, 2019). Figure 1 shows a map of AB 617 areas in the state.

3. The area is near major transportation corridors (Figure 2), highlighting the need to understand and address the local impact of trucks;

4. Many neighborhoods within these communities are near farmland, resulting in increased pesticide exposure, which has become a top priority for residents.

Figure 1. AB 617 communities and freight corridors in California, with Kern County shaded to indicate the county containing the communities studied in this report.

Our selection of Bakersfield, Shafter, Arvin, and Lamont communities is based on the urgency to address the unique combination of environmental challenges stemming from industrial operations, agricultural activities, and truck traffic. The prevalence of these issues, compounded by the communities' proximity to major transportation corridors (Figure 2) and farmlands, not only intensifies the local air pollution burden but also provides a comprehensive case study that mirrors similar environmental justice and public health challenges. The insights gained here will not only be instrumental in formulating specific, actionable policies for these localities but will also offer valuable precedents for addressing analogous challenges in other AB 617 communities and pollution-burdened areas worldwide. The chosen communities present a distinctive intersection of environmental stressors, offering an enriched context for in-depth analysis, thereby yielding findings that are both locally impactful and globally relevant.

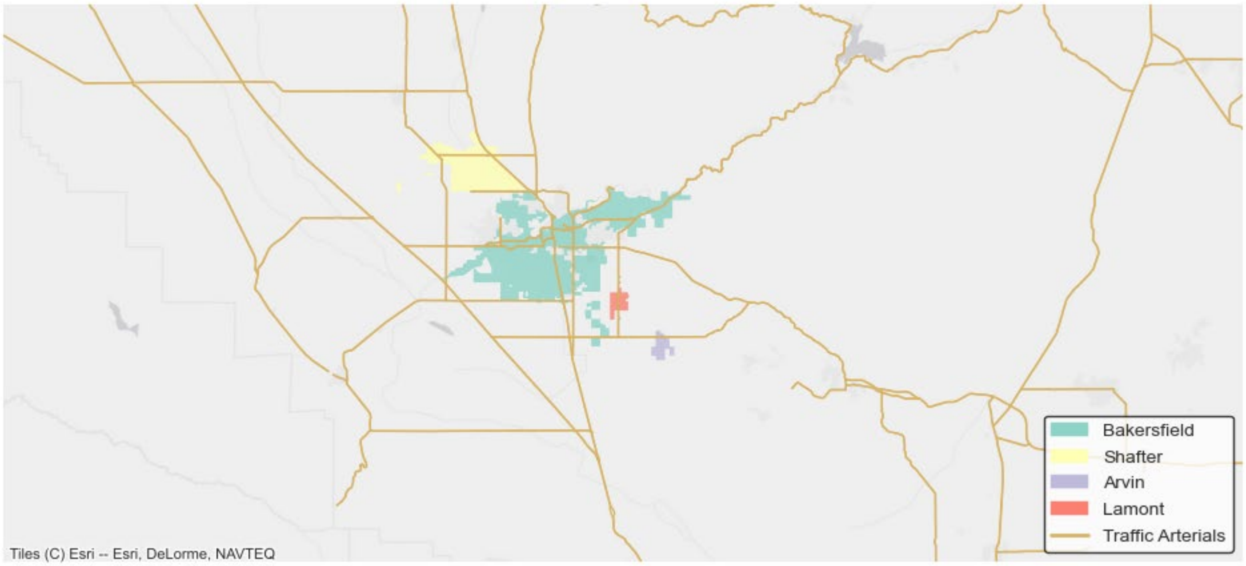


Figure 2. Map of the study area with four cities and major transportation corridors (traffic arterials) highlighted.

Figure 3 shows six key socioeconomic and environmental variables in the study area: total population, poverty rate, traffic density, population with limited education, pollution burden score, and unemployment rates. As shown, Bakersfield, Shafter, Arvin, and Lamont collectively house over 50% of the total residents in Kern County, a demographic scale that allows for a comprehensive analysis of pollution exposure and socioeconomic factors. These communities also face significant challenges related to socio-economic factors and environmental issues. The maps reveal that these cities have high poverty rates and unemployment levels, indicating a large proportion of disadvantaged residents. Furthermore, the maps show that these communities have low education levels, which could exacerbate the challenges related to socio-economic well-being and access to resources. In terms of environmental factors, Figure 3 displays a high pollution burden score in the four cities, emphasizing the severity of pollution exposure in these communities. Moreover, the traffic map indicates that these cities experience a considerable amount of traffic, which could contribute to air pollution and related health issues. The combination of these factors demonstrates that Bakersfield, Shafter, Arvin, and

Lamont are indeed disadvantaged communities, facing considerable challenges, which underscores the importance of focusing on these communites.

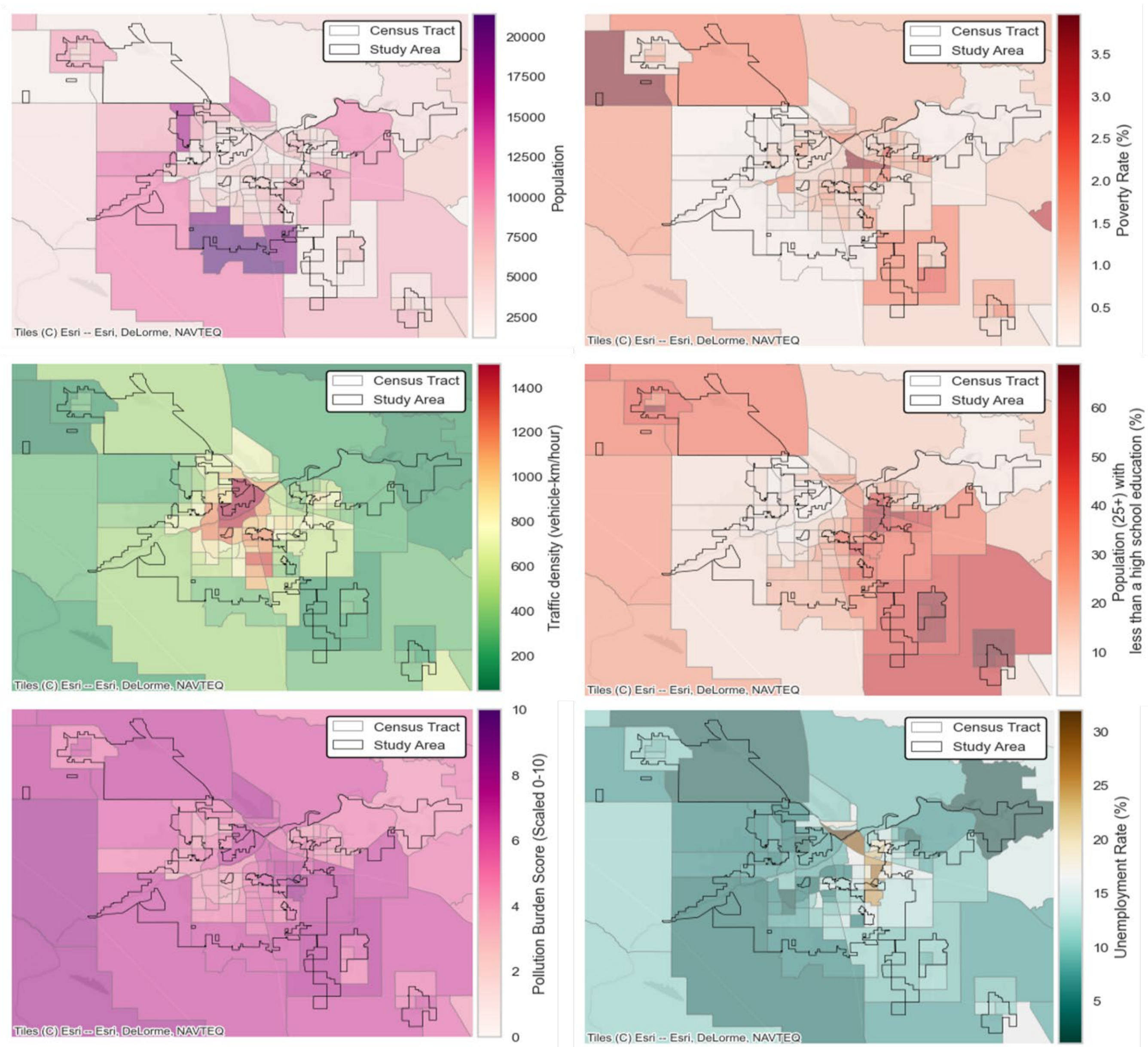


Figure 3. Distribution of sample socio-economic and environmental variables

Methods

This project employs comprehensive methods to analyze truck idling, searching for parking, and truck parking demand. Figure 4 shows the technical workflow of the analyses, from data preparation and preprocessing to results analysis and policy recommendations. The process starts with data preparation and preprocessing, followed by three parallel analysis branches, each focusing on a specific issue: idling, searching for parking, and truck parking demand. In each analysis, the random forests (RF) model, evaluates the model performance using various metrics and analyzes feature importance and model diagnostics. For the searching for parking analysis, the process also uses a fishnet grid of the study area to predict the average time to park. In the case of the truck parking demand analysis, we conduct a sensitivity analysis with varying truck percentages. After these parallel analyses, the process conducts a correlation analysis of emission and environmental factors using a Spatial Error Model (SEM). The workflow culminates in results analysis and discussion, followed by policy recommendations.

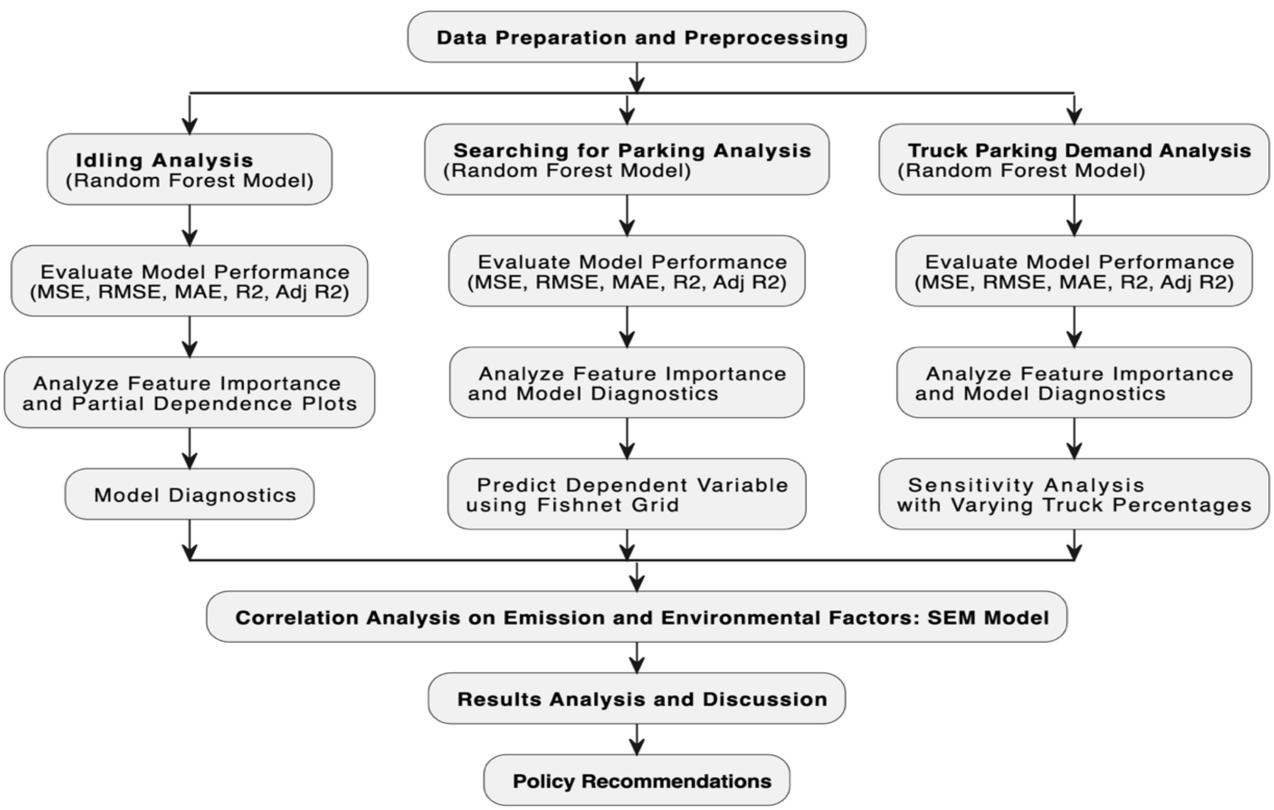


Figure 4. Technical workflow for modeling analyses (MSE, mean squared error; RMSE, root mean squared error; MAE, mean absolute error; R2, R-squared; Adj R2, adjusted R-squared; SEM, spatial error model)

Data Preparation

This project integrated the various data sources into a unified data framework, including Geotab (containing dependent variables), CES, 2018 Five-Year American Community Survey data, 2018 ZBP data, and OSM road network spatial data. The integration method uses datasets containing dependent variables, such as Geotab and traffic count data, as the parent datasets (spatial points). Other CES and Census data are interpolated from the census tract level to the parent datasets; ZBP data is converted from ZIP to points; and spatial data variables—such as road distance, the number of POIs within a buffer zone, and land use area–are generated (details in subsequent sections) to form the integrated dataset. The team then standardizes the integrated datasets for splitting, training, and modeling.

Mainly, spatial data processing is performed on integrated datasets, which involves handling spatial data, such as road networks, land use, and points-of-interest. The processing of this data includes data projection, buffer analysis, and overlay analysis. For instance, road network data can be projected to a common coordinate system to ensure consistency.

Model Comparison and Selection

In this process of model comparison and selection, the primary objective is to identify suitable models that can both effectively predict the dependent variables—average idling time (Y1), average time to park (Y2), and proximitized truck parking demand (Y3)—and provide interpretability to understand the underlying relationships between the independent variables (e.g., socio-demographic factors, environmental indicators, spatial characteristics) and the dependent variables.

The hypothesis is that the independent variables, such as traffic density, socio-demographic factors, and land use patterns, can help explain the variations in the dependent variables. The study compares and evaluates several popular machine learning models with interpretability and predictability as main concerns. These models have been chosen due to their varying strengths and characteristics, which may cater to different aspects of the analysis. Comparing and selecting the best-fitting model can ensure that results are accurate and reliable, help gain valuable insights into the factors that influence the variables of interest, and test our hypothesis.

Figure 5 illustrates the conceptual framework for our model comparison and selection process. The dependent variables are predicted and explained by the independent variables. We evaluate various models, including neural networks (NN), convolutional neural networks (CNN), decision trees / random forests (DT/RF), K-nearest neighbors (KNN), and Bayesian Ridge (see below for detailed descriptions), to test these hypotheses. The evaluation metrics, such as mean squared error (MSE), root mean squared error (RMSE), mean absolute error (MAE), R-squared (R2), and adjusted R-squared (Adj R2), are used to determine the predictability performance of each model. Finally, the study selects the best model that offers strong predictability and interpretability.

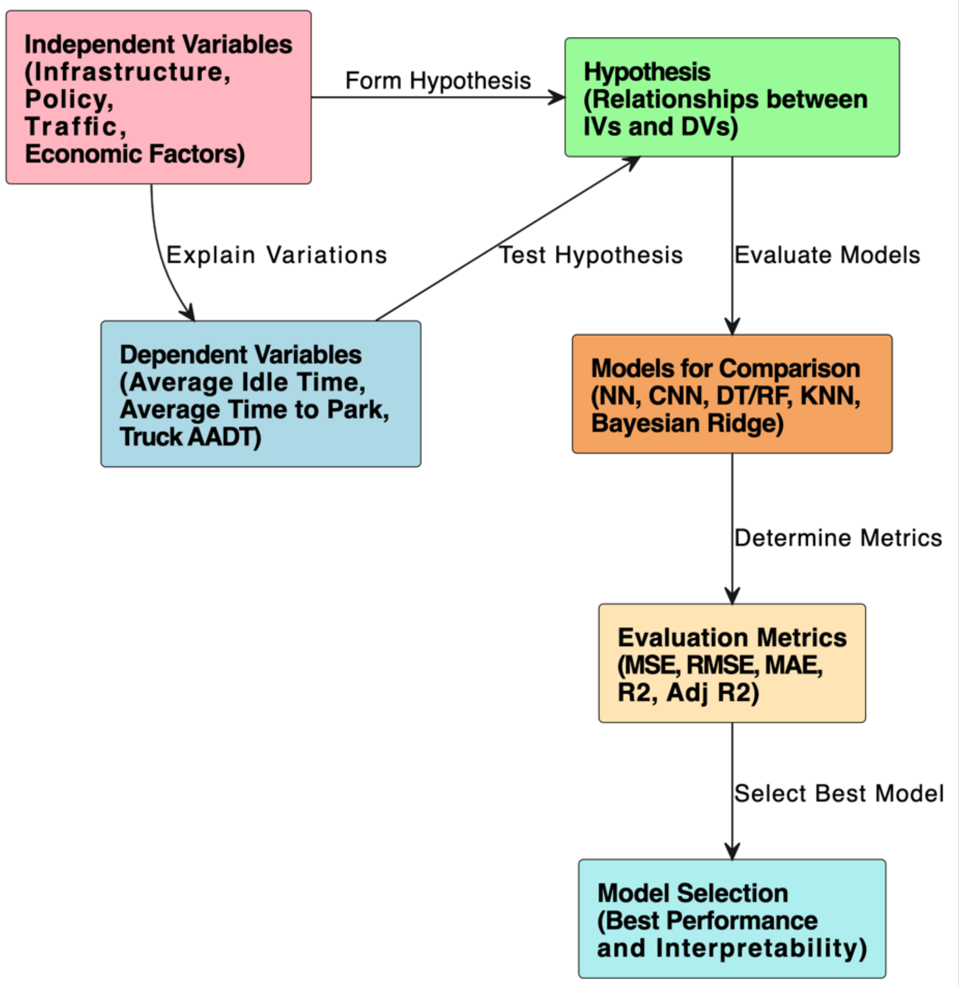


Figure 5. Conceptual framework for model comparison and selection process (AADT, annual average daily traffic; IV, independent variable; DV, dependent variable; NN, neural network; CNN, convolutional neural network; DT/RF, decision trees/random forests; KNN, K-nearest neighbor; MSE, mean squared error; RMSE, root mean squared error; MAE, mean absolute error; R2, R-squared; Adj R2, adjusted R-squared)

- Neural networks (NNs): Neural networks are a class of machine learning models that mimic the structure and function of the human brain. They consist of interconnected nodes or neurons, arranged in layers (Abdi et al., 1999). Neural networks excel at modeling complex data and nonlinear relationships, which could help capture the intricate relationships between various socio-demographic, economic, and spatial factors in our study.
- Convolutional neural networks (CNNs): Convolutional neural networks are a specialized type of neural networks designed to process grid-like data, such as images and time-series data (Gu et al., 2018; O'Shea and Nash, 2015). While the study does not primarily involve image or sequence data, CNNs can still be considered for their ability to capture local spatial patterns, which might be relevant in identifying spatially dependent factors influencing idling, parking search, and truck parking demand.
- Decision trees / random forests (DT/RF): Decision trees are hierarchical models that recursively split the input data based on feature values to make predictions (Myles et al., 2004). They could offer high interpretability, crucial for understanding the underlying relationships between the independent and dependent variables and informing policy recommendations. Random forests are an ensemble method that combines multiple decision trees to improve predictive performance and reduce the risk of overfitting (Rodriguez-Galiano et al., 2015; Segal, 2004).
- K-nearest neighbors (KNN): K-nearest neighbors is a simple, non-parametric algorithm that makes predictions based on the majority class or average value of the K-nearest data points in the feature space (Song et al., 2018). KNN works well with small datasets and relatively simple problems (Zhang et al., 2018). Its ability to capture local similarities among data points could potentially reveal spatial patterns or clusters of idling, parking search, and truck parking demand in our study.
- Bayesian Ridge: Bayesian Ridge regression is an extension of linear regression that incorporates prior knowledge about the coefficients through a probabilistic framework (Hackeling, 2017). It is particularly useful for regression problems with multicollinearity among the features, which might be present in our study given the diverse set of socio-demographic, economic, and spatial factors. Bayesian ridge can estimate the uncertainty of the coefficients and reduce overfitting by shrinking the coefficients towards zero, resulting in more robust predictions (Efendi and Effrihan, 2017), which can be suitable for datasets with less than 400 observations in the project.

Model Diagnostics and Performance Assessment

Model Refinement and Optimization

This project uses two techniques to refine and optimize the models: cross-validation (CV) and grid search–based CV (GridSearchCV). Fine-tuning the models involves adjusting their parameters and hyperparameters to achieve the best possible prediction accuracy and interpretability. This process is vital for the project because it enables identifying the most influential factors affecting idling time, time to search for parking, and truck parking demand and drawing reliable conclusions from the data. The explanations of CV and GridSearchCV are as follows.

- Cross-validation is a resampling technique that helps mitigate overfitting by assessing the model's performance on different subsets of the training data, thereby obtaining reliable performance estimates for each model (Browne, 2000). We used K-fold cross-validation, which involves dividing the dataset into K equally sized folds. The model is then trained on $(K-1)$ folds, with this process repeated K times. The average performance across all K iterations provides an unbiased estimation of the model's performance. We chose K-fold cross-validation because it reduces the risk of overfitting and provides a more accurate representation of the model's performance on unseen data (King et al., 2021).
- GridSearchCV is a technique that automates the process of hyperparameter tuning by exhaustively searching through a specified parameter grid, aiming to find the optimal combination of hyperparameters that yields the best model performance (Pedregosa et al., 2011). In the project, we implemented GridSearchCV for each model, including the selected hyperparameters and search space for each model. For example, in the case of the random forest model, we search through various combinations of parameters such as the number of trees, tree depth, and minimum samples per leaf. By leveraging GridSearchCV, we were able to identify the best hyperparameter settings for each model, ultimately resulting in improved model performance and increased predictive accuracy.

Model Performance Assessment

To evaluate the performance of the models in our study, we employ various metrics, including mean squared error (MSE), root mean squared error (RMSE), mean absolute error (MAE), R-squared (R2), and adjusted R-squared (Adj R2). In addition, the study employs several diagnostic tools and techniques to assess model performance, including residual analysis and feature importance analysis. The rationale behind using these tools and techniques is to comprehensively understand each model's performance and identify areas for potential improvement.

The approach to model diagnostics and performance assessment in the context of analyzing truck idling, searching for parking, and truck parking demand emphasizes the importance of selecting and refining models that offer not only high predictive accuracy but also provide meaningful insights into the relationships between the independent variables, such as socio-economic factors, environmental aspects, and spatial characteristics, and the dependent variables, which represent various aspects of truck activities.

Truck Activities and Emission Modeling Analyses

Idling Analysis

The aim of the truck idling analysis is to better understand truck idling behavior and its implications for fuel consumption, emissions, and local air quality. The dependent variable for this analysis is the average idle time (Y1), which represents the mean duration trucks spend idling at specific locations. This variable is significant as it helps us assess idling patterns and identify areas with high levels of idling, which may indicate insufficient parking facilities, traffic congestion, or other factors contributing to inefficient truck operations. We employ feature importance analysis and partial dependence plots techniques to explore the relationships between the key independent variables and Average Idle Time. Feature importance analysis allows for identifying the most influential variables in predicting idling time, which can help prioritize interventions or policies to address the underlying issues. Partial dependence plots, on the other hand, provide insights into the non-linear relationships between the independent variables and Average Idle Time, offering a deeper understanding of how these variables interact and influence idling behavior.

Searching for Parking Analysis

Searching for parking examines the time spent to park by all vehicles (including trucks and passenger vehicles). The dependent variable for this analysis is the average time to park (Y2), which represents the mean duration trucks spend searching for available parking spaces. This variable is crucial in understanding parking search behavior, as it can help identify areas where parking availability is limited, leading to increased search times and operational inefficiencies. We use feature importance analysis and model prediction techniques to investigate the relationships between the key independent variables and Average Time to Park. Furthermore, we employ the Fishnet Grid Analysis approach to better understand the spatial distribution of parking search times. Dividing the study area into a grid of uniform cells helps visualize the geographical patterns of parking search times and identify hotspots where targeted interventions may be most effective. This spatial analysis can guide the development of tailored solutions, such as new parking facilities or demand management strategies, to alleviate parking search challenges and improve urban freight efficiency.

Truck Parking Demand Analysis

In this part of the analysis, we investigate truck parking demand, a critical component of urban freight management with significant implications for traffic management, land use planning, and environmental impacts. The dependent variable is truck parking demand, truck AADT (Y3), which is an indirect measure of parking demand (AADT = annual average daily traffic). Truck AADT provides valuable insight into the volume of truck traffic in each area's nearby traffic sensors, which can help estimate parking needs and identify areas with potential parking shortages or oversupply. To explore the relationships between the key independent variables and truck AADT, the study employs feature importance analysis and sensitivity analysis with varying truck percentages (K%). Sensitivity analysis with varying truck percentages (K%) allows us to examine how changes in truck traffic volume, reflecting different levels of demand, can influence the relationships between independent variables and truck AADT.

Correlation Analysis on Emission and Environmental Factors

After conducting the first three parts of the modeling analysis, we examined the relationship between emissions, environmental factors, and the results obtained from the previous modeling steps. To achieve this, we employed the following steps:

1. Interpolate the dependent variables (Y1, Y2, Y3) from the idling analysis, searching for parking analysis, and truck parking demand analysis into the fishnet grid cells.

2. Interpolate the environmental, public health, and socioeconomic variables from the CES dataset into the fishnet grid cells, creating a comprehensive dataset containing all relevant variables.

3. Perform spatial error model (SEM) regression analysis on the combined dataset to explore the relationships between the environmental variables (with a focus on PM 2.5 concentration and toxic release to air [i.e., toxicity-weighted concentrations of toxic substances from facility emissions and off-site incineration][1]) and the dependent variables (Y1, Y2, Y3). We will compare the model performance using various metrics, including Akaike information criterion (AIC), log-likelihood (LL), and pseudo R-squared (PR2), and the p-values of Y1-Y3.

4. Develop two sets of SEM models for each environmental variable of interest (PM 2.5 concentration and toxic release to air): a simple model containing only Y1, Y2, and Y3, and a complex model that incorporates Y1, Y2, Y3, and the additional socioeconomic variables.

5. Analyze the results of the selected SEM models (simple and complex) for PM 2.5 concentration and toxic release to air to gain insights into the relationships between emissions, environmental factors, and the outcomes of the previous modeling steps.

[1] The term "toxic release to air" is adopted directly from the CalEnviroScreen 4.0.

Empirical Results

The study examines truck-related behaviors and requirements—referred to as “truck activities”—which include trucks’ active on-road behaviors such as idling and searching for parking, as well as the resultant parking demand. These activities collectively influence traffic dynamics, the use of parking infrastructure, and the environmental quality in the regions under study. The first subsection, Exploratory Spatial Modeling and Analyses, explores the spatial patterns and relationships identified in our data, providing a foundation for subsequent modeling efforts. The second subsection, Truck Activities and Emission Modeling, discusses the results of our selected models of the three truck activities, the performance of the models, and the environmental indicators from CalEnviroScreen associated with the truck activities. Through these analyses, we aim to provide insights into the various factors influencing the truck activities and provide a basis for informed decision-making and policy recommendations.

Exploratory Spatial Modeling and Analyses

Idling Analysis

The exploratory analysis of idling data, delineated in sections (a) to (c) below, demonstrate:

> (i) The prevalence of pronounced idling hotspots, particularly in urban centers like downtown Bakersfield and Shafter, underscoring a conspicuous concentration of idling activities, as depicted in the cumulative daily idling time KDE map (Figure 6).
>
> (ii) Heavy-duty trucks (HDTs) are the principal contributors to cumulative idling times, outpacing cars and lighter trucks, a phenomenon illustrated by the marked disparity in idling times across various vehicle types (Figure 8).
>
> (iii) The idling behaviors of these heavy-duty trucks, while concentrated, exhibit a pronounced variance, indicative of a heterogeneous pattern of idling times even within this specific vehicle category.
>
> (iv) Geospatial patterns and temporal fluctuations in idling behaviors are discernible, with certain locales and times having intensified idling, as evidenced in the daytime idling concentrated around commercial and industrial hubs.

In sum, the idling landscape is characterized by a substantial contribution from heavy-duty trucks, a phenomenon that is geographically concentrated but varies among individual vehicles. These findings spotlight the need for nuanced, targeted interventions, underpinned by a keen understanding of the spatial, temporal, and vehicular dimensions of idling behaviors. This in-depth analysis serves as a foundation for informed policy formulation, aimed at mitigating the environmental and health impacts of prolonged vehicle idling.

a. Cumulative Daily Idling Time

We focus on the analysis of cumulative daily idling time within the study area to better understand the spatial distribution of idling events and identify potential hotspots that may require targeted interventions to reduce idling times. We employed the kernel density estimation (KDE) method, a non-parametric technique widely used to estimate a random variable's probability density function (Węglarczyk, 2018). The KDE analysis on the "cumulative daily idling time" variable for the idling dataset revealed that downtown Bakersfield and Shafter exhibit significant idling hotspots. These areas indicate potential locations for implementing idling reduction measures.

Figure 6 shows the cumulative idling time KDE map, which showcases the spatial distribution of idling events within the study area. The map employs KDE to display the concentration of idling time at different locations, enabling us to identify idling hotspots.

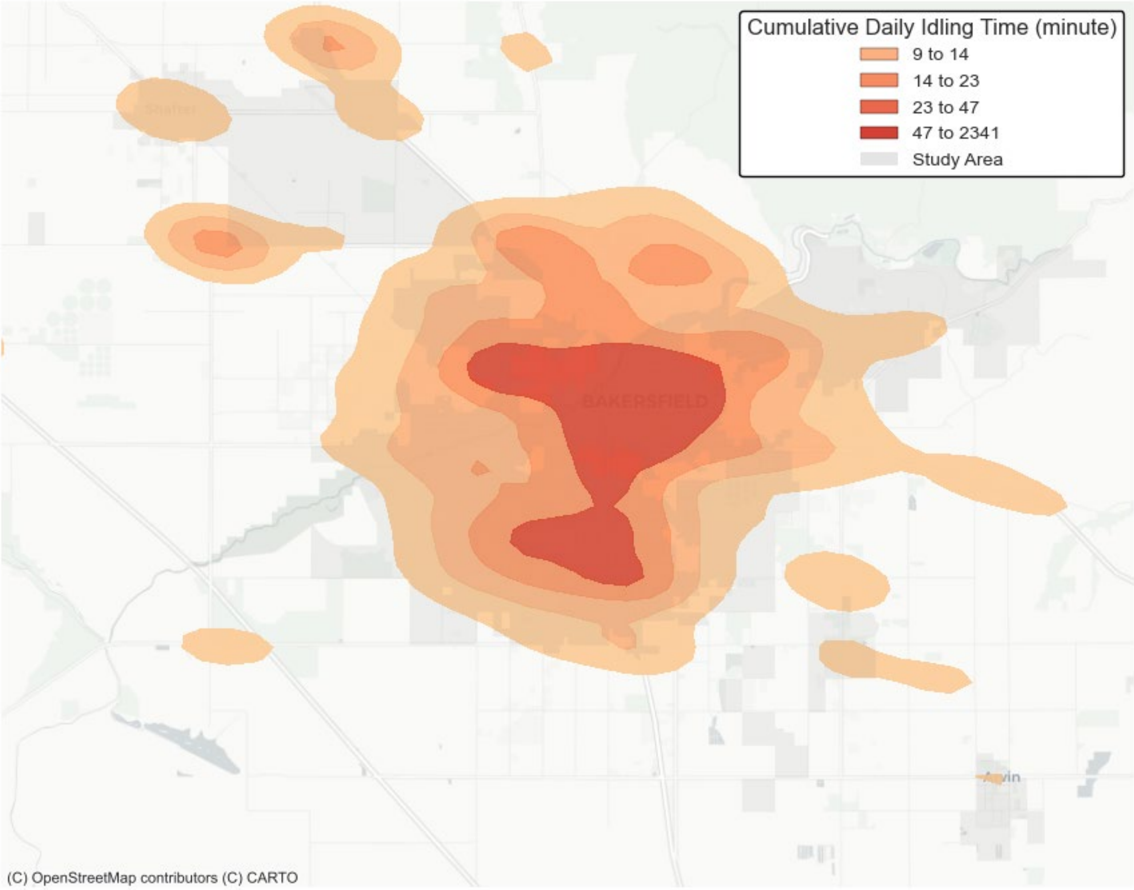


Figure 6. Cumulative daily idling time KDE map

Figure 7 presents a 2D density plot, illustrating the relationship between cumulative daily idling time and average idling time of all idling events for all vehicle types. The graph visualizes the spatial distribution of cumulative daily idling time of all vehicle types across the study area. The density patterns in the graph, reveal a positive correlation between cumulative and average idling times for different vehicle types, as shown by the dark blue dots and marked by the trend line. This relationship suggests that as the cumulative idling time increases, the average idling time also tends to increase, as would be expected. However, many data points have cumulative idling times significantly higher than the trend line—i.e., greater than would be suggested by the hypothetical proportional relationship between cumulative and average idling times. These data points indicate the presence of potential idling hotspots in the study area, as labeled in Figure 7.

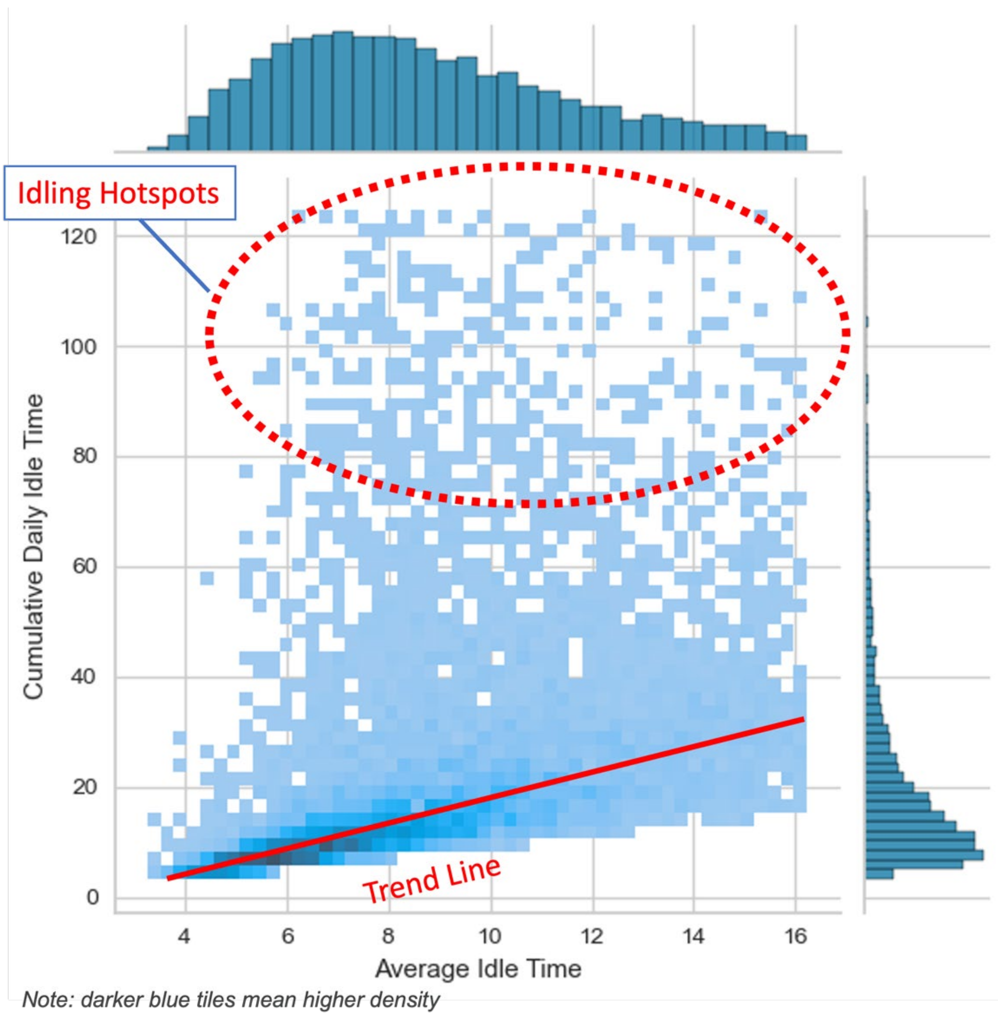


Figure 7. Density plot between cumulative daily idling time and average idling time

Figure 8 displays a series of boxplots, representing the average idling time for various vehicle types, including all types, cars only, light-duty trucks (LDTs), medium-duty trucks (MDTs), and heavy-duty trucks (HDTs). These boxplots provide an overview of each vehicle category's tendency and dispersion of idling times. We can identify which categories have the highest idling times and the greatest variability by comparing the median values and interquartile ranges across vehicle types. Notably, the average idling time for all vehicle types appears to correlate highly with trucks (of all weights) rather than cars, indicating that trucks contribute significantly to idling events.

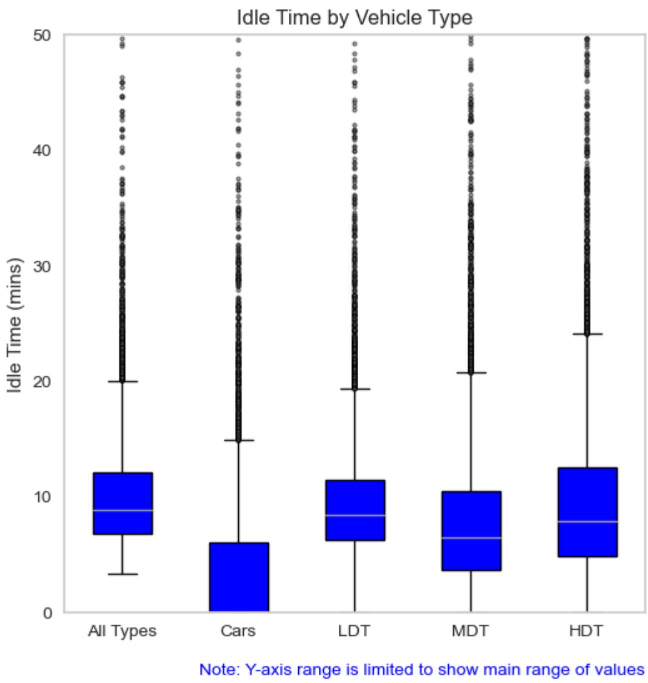


Figure 8. Boxplots of average idling time for various vehicle types (LDT, light-duty trucks; MDT, medium-duty trucks; HDT, heavy-duty trucks) (The box limits indicate the range of the central 50% of the data [interquartile range]; the central lines in the boxes indicate medians; the whiskers [solid lines projecting vertically from the boxes] extend to the furthest data point from the box that is within 1.5 times the height of the box [i.e., 1.5 × the interquartile range); dots beyond the whiskers are considered outliers.)

b. Average Vehicle Idling Time (Weighted by Vehicle Type Percentage)

Given that light-, medium-, and heavy-duty trucks are the main contributors to 'all vehicle' idling time (Figure 8), we wanted to identify the locations where each of these truck types are idling. We determined the average vehicle idling time, weighted by the percentage of each truck type (light-duty trucks [LDTs], medium-duty trucks [MDTs], and heavy-duty trucks [HDTs]), to identify trends in idling behavior for each truck type and explore spatial differences in idling times. Figure 9 displays the average idling times for the different truck types throughout the study area.

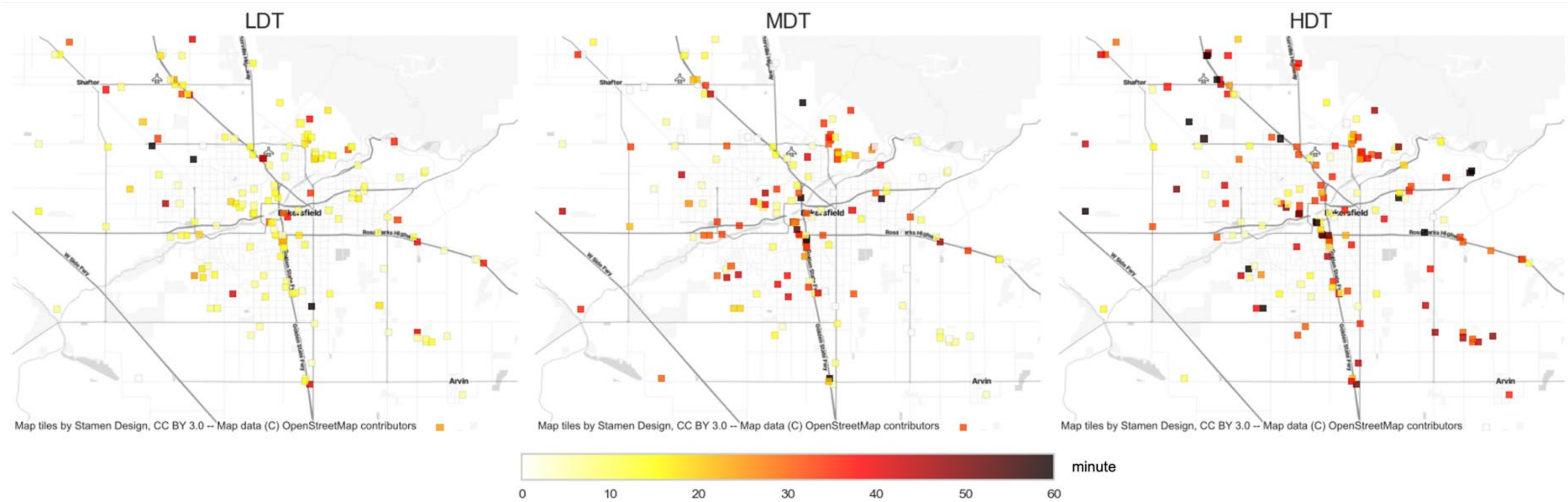


Figure 9. Spatial distribution of average vehicle idling time (light-duty trucks [LDTs], medium-duty trucks [MDTs], heavy-duty trucks [HDTs])

The analyses revealed several key findings:

- HDTs have the highest idling times, followed by MDTs, and then LDTs, as shown by a comparison of the three maps.
- Variability in idling times, indicated by the range of colors of the boxes in each panel, is higher for HDTs and MDTs than for LDTs
- There are geographical differences in idling times, with some geohashes consistently exhibiting higher or lower idling times across all three types of trucks. For instance, certain geohashes in the upper right-hand corner of the plot display higher idling times for all three truck types.

c. Hourly Idling Distributions

We analyzed the hourly distribution of truck idling activities in each geohash (150 m X 150 m square) within the study area to identify distinct time-of-day patterns and their spatial distribution. First, we employed the *k*-means algorithm, an unsupervised machine learning method, to cluster the data into meaningful groups. Then, based on the clustering results, we classified each geohash into one of three categories of truck idling: low idling (throughout the day and night), daytime idling (6 AM–6 PM), and nighttime idling (6 PM–6 AM). Figure 10 aids in determining the optimal number of clusters for our analysis. At $k = 3$ in the figure, there is an elbow in both distortion score and fit time, representing a local optimum. Consequently, we used three clusters for our analysis, capturing the most distinct patterns in idling activities.

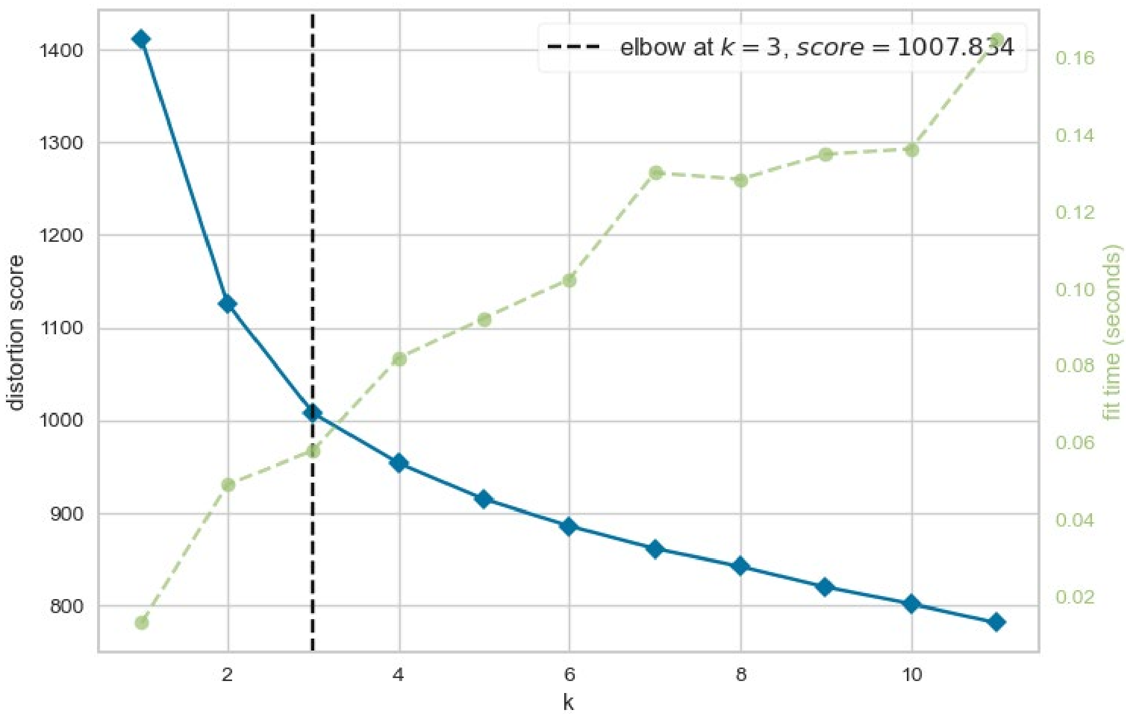


Figure 10. Distortion score elbow for *k*-means clustering

Figure 11 and Figure 12 present the distribution and hourly patterns of the three clusters (categories) of idling: low, daytime, and nighttime. Figure 11 displays the total count of geohashes associated with each idling category, illustrating that the low idling category has the highest count, significantly more than those for daytime active and nighttime active idling geohashes, with nighttime idling being the least frequent. Figure 12, with grouped boxplots, represents the percentage of vehicles idling across different times of the day for each category. This chart illustrates considerable variability in idling behavior within each cluster, across the 24-hour period. The low idling category is consistently represented throughout the day with less variance, while daytime active idling shows higher percentages predominantly during daylight hours, tapering off at night.

Nighttime active idling, conversely, demonstrates increased idling behavior during the evening hours, as evidenced by the higher median values and interquartile ranges during those times. These patterns highlight the clear distinctions between the idling categories, with low idling occurring more uniformly throughout the day, daytime idling peaking during daylight hours, and nighttime idling more prevalent during the late hours.

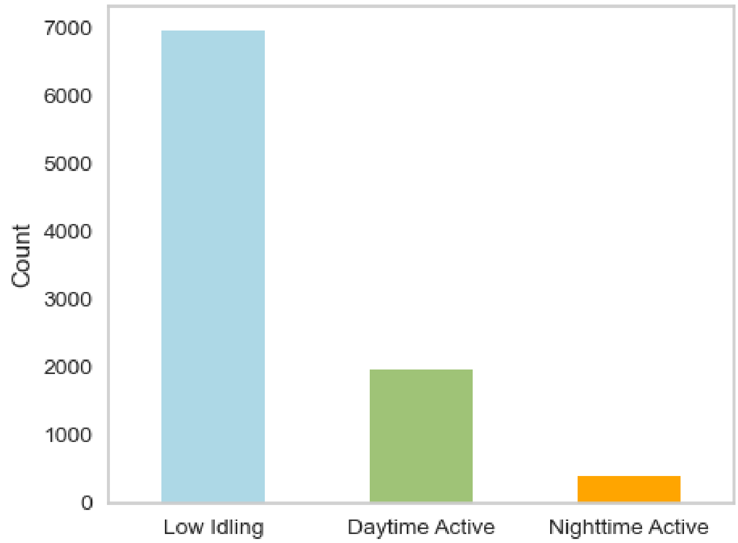


Figure 11. Distribution of each of three categories of geohashes.

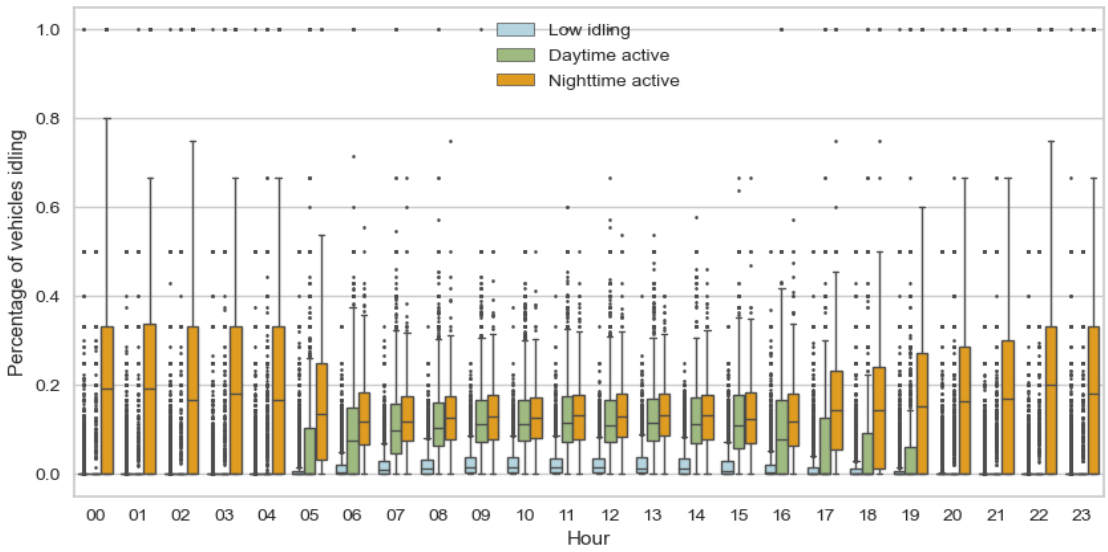


Figure 12. The percentages of total observed vehicles idling at each our of the day for each category of idling: low, daytime, and nighttime. (Each box indicates the central 50% of values [interquartile range]; the central line indicates the median; whiskers extend to 1.5 × interquartile range; the dots represent outliers.)

Figure 13 highlights the geographical patterns of low idling, daytime idling, and nighttime idling activities within the study area. Most daytime idling activities were concentrated in Bakersfield city and its vicinity, and could be associated with idling during truck distribution operations. This suggests that targeted interventions may be necessary to reduce idling during daytime hours in these areas. In contrast, nighttime idling activities were primarily distributed near traffic arteries, freeway entrances and exits, and industrial zones, which may indicate a pattern of trucks idling during rest periods or awaiting early morning deliveries. The concentration of nighttime idling activities in these areas can be attributed to drivers taking mandatory rest breaks and the absence of adequate truck parking facilities, leading to extended idling times. The low idling activities, on the other hand, were predominantly found in residential and less commercially active areas. These patterns may be due to the lower traffic volumes and reduced commercial activities in these locations, especially during nighttime hours.

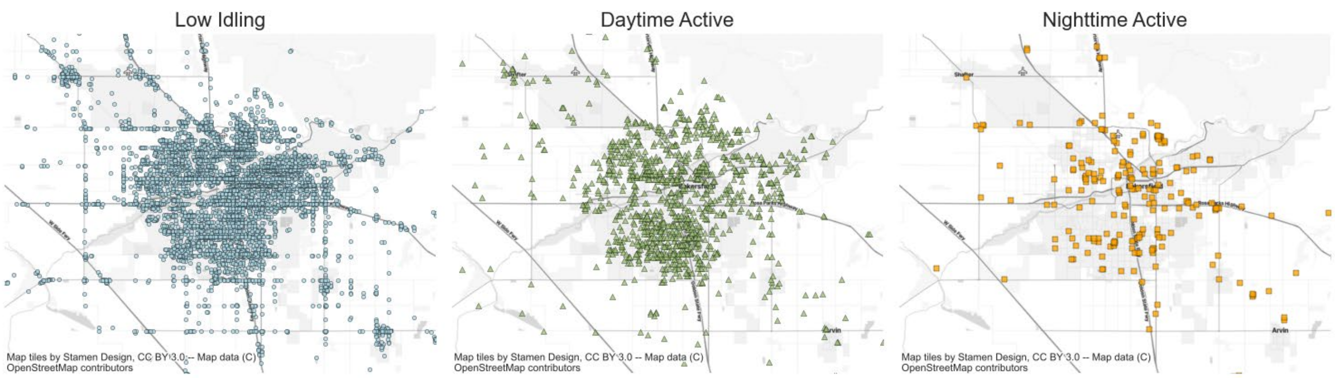


Figure 13. Spatial distribution of the three categories of idling.

Searching for Parking Analysis

In the intricate development of urban mobility, the quest for parking emerges as a pivotal concern. Our analysis, delineated in the subsequent sections, unveils a landscape marked by concentrated hotspots of parking searches and the pronounced impact of truck traffic. Key revelations encompass:

(i) The existence of parking search hotspots, particularly accentuated near arterial routes, freight yards, and highway ramps, unveiling a narrative of intensified search activities in these locales (Figure 14).

(ii) The interplay between the average time to park and the percentage of truck traffic, unveiling specific geohashes marked by heightened truck traffic and extended parking search durations, spotlighting areas ripe for intervention (Figure 15).

(iii) A temporal pattern characterized by a surge in parking searches during daytime, illuminating the daily rhythms of parking behaviors and indicating the window for targeted interventions (Figure 16).

(iv) A scatterplot that unravels the nuances between the total and unique geohashes traversed in the search for parking, spotlighting areas marked by intensified and spatially concentrated parking searches (Figure 17).

a. Parking Search Hotspots

We focus on the spatial distribution of parking search hotspots in Bakersfield, particularly concerning trucks, to better understand the patterns and areas where parking is most sought after. Figure 14 presents two kernal density estimation (KDE) contour maps, illustrating the average cruising time for parking searches in Bakersfield. The top and bottom maps represent, respectively, all geohashes and geohashes with 50% of traffic made up by trucks. The KDE maps show that parking search hotspots are concentrated near traffic arteries, freight yards, and highway ramps, primarily in Bakersfield downtown and its western suburbs. In these hotspots, the average time to park for trucks exceeds 5 minutes, which is relatively long and indicates potential parking challenges for truck drivers. The bottom map, which represents a higher filter selecting truck traffic, further emphasizes the most significant parking search hotspots.

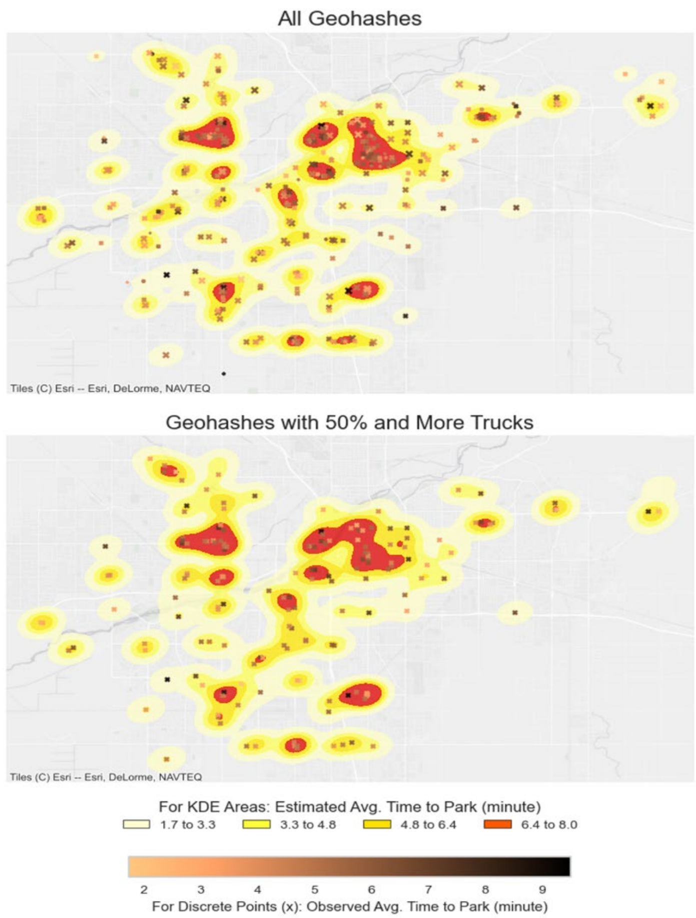


Figure 14. Maps of average time to park showing discrete data of geohashes and kernel-density estimation (KDE) for the colored areas around those discrete data points.

b. Identified Patterns and Clusters of Parking Search

We identified patterns and clusters of parking searches, focusing on the relationship between the average time to park and the percentage of truck traffic. Figure 15 displays a 2D KDE scatter plot and histograms showcasing the relationship between average time to park and the percentage of traffic consisting of trucks across geohashes. We can identify geohashes with a high proportion of truck traffic, which should be the focus areas for targeted interventions.

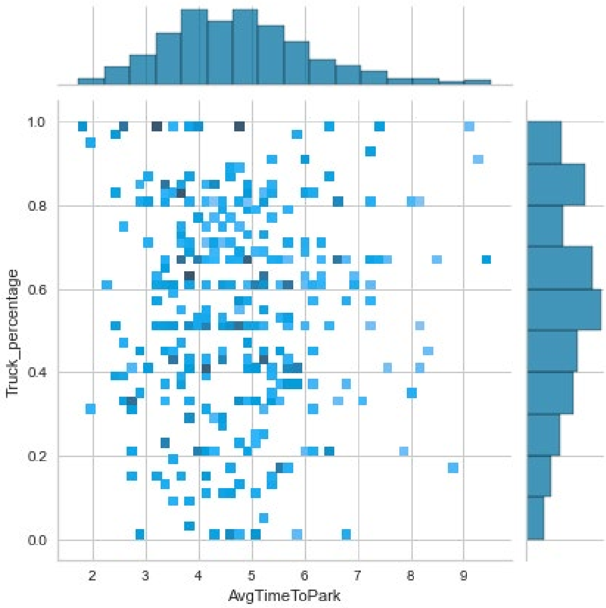


Figure 15. Density plot between average time to park and truck percentage

Figure 16 presents the hourly distributions of parking search activities with boxplots across the 24-hour day. As can be seen, parking searches mostly occur during the daytime, suggesting that targeted measures should prioritize addressing daytime parking search issues.

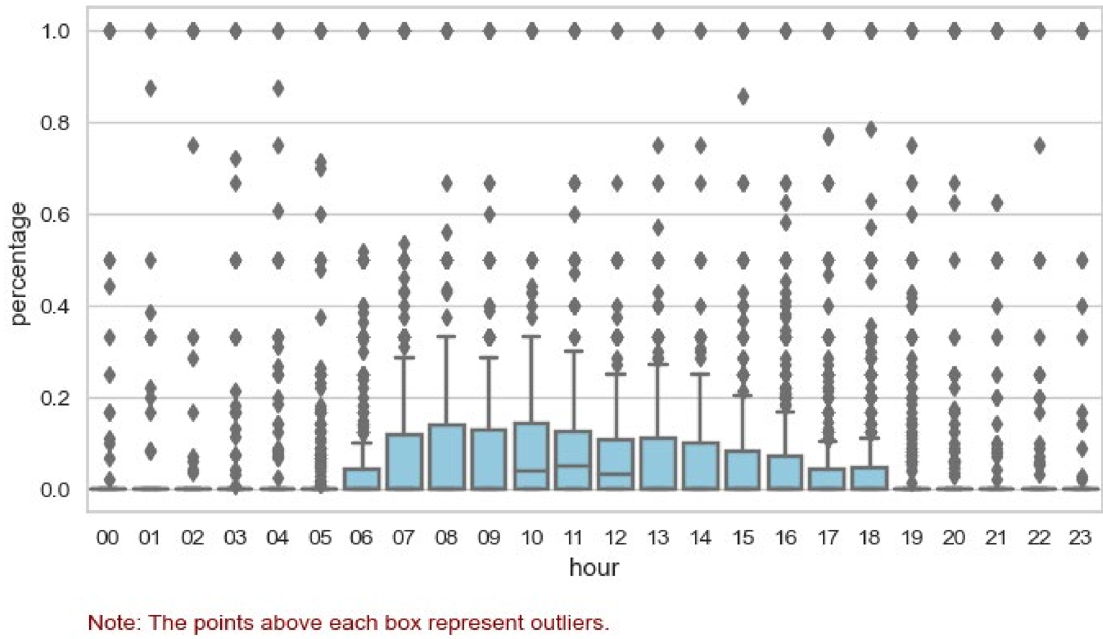


Figure 16. Hourly distributions of parking search (Boxes indicate the central 50% of values (interquartile range); lines within boxes indicate the medians, where shown; whiskers extend to 1.5 × the interquartile range)

Figure 17 presents a scatterplot designed to pinpoint specific geohashes where drivers are compelled to cover extensive distances within a confined area while searching for parking. This analytical visualization helps to discern the intricacies of parking search patterns by considering two metrics: the average total geohashes traversed and the average number of unique geohashes covered during each search for parking, including the final geohash where the parking spot is found.

From a disaggregate perspective, the *total geohashes traveled* (TGT) is computed as the cumulative sum of all geohashes a vehicle traverses before successfully securing a parking spot. It is mathematically represented as:

$$TGT = \sum_{i=1}^{n} G_i$$

where G_i denotes each geohash entered during the search and n is the total count of such geohashes.

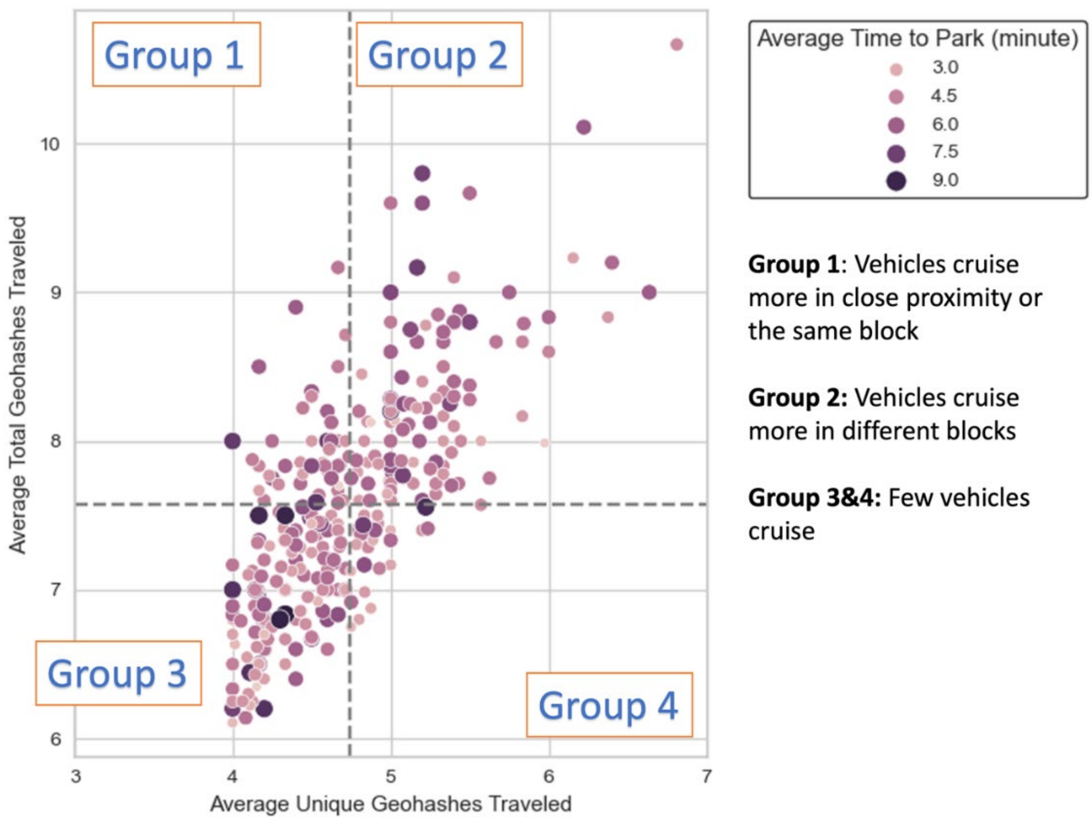
In contrast, *unique geohashes traveled* (UGT) accounts for the distinct geohashes entered, not counting repeated visits to the same geohash and representing the spatial expanse of the search. It is expressed as:

$$UGT = |\{G_1, G_2, \dots, G_n\}|$$

where each G_i is a unique geohash within the set.

For instance, consider a scenario where a driver initiates a parking search in geohash A, traverses through geohash B and C before revisiting A and ultimately securing a spot in geohash D. Here, the TGT is 5, and the UGT is 4. When the TGT and UGT are aggregated at the level of geohash and averaged, then they become the metrics analyzed in this study.

In Figure 17, data points are partitioned into four distinct clusters, grounded in the mean values of average UGT (x-axis) and average TGT (y-axis). Group 1 (low average UGT, high average TGT) and Group 2 (high average UGT, high average TGT) includes areas marked by intensive and protracted parking searches, indicative of parking scarcity. These two groups of geohashes are those where drivers must cover longer distances to find parking. The scatterplot is instrumental for urban planners and policymakers, illuminating specific geohashes characterized by exacerbated parking searches, thereby underscoring the imperative for targeted interventions to amplify parking supply and mitigate associated environmental externalities.



Group 1: Vehicles cruise more in close proximity or the same block

Group 2: Vehicles cruise more in different blocks

Group 3&4: Few vehicles cruise

Figure 17. Scatterplot of average TGT vs. average UGT

Truck Parking Demand Analysis

Figure 18 presents KDE density maps of hourly truck parking locations that illustrate the spatial and temporal distribution of truck parking activities on an hourly basis. Each shaded area (area within contour) is indicative of the concentration of trucks that have been observed to be parked for extended durations. This is not a percentage but rather a density measure derived from the KDE analysis of observed long-term truck parking behaviors within specific geohashes.

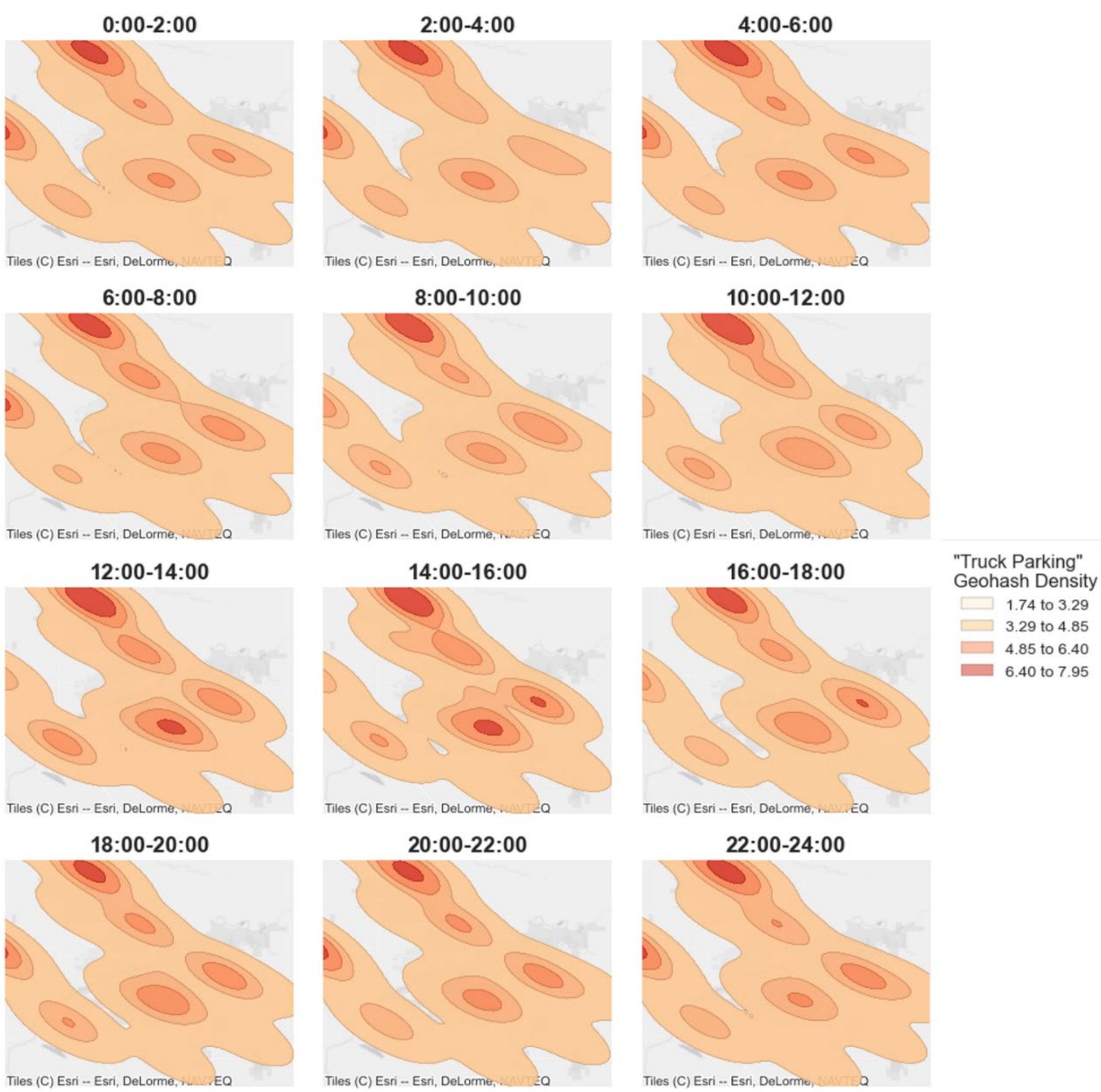


Figure 18. KDE maps showing hourly truck parking

The analysis reveals that nighttime parking hotspots are predominantly concentrated near the I-5 highway, while the SR-99 highway near Shafter exhibits high parking demand throughout the day. In contrast, daytime parking hotspots are primarily found in downtown Bakersfield, Lamont, and Arvin. These findings underscore the importance of considering both spatial and temporal dimensions when addressing truck parking challenges.

We executed a k-means clustering analysis on a dataset detailing truck parking locations, with the hourly truck parking percentages (specifically, the percentage of HDTs that are parked) in each geohash serving as the variable of interest. The objective was to categorize the geohashes into distinct clusters characterized by their periods of activity—resulting in two defined categories: "daytime active" and "nighttime active" truck parking.

Figure 19 represents the quantity of geohashes associated with each cluster type, providing a clear visual contrast between the two. Figure 20 shows the hourly distribution of truck parking activities, displayed through intricate boxplots that convey variations in parking percentages at different times of the day. Figure 21 complements this data by mapping the spatial distribution of the cluster labels, offering a geographical context to the behavioral patterns observed.

Key Insights:

- A predominance of "daytime active" geohashes was noted, indicating a higher frequency of truck parking activities during the day as compared to the night.
- The geohashes exhibited a distinct clustering pattern, bifurcating into areas characterized by heightened activity during the day and others more active during the night. The "nighttime active" geohashes recorded a peak parking percentage of up to 13.8% during the early morning hours (0-6 AM). In contrast, the "daytime active" geohashes demonstrated elevated parking percentages from 8 AM to 8 PM, with mean values surpassing those of the "nighttime active" category during these hours.
- The discrepancy in parking patterns between the two types of geohashes points towards the existence of distinct characteristics and behaviors associated with each cluster. This differentiation could be indicative of varying operational schedules, regulatory restrictions, or the influence of specific industrial or commercial activities in the vicinity.

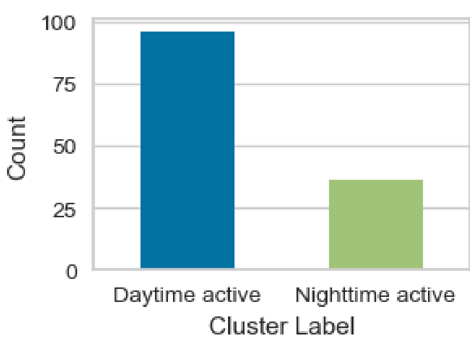


Figure 19. Number of geohashes within clusters with predominantly daytime or nighttime truck parking

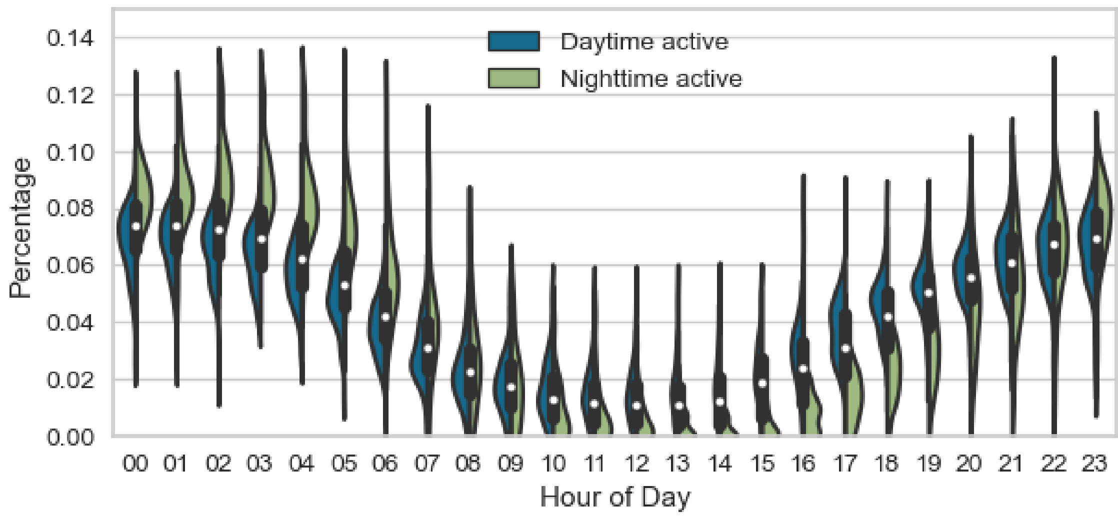


Figure 20. Hourly distribution of truck parking location

Figure 21. Spatial distribution of daytime and nighttime truck parking

In addition, the daytime-active type is less prevalent than the nighttime-active type (Figure 19) and is mainly distributed in Shafter, Arvin, Lamont, and near Bakersfield downtown. These findings highlight the distinct parking demand patterns between daytime-active and nighttime-active truck parking locations, with daytime-active locations being less common but concentrated in specific areas. This information can be valuable for policy and planning efforts to improve truck parking availability and utilization, and it emphasizes the need to consider both the spatial and temporal dimensions of truck parking demand. The noticeable concentration of daytime-active parking in specific areas presents an opportunity for focused interventions. There exists a potential to significantly enhance parking efficiency by zoning those hotspots and setting up commercial vehicle only parking spaces. The insight gained underscores the necessity of a nuanced, location-specific approach in devising strategies to mitigate parking challenges, aligning resources and efforts where they can yield the most impact.

Truck Activities Predictive Modeling

In this second part of the analysis, we explore the predictive modeling of truck activities, including truck idling and parking.

In the following subsections, we further discuss and analyze the performance of these models and the implications of the features with high importance for each predictive modeling task.

Idling Modeling

a. Idling Model Comparison and Selection Results

In this process of idling modeling, we compare the performance of several models in predicting truck idling behavior. Table 2 presents the performance metrics for each model: mean squared error (MSE), root mean squared error (RMSE), mean absolute error (MAE), R-squared (R2), and adjusted R2. Among the evaluated models, the cross-validated random forest (CV-RF) model outperforms the others in terms of R-squared and adjusted R2, with values of 0.3016 and 0.2797, respectively. Additionally, the CV-RF model has the lowest MAE (2.9287) among all models, indicating its superior ability to predict idling behavior accurately. Although the cross-validated convolutional neural network (CV-CNN) model shows a slightly lower MSE (26.1052) than the CV-RF model (26.7320), the CV-RF model still has the better overall performance.

Table 2. Idling Model Performance Metrics

Model	MSE	RMSE	MAE	R2	Adj. R2	# of Obs.	# of Feat-ures	Training/ Testing Split Ratio
Bi-layer neural network (NN)	37.4609	6.1205	3.7841	0.0213	0.0001	9368	57	80%/20%
Decision tree (DT)	52.6250	7.2543	4.1876	0.0001	0.0001	9368	57	80%/20%
Convolutional neural network (CNN)	30.2247	5.4977	3.0512	0.2103	0.1855	9368	57	80%/20%
Cross-validated CNN (CV-CNN)	26.1052	5.0922	3.2219	0.2186	0.1876	9368	57	80%/20%
K-nearest neighbors (KNN)	27.8410	5.2765	3.2100	0.2726	0.2498	9368	57	80%/20%
Cross-validated random forest (CV-RF)	26.7320	5.1703	2.9287	0.3016	0.2797	9368	57	80%/20%

The random forest model's inherent feature selection capability and low risk of overfitting to overfitting make it an ideal choice for the idling analysis. The cross-validation aspect of the CV-RF model ensures that the model's performance is consistent across different training and testing sets, reducing the risk of overfitting and enhancing generalizability. Considering these factors, we chose the CV-RF model as the primary modeling approach for idling analysis in this study.

b. Feature Importance Analysis of Idling Model

Table 3 lists the feature importance of the selected CV-RF model. The top six most important features are: (1) the percentage of HDTs, (2) the acrage of an area that is dedicated to a specific land use and is within 800 m (a proxy for a 10-minute walk), (3) distance to nearest POI (m), (4) Gini index of income inequality, (5) population aged between 10 and 64, and (6) distance to major road (m).

- Percentage of HDTs (PercentHDT): As the proportion of heavy-duty trucks increases, the idling time also tends to increase. This positive relationship becomes more pronounced at higher HDT percentages, indicating that areas with a larger share of heavy-duty trucks might experience more significant idling issues.
- Specific land use area in acre (within 800 m): The partial dependence plot (PDP) reveals a positive relationship between specific land use areas (commercial, industrial, retail, and construction) and idling time. However, this relationship plateaus when the specific land use area reaches around 14,000 acres, suggesting that the impact of land use on idling might diminish beyond this threshold.
- Distance to point-of-interest (m): The PDP demonstrates a negative relationship between the distance to the nearest point-of-interest and idling time. As the distance to the nearest POI increases, idling time decreases. This finding implies that areas closer to POIs may experience more idling issues.
- Gini index of income inequality: The PDP shows a slight positive relationship between the Gini index of income inequality and idling time, suggesting that areas with higher income inequality might experience more significant idling problems.
- Population ages 10–64: The PDP indicates a positive relationship between the population of ages 10–64 and idling time. This finding implies that areas with a larger working-age population may be more prone to idling issues.
- Distance to major road (m): The PDP reveals a positive relationship between the distance to the major road and idling time. As the distance to the major road increases, idling time tends to increase as well.

Table 3. Idling Model Feature Importances

Feature Name	Feature Importance	Permutation Importance
Percentage of HDTs	0.099545	0.209899
Specific land use area in acres within 800 m	0.100089	0.150485
Distance to nearest POI (m)	0.094304	0.075661
Gini index of income inequality	0.039882	0.053827
Population ages 10–64 yr	0.040119	0.037979
Distance to major road (m)	0.078926	0.037873
Linguistic isolation	0.032857	0.029165
Population with some college, less than 1 Year	0.016716	0.024541
Percentage of MDTs	0.037234	0.023874
Population with associate's degree	0.030646	0.021047

c. Idling Model Diagnostics

After refining the CV-RF model by reducing the number of features to 20, we obtained the performance metrics listed in Table 4. The refined model's performance is relatively good, considering the complexity of the idling behavior and the limited number of features used. The R-squared and adjusted R2 values respectively indicate that the model explains approximately 29.51% and 28.75% of the variability in the idling time. Moreover, the average cross-validation score of 0.2759 suggests that the model has a reasonable predictive ability and is not overfitting the data.

Table 4. Performance Metrics of Idling Model with Reduced Features

Metric	Value
Average CV Score	0.2759
Mean squared error (MSE)	26.9790
Root MSE (RMSE)	5.1941
Mean absolute error (MAE)	2.9462
R-squared	0.2951
Adjusted R2	0.2875

The relatively low MSE, RMSE, and MAE values indicate that the model's predictions are relatively close to the observed values. These metrics demonstrate the model's potential utility in understanding and predicting idling behavior, which could be beneficial for policymakers and urban planners looking to reduce idling-related emissions and improve air quality.

While the refined model shows satisfactory performance, there is still room for improvement. Further research could focus on incorporating additional relevant features or exploring alternative modeling techniques to enhance the model's predictive power and better understand the factors driving truck idling behavior.

Searching for Parking Modeling

a. Searching for Parking Model Comparison and Selection Results

In this part, we compared the performance of different models for predicting truck parking search time. Notably, we fitted two separate datasets, the full dataset containing all geohashes and the truck-prioritized dataset containing only geohashes with 50% or more trucks, to assess the performance of the models under different circumstances. The rationale behind this decision is to investigate how the models behave when faced with areas with a higher concentration of trucks, which are of greater interest for truck parking management and policymaking. Fitting the models on both datasets allows us to evaluate their performance and robustness in capturing the parking search dynamics in areas with high truck parking demand. Table 5 presents the performance metrics for each model.

Table 5. Searching for Parking Model Performance Metrics

Model	Dataset	# of Obs.	# of Features	MSE	RMSE	MAE
DT	Full	322	60	4.5380	2.1303	1.6200
	Truck-prioritized	173	60	2.1865	1.4787	1.1460
RF	Full	322	60	1.3463	1.1603	0.9585
	Truck-prioritized	173	60	2.2365	1.4955	1.1350
Bayesian Ridge	Full	322	60	1.4704	1.2126	0.9976
	Truck-prioritized	173	60	2.0975	1.4483	1.1012
KNN	Full	322	60	1.5218	1.2336	1.0127
	Truck-prioritized	173	60	2.2472	1.4991	1.0949
CNN	Full	322	60	1.4856	1.2189	1.0183
	Truck-prioritized	173	60	2.3148	1.5215	1.1249

DT, decision trees; RF, random forest; KNN, *k*-nearest-neighbor; CNN, convolutional neural network

Based on the performance metrics, the random forest (RF) model outperforms the other models on the full dataset, with the lowest values for MSE, RMSE, and MAE (1.3463, 1.1603, and 0.9585, respectively). This indicates that the RF model has the best predictive accuracy among the tested models for the full dataset. For the truck-prioritized dataset, the Bayesian Ridge model has the lowest MSE (2.0975) and RMSE (1.4483), while the RF model has the lowest MAE (1.1350). However, the differences in performance metrics between the Bayesian Ridge model and the RF model are relatively small. Given the RF model's superior performance on the full dataset and its comparable performance on the truck-prioritized dataset, we chose the RF model as the most suitable model for predicting parking search time. The RF model's superior performance can be attributed to its ability to handle complex relationships between the features and the target variable. Additionally, the RF model is more resistant to overfitting.

b. Feature Importance Analysis of Searching for Parking Model

In the feature importance analysis, the most important features (see Table 6) in the RF model for the full dataset and the truck-prioritized dataset are analyzed.

Table 6. Searching for Parking Model Feature Importances

Rank	Full Dataset		Truck-Prioritized Dataset	
	Feature Name	**Importance**	**Feature Name**	**Importance**
1	Distance to Major Road (m)	0.066	Percentage of LDTs	0.104
2	Employees of Industries not classified	0.060	Employees of Agriculture, Forestry, Fishing and Hunting	0.100
3	Employees of Agriculture, Forestry, Fishing and Hunting	0.058	Distance to Major Road (m)	0.060
4	Percentage of Other Vehicle Types	0.053	Population with Regular High School Diploma	0.056
5	Percentage of LDTs	0.047	Employees of Industries not classified	0.053
6	Linguistic Isolation	0.039	Specific Land Use Area in Acre (800m)	0.044
7	Employees of Educational Services	0.038	GED or Alternative Credential	0.042
8	Total Housing Units	0.035	Percentage of MDTs	0.038
9	Specific Land Use Area in Acre (800m)	0.035	Population with Some College, Less than 1 Year	0.037
10	Population with Regular High School Diploma	0.030	Gini Index of Income Inequality	0.033

For the full dataset, the top five important features are:

- Distance to Major Road (m): A shorter distance to major roads implies easier access to parking spaces, which may result in a shorter search time for parking.
- Employees of Industries not classified: A higher number of employees in industries not classified may indicate more diverse land use or parking demand, making it harder to find available parking spaces.
- Employees of Agriculture, Forestry, Fishing and Hunting: A higher number of employees in this industry might be related to the presence of larger, undeveloped areas, which could affect the availability of parking.

- Percentage of Other Vehicle Types: A higher percentage of other vehicle types indicates a more diverse mix of vehicles competing for parking spaces, potentially increasing search time.
- Percentage of LDTs: An increase in the percentage of LDTs can lead to higher demand for parking spaces, affecting search time.

For the truck-prioritized dataset, the top five important features are:

- Percentage of LDTs: A higher percentage of LDTs implies higher demand for parking spaces, making it more challenging to find available spots.
- Employees of Agriculture, Forestry, Fishing and Hunting: Similar to the full dataset, more employees in this industry might affect parking availability.
- Distance to Major Road (m): Like in the full dataset, a shorter distance to major roads might lead to easier access to parking spaces and shorter search times.
- Population with Regular High School Diploma: This might indicate socio-economic factors that affect the parking situation in the area.
- Employees of Industries not classified: As mentioned for the full dataset, more employees in industries not classified can indicate more diverse land use or parking demand.

c. Searching for Parking Model Diagnostics

We analyzed the performance of the selected RF model in predicting parking search time for both the full and truck-prioritized datasets. For the full dataset model, the MSE is 1.3463, RMSE is 1.1603, and MAE is 0.9585. These relatively low error values indicate that the model demonstrates satisfactory performance in predicting parking search time for the full dataset, capturing the underlying patterns in the data. For the truck-prioritized dataset model, the MSE is 2.2365, RMSE is 1.4955, and MAE is 1.1350. Although the error values are higher than those for the full dataset model, they still suggest a reasonably accurate model for predicting parking search time for areas with a higher concentration of trucks.

Overall, the RF model satisfactorily predicts parking search time for both datasets. The model's ability to capture the underlying patterns in the data can provide valuable insights for policymakers and city planners to make informed decisions in optimizing parking infrastructure and addressing parking issues, particularly in areas with a higher concentration of trucks.

d. Fishnet Grid Prediction of Parking Search Time

In our analysis, we adopted a sophisticated approach to quantify the relationship between road density and the propensity for parking search behavior, specifically within each 400m x 400m grid. To achieve a more refined, yet intuitive representation, we employed the sigmoid transformation function, as illustrated in Figure 22. This mathematical technique is particularly adept at handling scenarios in our study, where the goal is to encapsulate the nuanced correlation between road density (expressed in miles per squared mile) and the increased likelihood of parking search occurrences.

The sigmoid transformation function is notable for its capacity to map a wide range of road density values onto a more manageable scale of 0.05 to 1.0. This conversion is not arbitrary but is instrumental in deriving meaningful insights from the complex interplay between road density and parking search behavior. As the road density intensifies, the sigmoid function responds by indicating a proportional increase in the propensity for parking search, offering a nuanced view of this dynamic relationship. Utilizing the term "propensity" underscores the probabilistic nature of parking search behavior. It encapsulates the likelihood, rather than certainty, of such activities occurring within the specified grids, enabling us to paint a more realistic and nuanced picture of parking search patterns. This is especially pertinent given the multifaceted determinants that influence parking search behavior, including but not limited to, road density.

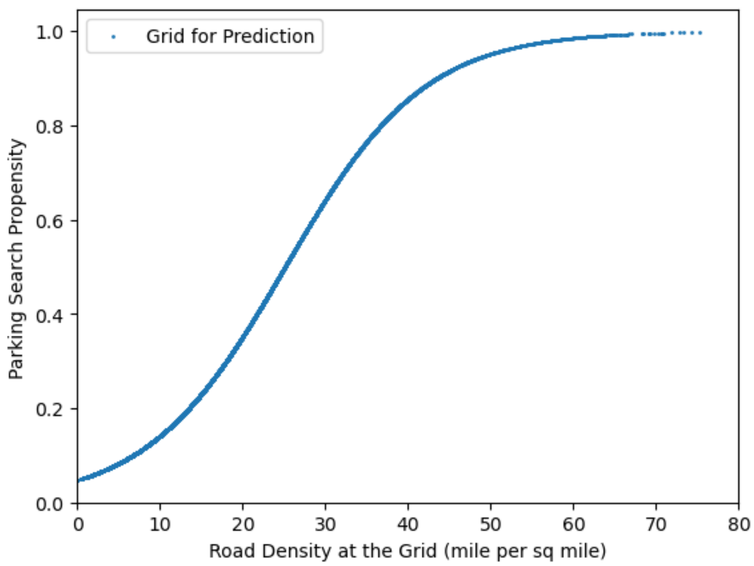
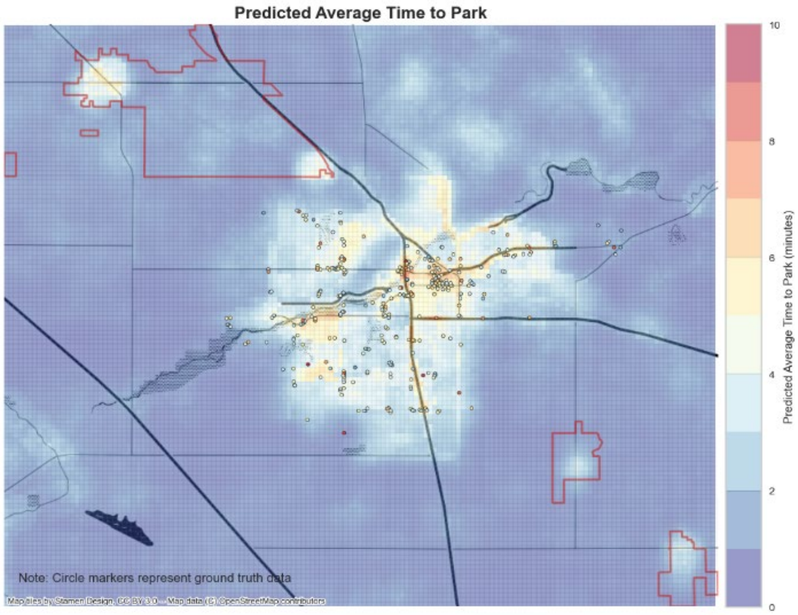


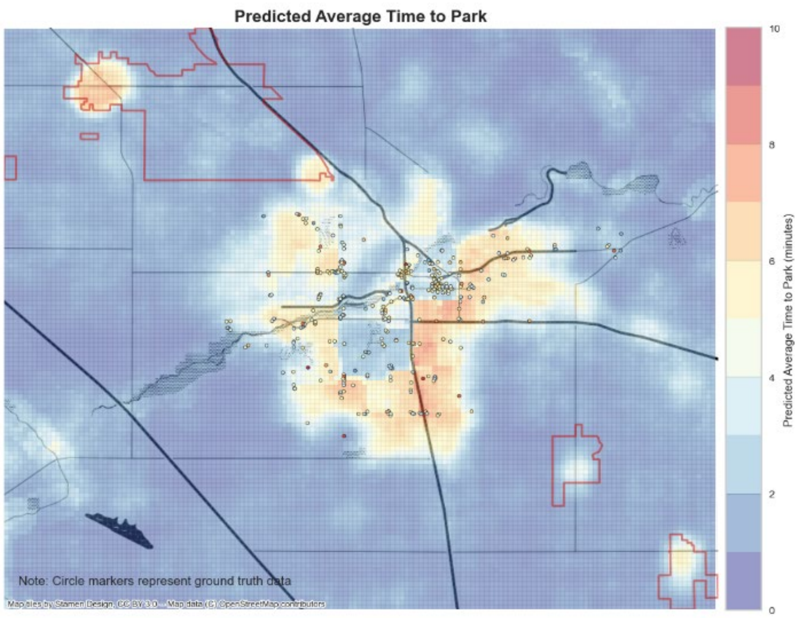
Figure 22. Sigmoid transformation function of road density

Figure 23 presents fishnet grid prediction maps of the average time to park for the full model (a) and the truck-prioritized model (b). For the full model, parking search hotspots are primarily distributed in the Oleander/Sunset, Wible Orchard, Campus Park, and River Walk Park areas near the SR-99/SR-58 interchange in Bakersfield. These areas have a relatively high concentration of freight transportation and warehousing points of interest, and commercial land use. However, the truck-prioritized model's prediction results show parking search hotspots in downtown Bakersfield and River Walk Park. The differences in the hotspots identified by the two models can be attributed to the truck-prioritized model focusing on areas with higher concentrations of trucks, which are more likely to have parking issues in dense urban settings such as downtown Bakersfield.

Both the full and truck-prioritized models also reveal additional parking search hotspots in Shafter downtown, Lamont, and Arvin. These areas might experience parking search issues due to a combination of factors such as road density, land use, and the presence of various vehicle types competing for limited parking spaces. Understanding the spatial distribution of parking search hotspots can help policymakers and city planners better target their interventions and improve the overall efficiency of the parking system, particularly in areas with high truck concentrations.



a) Full model



b) Truck-prioritized model

Figure 23. Fishnet grid prediction maps of average time to park

Truck Parking Demand Modeling

a. Feature Importance Analysis of Truck Parking Demand Model

Here, we analyze the most important features from Table 7 and discuss their relationship with truck parking demand, as represented by truck AADT (AADT_K).

Table 7. Truck Parking Location Model Feature Importances

Feature Name	Importance
Commercial land use area (within 800 m)	0.2077
Distance to major road (m)	0.1061
Gini Index of Income Inequality	0.0766
Total population	0.0756
Population with some college, less than 1 year	0.0654
Population with GED or alternative credential	0.0457
Total occupied housing units	0.0433
Owner-occupied housing units	0.0422
Population with regular high school diploma	0.0405
Employees of agriculture, forestry, fishing and hunting	0.0381

- Commercial land use area (0.2077 importance): An increase in commercial land use area typically corresponds with a higher demand for truck parking due to the need for goods delivery and transportation services. This highlights the importance of ensuring adequate parking facilities in commercial zones.
- Distance to major road (m) (0.1061 importance): The proximity to major roads is crucial in determining truck parking demand. Shorter distances to major roads increase the accessibility of parking spaces and make it more convenient for truck drivers to find parking, which in turn affects overall truck parking demand.
- Gini Index of income inequality (0.0766 importance): A higher Gini Index might be associated with a higher concentration of low-income households, which could lead to an increased demand for truck parking, as these households are more likely to rely on goods and services provided by trucks.
- Total population (0.0756 importance): An increase in the total population generally leads to a higher demand for goods and services, which in turn translates into a higher demand for truck parking spaces.
- Population with some college, less than 1 year (0.0654 importance): This demographic may be associated with a higher demand for truck parking, possibly due to their occupational choices or their consumption patterns.

- Population with GED or alternative credential (0.0457 importance): Similar to the population with some college, individuals with a GED or alternative credential may be associated with higher truck parking demand based on their occupational choices or consumption patterns.

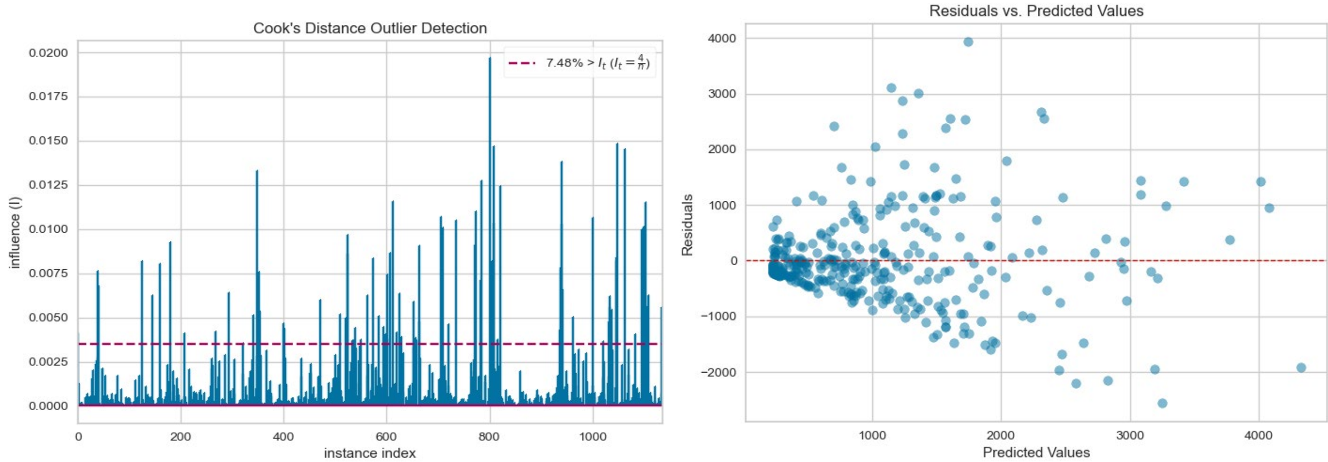
Other important features include the number of occupied housing units, owner-occupied housing units, the population with a high school diploma, employees in agriculture, forestry, fishing, and hunting, linguistic isolation, and the percentage of income spent on rent. These factors could also be related to truck parking demand through their influence on the local economy, land use patterns, and overall demand for goods and services.

b. Truck Parking Demand Model Diagnostics

We evaluated the performance of the optimized RF model using the following metrics:

- Mean squared error (MSE): 731248.0417
- Root mean squared error (RMSE): 855.1304
- R-squared: 0.4561

These performance metrics indicate that the RF model has a moderate level of predictive accuracy, as the R-squared value of 0.4561 suggests that approximately 45.61% of the variation in truck parking demand can be explained by the model. Although the model shows some level of predictive power, there is still room for improvement to increase its accuracy. Figure 24 presents the truck parking demand model diagnostics, including the residual plot (a) and Cook's distance diagram (b). The residual plot shows no apparent patterns or relationships, suggesting that the model assumptions of independence and constant variance are generally met. However, the errors tend to increase as the predicted values increase, indicating that the model might be less accurate for higher truck parking demand predictions. The Cook's distance diagram shows that the maximum influence value is only 0.02, indicating that no significant outliers could impact the model's performance. This suggests that the model is robust and not overly influenced by individual data points.



a) Residual plot b) Cook's distance diagram

Figure 24. Truck Parking Demand Model Diagnostics

c. Sensitivity Analysis with Varying Truck Percentages

In the truck parking demand modeling section, we performed a sensitivity analysis to investigate the impact of different truck percentage thresholds (K%) on the model's performance.

Table 8 shows the performance metrics (MSE, RMSE, and R-squared) for various K% thresholds ranging from 5% to 50%. As K% increases, the MSE and RMSE values also increase, indicating that the model's error grows as the threshold for truck parking demand rises. However, the R-squared values remain relatively stable at around 0.46, suggesting that the proportion of variance in truck parking demand explained by the model does not change significantly with different K% thresholds.

Table 8. Sensitivity Analysis of K% Threshold on Truck Parking Demand Model Diagnostics

K%	MSE	RMSE	R-squared
5%	80661.2	284.0	0.4601
10%	322644.7	568.0	0.4601
15%	724295.5	851.1	0.4613
20%	1290578.7	1136.0	0.4601
25%	2024480.4	1422.8	0.4579
30%	2897182.1	1702.1	0.4613
35%	3953224.0	1988.3	0.4599
40%	5162314.8	2272.1	0.4601
45%	6546071.3	2558.5	0.4590
50%	8097921.7	2845.7	0.4579

Figure 25 illustrates the sensitivity analysis results with varying truck percentages. The left y-axis represents the log10(MSE) values, and the right y-axis represents the R-squared values. The figure confirms that as K% increases, the MSE values increase, but the R-squared values remain stable around 0.46.

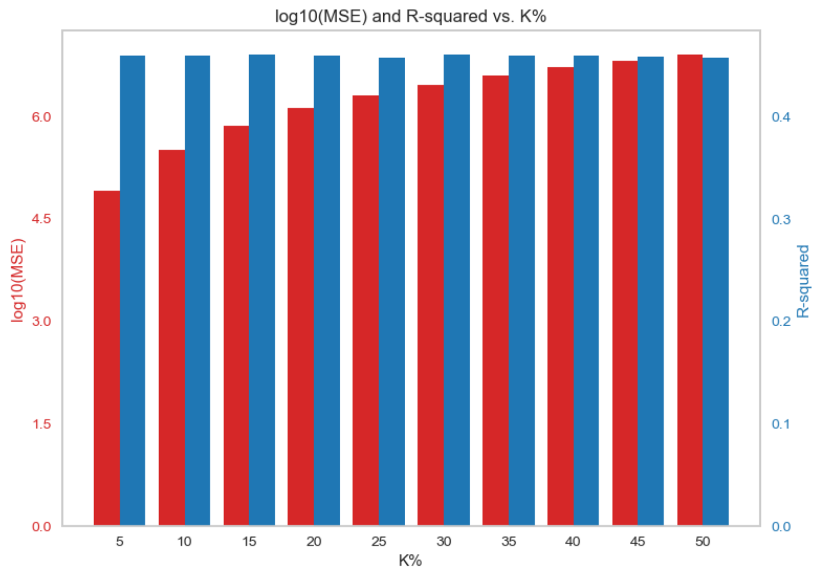


Figure 25. Sensitivity analysis results with varying truck percentages

Based on these findings, we chose a K% threshold of 15% for the following reasons:

- At 15%, the model achieves a relatively low MSE (724295.5) and RMSE (851.1) compared to higher K% thresholds, suggesting that the model has a lower error at this threshold.
- The R-squared value at 15% (0.4613) is slightly higher than other thresholds, indicating that the model explains a marginally greater proportion of variance in truck parking demand at this threshold.
- Selecting a threshold of 15% provides a balance between capturing a significant proportion of truck parking demand and minimizing the model's error.

By choosing a K% threshold of 15%, we aim to optimize the model's performance while accounting for the sensitivity of the model to different truck parking demand thresholds.

Correlation Analysis on Emission and Environmental Factors

a. SEM Model Performance Comparison

Using spatial error model (SEM), we compare the performance of the simple model (Y1, Y2, and Y3 only) and the pooled model (Y1, Y2, Y3, and other socio-economic variables) based on the model performance metrics provided in Table 9.

Table 9. SEM Model Performance Comparison

Dependent Variable	Simple Model (Y1, Y2 and Y3 only)					
	AIC	LL	PR2	Y1 p-value	Y2 p-value	Y3 p-value
CalEnviroScreen Score	-290.51	149.25	0.021	0.5955	0.5385	0.9567
Ozone Concentration	-442.32	225.16	0.152	0.4906	0.2891	0.2832
PM2.5 Concentration	119.19	-55.59	0.080	0.4235	0.1235	0.4102
Diesel PM Emissions	805.01	-398.51	0.044	0.8544	0.5642	0.2933
Pesticide Use	2,959.02	-1,475.51	0.012	0.7702	0.0285	0.6275
Toxic Release to Air	453.53	-222.77	0.005	0.3114	0.0104	0.9464
Drinking Water Contamination	30.52	-11.26	0.163	0.5831	0.3902	0.3234
Lead Risk	-455.38	231.69	0.041	0.5498	0.5468	0.2208
Pollution Burden Score	-204.70	106.35	0.061	0.5076	0.2954	0.6871
Asthma Prevalence	-689.56	348.78	0.055	0.6035	0.3279	0.6375
Low Birth Weight Incidence	-599.82	303.91	0.044	0.4007	0.3159	0.7602
Cardiovascular Disease Prevalence	-768.91	388.46	0.156	0.5930	0.4301	0.3641
	Pooled Model (Y1, Y2, Y3 and other socio-economic variables)					
CalEnviroScreen Score	-1,681.11	854.55	0.840	0.6531	0.2560	0.2014
Ozone Concentration	-2,503.06	1,265.53	0.923	0.8442	0.1807	0.0389
PM2.5 Concentration	-2,082.57	1,055.28	0.928	0.9487	0.4777	0.0143
Diesel PM Emissions	493.15	-232.57	0.679	0.5290	0.6130	0.6098
Pesticide Use	2,868.51	-1,420.26	0.074	0.7471	0.0352	0.6127
Toxic Release to Air	-354.03	191.01	0.708	0.1977	0.0706	0.7057
Drinking Water Contamination	-2,101.00	1,064.50	0.906	0.5790	0.1012	0.0367
Lead Risk	-1,156.30	592.15	0.686	0.9797	0.0438	0.1407
Pollution Burden Score	-2,302.92	1,165.46	0.905	0.7538	0.2194	0.1175
Asthma Prevalence	-1,794.74	911.37	0.779	0.6090	0.4374	0.4183
Low Birth Weight Incidence	-1,875.93	951.96	0.862	0.9190	0.3620	0.5420
Cardiovascular Disease Prevalence	-2,457.38	1,242.69	0.885	0.6137	0.1585	0.0501

From the table, we can observe that the pooled model generally outperforms the simple model in terms of Akaike information criterion (AIC), log-likelihood (LL), and pseudo R-squared (PR2) values. The pooled model has lower AIC and higher LL values, indicating better goodness-of-fit. Additionally, the pooled model has higher pseudo R-squared values for most of the dependent variables, which means the complex model explains a larger proportion of the variance in these variables. However, a comparison of the p-values for Y1, Y2, and Y3 between the two models does not consistently indicate which model performs better. For some dependent variables, the simple model has lower p-values, while for others, the complex model has lower p-values. Lower p-values indicate a stronger relationship between the independent and dependent variables.

Specifically, for PM 2.5 concentration, the pooled model shows a significant improvement in AIC (-2,082.57), LL (1,055.28), and pseudo R-squared (0.928) values compared to the simple model. Moreover, the p-value for Y3 in the pooled model (0.0143) is considerably lower than in the simple model (0.4102), indicating a stronger relationship between Y3 and PM 2.5 concentration in the pooled model. This suggests that the pooled model captures the effects of PM 2.5 concentration better than the simple model.

In short, the pooled model generally performs better regarding goodness-of-fit and explaining the variance in the dependent variables. Although the p-values do not consistently favor one model over the other, the complex model shows a significant improvement for PM 2.5 concentration, so we chose this model for further analysis.

b. Simple and Pooled Models for PM 2.5 Concentration and Toxic Release to Air

We specifically look at the models for PM 2.5 concentration and toxic release into the air. We discuss the effects of Y1, Y2, and Y3 in the simple and complex models. We also explain the relationship between other variables and PM 2.5 concentration and toxic release to air, based on Table 10 and 11.

For PM 2.5 concentration, the pooled model exhibits a much higher pseudo R-squared value (0.9283) than the simple model (0.0801), indicating that the pooled model explains a larger proportion of the variance. In the pooled model, Proximitized Truck Parking Demand (Y3) has a significant positive relationship with PM 2.5 concentration ($p < 0.01$). At the same time, the effects of Average Idle Time (Y1) and Average Time to Park (Y2) are not statistically significant. This suggests that, in the pooled model, truck parking demand plays a more important role in PM 2.5 concentration than Y1 and Y2. In addition, the pooled model highlights significant relationships between PM 2.5 concentration and several socio-economic variables, such as the number of employees in agriculture, forestry, fishing and hunting; in construction; in manufacturing; and in retail trade. These relationships help to explain the variation in PM 2.5 concentration further.

Table 10. Parameter Estimates for Simple and Pooled Models for PM 2.5 Concentration

Model Metrics	Simple Model		Pooled Model	
Pseudo R-squared		0.0801		0.9283
Sigma-square ML		0.048		0.007
Log likelihood		-55.595		1055.283
Akaike info criterion		119.189		-2082.566
Schwarz criterion		139.369		-2011.937
Variable	**Coefficient**	**Std.Error**	**Coefficient**	**Std.Error**
Constant	-0.004	1.435	0.013	0.279
Average Idle Time	-0.006	0.007	0.000	0.003
Average Time to Park	0.023	0.015	-0.004	0.006
Proximitized Truck Parking Demand	0.007	0.009	0.008**	0.003
Employees of Agriculture, Forestry, Fishing and Hunting			0.028***	0.003
Employees of Mining			-0.016***	0.005
Employees of Utilities			0.056***	0.008
Employees of Construction			0.043***	0.007
Employees of Manufacturing			0.036***	0.011
Employees of Wholesale Trade			-0.011	0.012
Employees of Retail Trade			0.134***	0.011
Employees of Transportation and Warehousing			-0.028***	0.009
Total Population			0.259***	0.016
Median Houshold Income			0.520***	0.020
Lambda	0.995***	0.002	0.991***	0.003

Note: *: $p<0.05$, **: $p<0.01$, ***: $p<0.001$.

The pooled model also shows a much higher pseudo-R-squared value (0.7077) for toxic release to air than the simple model (0.0048). In this case, Average Time to Park (Y2) has a significant positive relationship with toxic release to air in the simple model ($p<0.01$). Still, the effect becomes insignificant in the pooled model. Neither Y1 nor Y3 shows significant relationships with toxic release to air in either model. In the pooled model, several socio-economic variables have significant relationships with toxic release into the air. For example, employees in agriculture, forestry, fishing, hunting; in construction; and in wholesale trade significantly affect toxic release into the air.

Table 11. Parameter Estimates for Simple and Pooled Models for Toxic Release to Air

Model Metrics	Simple Model		Pooled Model	
Pseudo R-squared		0.0048		0.7077
Sigma-square ML		0.065		0.032
Log likelihood		-222.766		191.015
Akaike info criterion		453.533		-354.03
Schwarz criterion		473.712		-283.401
Variable	**Coefficient**	**Std.Error**	**Coefficient**	**Std.Error**
Constant	-0.141	0.898	-0.062	0.377
Average Idle Time	-0.008	0.008	-0.007	0.006
Average Time to Park	0.045**	0.018	0.022	0.012
Proximitized Truck Parking Demand	-0.001	0.010	-0.003	0.007
Employees of Agriculture, Forestry, Fishing and Hunting			0.046***	0.007
Employees of Mining			-0.017	0.010
Employees of Utilities			-0.004	0.016
Employees of Construction			0.086***	0.015
Employees of Manufacturing			0.026	0.023
Employees of Wholesale Trade			-0.052**	0.026
Employees of Retail Trade			-0.230***	0.022
Employees of Transportation and Warehousing			0.004	0.020
Total Population			0.067	0.035
Median Houshold Income			0.875***	0.043
Lambda	0.992***	0.003	0.986***	0.005

Note: *: p<0.05, **: p<0.01, ***: p<0.001.

Discussion and Policy Recommendations

This section examines the results and findings of the various analyses and their implications for understanding truck activities, air pollution and formulating policy recommendations. We focus on the practical implications of the models, highlighting the real-world impact of the factors identified in each analysis.

Idling Analysis Discussion

In the idling analyses, both random forest (RF) and convolutional neural network (CNN) models show promising predictive performance. They can accurately predict average idle time using only socio-economic, road network, land use, and point-of-interest data from the prediction site's census tract, with a low risk of overfitting. We ultimately chose the RF model for its balance of predictability and interpretability, allowing for a more convenient and intuitive identification of factors influencing average idle time.

Trucks, particularly heavy-duty trucks (HDTs), contribute significantly to idling activities, necessitating measures to reduce the impact of their idling. These measures may include improving truck engines, using cleaner fuels, and encouraging truck drivers to turn off their engines when appropriate. The exploratory analysis underscores the complexity of idling behavior. HDTs are not only the principal contributors to cumulative idling times but also exhibit a pronounced variance in their idling patterns. This heterogeneous behavior within a specific vehicle category amplifies the need for tailored strategies that address the nuanced variations in idling times. Furthermore, the spatial and temporal dynamics of idling behavior are integral to developing effective intervention strategies. The concentration of daytime idling around commercial and industrial hubs necessitates a location-specific approach to mitigate idling and its ensuing environmental impact. The insights gleaned from the analysis underscore the multifaceted nature of idling behavior, weaving together geographic location, and temporal dynamics.

Disadvantaged communities are often more susceptible to the negative impacts of idling activities than other communities, as they tend to be located near major roads and logistics centers with higher truck traffic. Thus, reducing truck idling activities is particularly important for improving these communities' air quality and environmental health conditions.

There are differences in the average idle time among different types of trucks: HDTs have the highest average idle time, followed by medium-duty trucks (MDTs) and light-duty trucks (LDTs). HDTs and MDTs exhibit greater variability in idling time than LDTs, which can be demonstrated using scatter plots and box plots.

Geographical disparities in idling activities are pronounced, especially concerning HDTs, which are identified as the primary contributors to extensive idling times. Specific areas exhibit intensified idling activities, marked by a combination of the volume of HDTs and the duration of their idling. The cumulative idling times are not uniform but vary significantly, underlining a heterogeneous pattern of idling behaviors even within the HDT category. Hourly idling distributions reveal distinct patterns: low idling throughout the day and night, heightened daytime idling, and increased nighttime idling. These patterns are not random but are geographically clustered. Daytime idling is notably concentrated in commercial and industrial hubs like Bakersfield and its surrounding areas, reflecting the intersection of truck activities and urban centers. In contrast, nighttime idling is predominantly located near major roads and highway entrances and exits, likely associated with rest periods and early morning deliveries.

Subsequent idling modeling results reveal the top six influencing factors through feature importance analysis: (1) the percentage of HDTs, (2) the acrage of an area that is dedicated to a specific land use and is within 800 m (a proxy for a 10-minute walk) (3) distance to the nearest point-of-interest (m), (4) Gini index of income inequality, (5) population aged between 10 and 64, and (6) distance to major road (m), reflecting traffic location.

The percentage of HDTs impacts idle time because HDTs generally have longer idle times, leading to higher average idle times in areas with more HDTs. Land use and traffic location influence idle time, as areas with specific land use types, such as logistics centers or industrial zones, and proximity to major roads may experience higher idling activities. The number of people and income inequality Gini index can affect idle time, as higher population density and income inequality may lead to more significant truck activities, especially in disadvantaged communities, resulting in longer idle times.

Based on these findings, a preliminary policy recommendation for addressing idling phenomena could be the implementation of targeted anti-idling campaigns and regulations in AB617 communities. This may involve raising public awareness of the negative impacts of idling, promoting behavioral changes among truck drivers, and enforcing penalties for excessive idling. In addition, providing designated rest areas and parking spaces with alternative power sources, such as electric charging stations, could encourage truck drivers to reduce engine idling while resting or waiting. Focusing on these initial strategies can improve air quality and public health while laying the groundwork for more comprehensive policy interventions in the future.

Searching for Parking Analysis Discussion

In the searching for parking analyses, we used various models to predict truck parking search times, comparing and evaluating these models using both the full and truck-prioritized datasets. The random forest (RF) model performed best for the full dataset. In contrast, the Bayesian Ridge Regression model demonstrated optimal performance in terms of MSE and RMSE for the truck-prioritized dataset, with the RF model excelling in MAE. Consequently, we selected the RF model as the most suitable for predicting parking search times.

Our exploratory analysis revealed insights into the challenges and issues faced by trucks during parking searches. Parking hotspots concentrate around transportation arteries, freight yards, and highway ramps, with longer parking search times, particularly in the Bakersfield downtown area and its western suburbs. Thus, it is crucial to implement measures in these areas to alleviate excessive parking search times, enhancing traffic efficiency and reducing congestion.

Key factors influencing parking search times were identified. For the full dataset, the most crucial features included distance to main roads, the number of employees in unclassified industries, the number of employees in agriculture, forestry, fishing, and hunting industries, the percentage of other vehicle types, and the percentage of large trucks. For the truck-prioritized dataset, the most critical features encompassed the percentage of large trucks, the number of employees in agriculture, forestry, fishing, and hunting industries, distance to main roads, the population with high school diplomas, and the number of employees in unclassified industries.

We discovered the spatial distribution of parking search hotspots by employing a sigmoid transformation function to convert road density into parking search behavior inclination and utilizing fishnet grid prediction maps to display parking search hotspot areas. This knowledge enables decision-makers and urban planners to target interventions better and improve the overall efficiency of parking systems.

Preliminary policy recommendations for improving truck parking facilities and management include constructing additional parking lots and providing more parking spaces to alleviate lengthy parking search times. Moreover, the parking needs of disadvantaged communities should be considered by offering dedicated loading and unloading berths and ample truck parking facilities away from residential areas, reducing the inconvenience and pressure caused by truck parking searches. Also, limiting truck operating hours in disadvantaged communities can effectively reduce their environmental impact.

For non-truck parking search behavior, such as private cars and MPVs, promoting emerging parking approaches like shared parking can help better use existing parking resources and minimize environmental impact. Efforts should also be made to reduce truck emissions during parking searches. This can be achieved by improving fuel efficiency, promoting electric vehicles, and implementing stricter emissions standards.

Truck Parking Demand Analysis Discussion

The discussion on truck parking demand analysis centers around the performance and importance of the random forest model in understanding the factors influencing truck parking demand and their implications for air quality.

Our study identified several significant factors affecting truck parking demand, including commercial land area, distance to main roads, Gini coefficient of income inequality, total population, population with some college education, and population with GED or alternative certificates. These factors offer valuable insights into the dynamics of truck parking demand, which can inform policymakers and urban planners in developing effective strategies to address truck parking issues and reduce truck parking search times and associated emissions.

The optimized RF model demonstrates a moderate predictive accuracy, highlighting the complexity of truck parking demand and suggesting the potential need for further improvements in the model. The increased errors in higher prediction values indicate that additional factors might be at play, and future studies could incorporate these factors to improve the model's performance and accuracy.

The sensitivity analysis of changing truck percentage thresholds reveals that a 15% threshold provides the best balance between capturing a significant proportion of truck parking demand and minimizing model errors. This finding can guide the selection of appropriate thresholds for future studies and policy interventions, ensuring they effectively address truck parking demand while minimizing unintended consequences for air quality.

We propose preliminary policy recommendations for addressing truck parking demand based on these insights. Key suggestions include improving truck parking facilities, promoting efficient truck routing, and incentivizing the adoption of cleaner fuels and technologies. By implementing these policy recommendations, decision-makers can better address truck parking demand, improving air quality and living conditions in AB 617 communities.

Discussion of Correlation Analysis on Emission and Environmental Factors

In our analysis, we employed a spatial error model (SEM) to compare the performance of a simple model (Y1, Y2, and Y3 only) and a pooled model (Y1, Y2, Y3, and other socio-economic variables) based on model performance metrics. It provides more insights into the relationships between Y1, Y2, Y3, and other socio-economic variables. The results indicate that Proximitized Truck Parking Demand (Y3) plays an essential role in PM 2.5 concentration, while socio-economic variables significantly explain the variations in both environmental indicators. These findings can be helpful for policymakers and stakeholders in developing effective strategies to address air pollution and its associated health impacts.

Policy Recommendations

Considering the findings from the previous discussions, a comprehensive policy framework addressing truck parking facilities, management, and air pollution control is crucial for improving air quality and living conditions in AB 617 communities. The following policy recommendations aim to address these intertwined issues and contribute to the overall well-being of the affected communities:

- Improve truck parking facilities: Develop designated truck parking areas with appropriate amenities to reduce the time trucks spend searching for parking, thereby minimizing emissions associated with idling and congestion. Additionally, encourage the use of intelligent parking systems that provide real-time information on parking availability to further reduce truck parking search times.

- Enhance truck routing efficiency: Implement intelligent transportation systems and use real-time traffic data to optimize truck routing, avoid congested areas, and promote off-peak delivery schedules. This will help reduce emissions caused by congestion and idling, improving overall air quality.
- Implement stricter emission standards: Promote the adoption of cleaner fuels, electric trucks, and advanced emission control technologies to reduce the environmental impact of truck activities. Provide incentives and financial support for the transition to cleaner vehicles, accelerating the shift towards more sustainable transportation systems.
- Strengthen land-use planning and zoning regulations: Create buffer zones between industrial, commercial, and residential areas, and promote mixed-use developments that encourage walking, cycling, and public transportation use. This will help reduce the exposure of vulnerable communities to air pollution and improve overall living conditions.

The implementation of these policy recommendations is vital for improving air quality and enhancing the quality of life in AB 617 communities. By addressing the complex interplay of factors contributing to air pollution, these policies can significantly impact community well-being.

However, several challenges and considerations must be taken into account during the implementation process. One challenge is coordinating efforts among various stakeholders, including government agencies, transportation companies, and local communities, to ensure a holistic and integrated approach to addressing air pollution issues. Additionally, securing adequate funding and resources for implementing the proposed policies may also pose a challenge.

Another consideration is the potential displacement of trucking activities to other areas, which may lead to unintended consequences for neighboring communities. To avoid this, it is important to carefully assess the potential impacts of policy interventions and incorporate appropriate mitigation measures where necessary.

In conclusion, a comprehensive policy framework addressing truck parking facilities, management, and air pollution control is essential for improving air quality and living conditions in AB 617 communities. By considering the challenges and considerations associated with policy implementation, decision-makers can develop effective strategies that promote sustainable development and improve the well-being of community residents.

Conclusion

This report sheds light on the complex interplay between truck idling, searching for parking, and air pollution in disadvantaged communities in Kern County, California (California Air Resources Board, 2021). The project aimed (i) to identify factors contributing to truck idling and searching for parking, with a particular emphasis on these communities; and (ii) to develop targeted planning and policy recommendations for local and state agencies to reduce the negative impacts of these activities on disadvantaged communities.

In conclusion, our study provides valuable insights into the factors affecting truck idling behavior, parking search times, and truck parking demand in AB617 communities and their implications for air quality and public health. We identified key contributing factors, such as the percentage of heavy-duty trucks, commercial, and transportation and warehousing land use, traffic location, and socio-economic variables, which can help inform policy recommendations to address these challenges. By implementing targeted anti-idling campaigns and regulations, improving truck parking facilities and management, and promoting cleaner fuels and technologies, policymakers can work towards improving air quality and living conditions in disadvantaged communities.

Furthermore, our analysis reveals the significant role that Proximitized Truck Parking Demand (Y3) plays in PM 2.5 concentration and the importance of socio-economic variables in explaining variations in environmental indicators. To address air pollution, policy recommendations include enhancing truck routing efficiency, implementing stricter emission standards, promoting cleaner fuels and electric trucks, and strengthening land-use planning and zoning regulations. By incorporating these strategies, decision-makers can contribute to reducing air pollution.

While this study provides valuable insights, it is essential to acknowledge its limitations and areas for future research. For example, the model to predict truck parking demand could be further improved to capture additional factors influencing demand (Guerrero et al., 2023). Moreover, future research should focus on evaluating the effectiveness of the recommended policy interventions, exploring the long-term impact of these interventions on air quality and community well-being, and investigating possible synergies between different policy measures to maximize their overall benefits. Further research could explore the potential for innovative technologies, such as autonomous vehicles and smart transportation systems, to address the challenges associated with truck parking demand and air quality.

By addressing these issues through informed policymaking and targeted interventions, it is possible to mitigate truck idling and parking emissions. This study can serve as a stepping stone towards a more sustainable and equitable future for all, emphasizing the importance of continuous research and collaboration between researchers, policymakers, and practitioners to develop and implement innovative solutions to address the pressing challenges of truck idling and parking issues and air quality in our communities.

References

Abdi, H., Valentin, D., Edelman, B., 1999. Neural Networks. SAGE.

Alho, A., Oh, S., Seshadri, R., Dalla Chiara, G., Chong, W.H., Sakai, T., Cheah, L., Ben-Akiva, M., 2022. An agent-based simulation assessment of freight parking demand management strategies for large urban freight generators. Res. Transp. Bus. Manag. 43, 100804. https://doi.org/10.1016/j.rtbm.2022.100804

Anti-Idling Laws Gain Momentum Around the Country [WWW Document], n.d. URL https://sustainableamerica.org/blog/anti-idling-laws-gain-momentum-around-the-country/ (accessed 1.30.23).

Bailey, D., Solomon, G., 2004. Pollution prevention at ports: clearing the air. Environ. Impact Assess. Rev., Environment and Health: new answers, new questions. 24, 749–774. https://doi.org/10.1016/j.eiar.2004.06.005

Boggs, A.M., Hezaveh, A.M., Cherry, C.R., 2019. Shortage of Commercial Vehicle Parking and Truck-Related Interstate Ramp Crashes in Tennessee. Transp. Res. Rec. 2673, 153–163. https://doi.org/10.1177/0361198119849586

Browne, M.W., 2000. Cross-Validation Methods. J. Math. Psychol. 44, 108–132. https://doi.org/10.1006/jmps.1999.1279

California Air Resources Board, 2021. San Joaquin Valley Air Pollution Control District AB 617 Community Emission Reduction Program.

California Air Resources Board, 2019. Community Emissions Reduction Program for Arvin/Lamont AB 617 Community approved by the California Air Resources Board | California Air Resources Board [WWW Document]. URL https://ww2.arb.ca.gov/news/community-emissions-reduction-program-arvinlamont-ab-617-community-approved-california-air (accessed 4.19.23).

California Air Resources Board, 2004. Airborne Toxic Control Measure to Limit Diesel-Fueled Commercial Motor Vehicle Idling | California Air Resources Board [WWW Document]. URL https://ww2.arb.ca.gov/our-work/programs/atcm-to-limit-vehicle-idling/about (accessed 1.30.23).

California Vehicle Code Section 2462 [WWW Document], n.d. URL https://california.public.law/codes/ca_veh_code_section_2462 (accessed 1.30.23).

CHAPTER 9. STOPPING, STANDING, AND PARKING - Sections 22500-22526 :: California Vehicle Code :: 2005 California Code :: California Law :: California Code :: US Codes and Statutes :: US Law :: Justia [WWW Document], n.d. URL https://law.justia.com/codes/california/2005/veh/22500-22526.html (accessed 1.30.23).

Conway, A., Tavernier, N., Leal-Tavares, V., Gharamani, N., Chauvet, L., Chiu, M., Bing Yeap, X., 2016. Freight in a Bicycle-Friendly City: Exploratory Analysis with New York City Open Data. Transp. Res. Rec. 2547, 91–101. https://doi.org/10.3141/2547-13

Debnath, A., Singh, P.K., 2018. Environmental traffic noise modelling of Dhanbad township area – A mathematical based approach. Appl. Acoust. 129, 161–172. https://doi.org/10.1016/j.apacoust.2017.07.023

Dhariwal, S., Sharma, A., Raipuria, A., 2022. Algorithms and Difficulties for Autonomous Cars Based on Artificial Intelligence, in: Autonomous Vehicles Volume 1. John Wiley & Sons, Ltd, pp. 27–37. https://doi.org/10.1002/9781119871989.ch2

Divekar, P., Han, X., Zhang, X., Zheng, M., Tjong, J., 2023. Energy efficiency improvements and CO2 emission reduction by CNG use in medium- and heavy-duty spark-ignition engines. Energy 263, 125769. https://doi.org/10.1016/j.energy.2022.125769

Efendi, A., Effrihan, 2017. A simulation study on Bayesian Ridge regression models for several collinearity levels. AIP Conf. Proc. 1913, 020031. https://doi.org/10.1063/1.5016665

Farahani, R., 2011. Logistics Operations and Management: Concepts and Models. Elsevier.

Farzaneh, R., Johnson, J., Jaikumar, R., Ramani, T., Zietsman, J., 2020. Use of Vehicle Telematics Data to Characterize Drayage Heavy-Duty Truck Idling. Transp. Res. Rec. 2674, 542–553. https://doi.org/10.1177/0361198120945990

Frey, H.C., Kuo, P.-Y., 2009. Real-World Energy Use and Emission Rates for Idling Long-Haul Trucks and Selected Idle Reduction Technologies. J. Air Waste Manag. Assoc. 59, 857–864. https://doi.org/10.3155/1047-3289.59.7.857

Frey, H.C., Kuo, P.-Y., Villa, C., 2009. Effects of Idle Reduction Technologies on Real World Fuel Use and Exhaust Emissions of Idling Long-Haul Trucks. Environ. Sci. Technol. 43, 6875–6881. https://doi.org/10.1021/es900186e

Gaines, L., Levinson, T., 2011. Idling: Cruising the Fuel Inefficiency Expressway.

Gaines, L., Vyas, A., Anderson, J.L., 2006. Estimation of Fuel Use by Idling Commercial Trucks. Transp. Res. Rec. 1983, 91–98. https://doi.org/10.1177/0361198106198300113

Gao, K., Tu, H., Sun, L., Sze, N.N., Song, Z., Shi, H., 2020. Impacts of reduced visibility under hazy weather condition on collision risk and car-following behavior: Implications for traffic control and management. Int. J. Sustain. Transp. 14, 635–642. https://doi.org/10.1080/15568318.2019.1597226

Giuliano, G., O'Brien, T., 2007. Reducing port-related truck emissions: The terminal gate appointment system at the Ports of Los Angeles and Long Beach. Transp. Res. Part Transp. Environ. 12, 460–473. https://doi.org/10.1016/j.trd.2007.06.004

Gopalakrishnan, R., Alho, A.R., Sakai, T., Hara, Y., Cheah, L., Ben-Akiva, M., 2020. Assessing Overnight Parking Infrastructure Policies for Commercial Vehicles in Cities Using Agent-Based Simulation. Sustainability 12, 2673. https://doi.org/10.3390/su12072673

Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., Liu, T., Wang, X., Wang, G., Cai, J., Chen, T., 2018. Recent advances in convolutional neural networks. Pattern Recognit. 77, 354–377. https://doi.org/10.1016/j.patcog.2017.10.013

Guerrero, S.E., Pulikanti, S., Wieghart, B., Bryan, J.G., Strow, T., 2023. Modeling Truck Parking Demand at Commercial and Industrial Establishments. Transp. Res. Rec. J. Transp. Res. Board 2677.

Hackeling, G., 2017. Mastering Machine Learning with scikit-learn. Packt Publishing Ltd.

Harris, N., 2014. Particulate Matter (PM) Exposure Assessment in a Local Highway Environment: Commuters' Perspective (M.S.). ProQuest Diss. Theses. North Carolina Central University, United States -- North Carolina.

Hasani, M., Jahangiri, A., Sener, I.N., Munira, S., Owens, J.M., Appleyard, B., Ryan, S., Turner, S.M., Ghanipoor Machiani, S., 2019. Identifying High-Risk Intersections for Walking and Bicycling Using Multiple Data Sources in the City of San Diego [WWW Document]. J. Adv. Transp. https://doi.org/10.1155/2019/9072358

Hopkins, J.S., 2017. The invisible hazard afflicting thousands of schools [WWW Document]. Cent. Public Integr. URL http://publicintegrity.org/environment/energy/carbon-wars/the-invisible-hazard-afflicting-thousands-of-schools/ (accessed 4.12.23).

Jaller, M., Holguín-Veras, J., Hodge, S.D., 2013. Parking in the City: Challenges for Freight Traffic. Transp. Res. Rec. 2379, 46–56. https://doi.org/10.3141/2379-06

Jaller, M., Pineda, L., Gueldas, Y., Alemi, F., Otay, I., 2020. Fostering the Use of Zero and Near Zero Emission Vehicles in Freight Operations. https://doi.org/10.7922/G2QC01SB

James, E., 2020. Assessing the Feasibility, Costs, and Benefits of Transitioning Part of the University of South Carolina Shuttle Fleet to an Alternative Fuel Source and Promoting Anti-idling Strategies. Sr. Theses 1–63.

Jonson, J.E., Borken-Kleefeld, J., Simpson, D., Nyíri, A., Posch, M., Heyes, C., 2017. Impact of excess NOx emissions from diesel cars on air quality, public health and eutrophication in Europe. Environ. Res. Lett. 12, 094017. https://doi.org/10.1088/1748-9326/aa8850

Karimi, S., Zadeh, M., Suul, J.A., 2020. Shore Charging for Plug-In Battery-Powered Ships: Power System Architecture, infrastructure, and Control. IEEE Electrification Mag. 8, 47–61. https://doi.org/10.1109/MELE.2020.3005699

Kawamura, K., Sriraj, P.S., Surat, H.R., Menninger, M., 2014. Analysis of Factors that Affect the Frequency of Truck Parking Violations in Urban Areas. Transp. Res. Rec. 2411, 20–26. https://doi.org/10.3141/2411-03

Kim, W., Wang, X. (Cara), 2021. Double parking in New York city: a comparison between commercial vehicles and passenger vehicles. Transportation. https://doi.org/10.1007/s11116-021-10212-5

King, R.D., Orhobor, O.I., Taylor, C.C., 2021. Cross-validation is safe to use. Nat. Mach. Intell. 3, 276–276. https://doi.org/10.1038/s42256-021-00332-z

Lin, C.-C., Peng, H., Grizzle, J.W., Kang, J.-M., 2003. Power management strategy for a parallel hybrid electric truck. IEEE Trans. Control Syst. Technol. 11, 839–849. https://doi.org/10.1109/TCST.2003.815606

Lust, E.E., Horton, W.T., Radermacher, R., 2009. A Review and Cost Comparison of Current Idle-Reduction Technology. Presented at the ASME 2008 Power Conference, American Society of Mechanical Engineers Digital Collection, pp. 715–726. https://doi.org/10.1115/POWER2008-60142

Lutsey, N., Brodrick, C.-J., Sperling, D., Oglesby, C., 2004. Heavy-Duty Truck Idling Characteristics: Results from a Nationwide Truck Survey. Transp. Res. Rec. 1880, 29–38. https://doi.org/10.3141/1880-04

MacDonald, C., Douglas, R., Tamayol, A., Bahrami, M., 2013. A Feasibility Study of Auxiliary HVAC Systems for Reducing Idling Time of Long Haul Trucks. Presented at the ASME 2012 Heat Transfer Summer Conference collocated with the ASME 2012 Fluids Engineering Division Summer Meeting and the ASME 2012 10th International Conference on Nanochannels, Microchannels, and Minichannels, American Society of Mechanical Engineers Digital Collection, pp. 323–330. https://doi.org/10.1115/HT2012-58353

Mahmud, S., Hernandez, S., Mitra, S., 2022. Impact of Truck Parking Facilities on Commercial and Industrial Land Values: A Spatial Hedonic Model. Transp. Res. Rec. 2676, 328–341. https://doi.org/10.1177/03611981211051355

Myles, A.J., Feudale, R.N., Liu, Y., Woody, N.A., Brown, S.D., 2004. An introduction to decision tree modeling. J. Chemom. 18, 275–285. https://doi.org/10.1002/cem.873

Nevland, E.A., Gingerich, K., Park, P.Y., 2020. A data-driven systematic approach for identifying and classifying long-haul truck parking locations. Transp. Policy 96, 48–59. https://doi.org/10.1016/j.tranpol.2020.04.003

O'Shea, K., Nash, R., 2015. An Introduction to Convolutional Neural Networks. https://doi.org/10.48550/arXiv.1511.08458

Park, S.S., Kozawa, K., Fruin, S., Mara, S., Hsu, Y.-K., Jakober, C., Winer, A., Herner, J., 2011. Emission Factors for High-Emitting Vehicles Based on On-Road Measurements of Individual Vehicle Exhaust with a Mobile Measurement Platform. J. Air Waste Manag. Assoc. 61, 1046–1056. https://doi.org/10.1080/10473289.2011.595981

Passchier, -Vermeer W., Passchier, W.F., 2000. Noise exposure and public health. Environ. Health Perspect. 108, 123–131. https://doi.org/10.1289/ehp.00108s1123

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, É., 2011. Scikit-learn: Machine Learning in Python. J. Mach. Learn. Res. 12, 2825–2830.

Peng, J., Liu, D., Parnell, J., Kessissoglou, N., 2019. Influence of translational vehicle dynamics on heavy vehicle noise emission. Sci. Total Environ. 689, 1358–1369. https://doi.org/10.1016/j.scitotenv.2019.06.426

Porter, C., Beagan, D.F., Koupal, J., Wayson, R.L., Welch, E., Cambridge Systematics, Inc., Eastern Research Group, Inc., AECOM, Energy and Environmental Research Associates, 2022. Addressing Truck Emissions and Noise at Truck Freight Bottlenecks Final Report (No. FHWA-HEP-22-026).

Profits, B.U., 2010. Long-Haul Truck Idling [WWW Document]. URL https://scholar.googleusercontent.com/scholar?q=cache:aLk2Tn4OgUUJ:scholar.google.com/+Long-Haul+Truck+Idling+Burns+Up+Profits&hl=en&as_sdt=0,5 (accessed 4.12.23).

Propper, R., Wong, P., Bui, S., Austin, J., Vance, W., Alvarado, Á., Croes, B., Luo, D., 2015. Ambient and Emission Trends of Toxic Air Contaminants in California. Environ. Sci. Technol. 49, 11329–11339. https://doi.org/10.1021/acs.est.5b02766

Rahman, S.M.A., Masjuki, H.H., Kalam, M.A., Abedin, M.J., Sanjid, A., Sajjad, H., 2013. Impact of idling on fuel consumption and exhaust emissions and available idle-reduction technologies for diesel vehicles – A review. Energy Convers. Manag. 74, 171–182. https://doi.org/10.1016/j.enconman.2013.05.019

Rodriguez-Galiano, V., Sanchez-Castillo, M., Chica-Olmo, M., Chica-Rivas, M., 2015. Machine learning predictive models for mineral prospectivity: An evaluation of neural networks, random forest, regression trees and support vector machines. Ore Geol. Rev. 71, 804–818. https://doi.org/10.1016/j.oregeorev.2015.01.001

Samsun, R.C., Krupp, C., Baltzer, S., Gnörich, B., Peters, R., Stolten, D., 2016. A battery-fuel cell hybrid auxiliary power unit for trucks: Analysis of direct and indirect hybrid configurations. Energy Convers. Manag. 127, 312–323. https://doi.org/10.1016/j.enconman.2016.09.025

SEC. 80.69.2. PARKING - COMMERCIAL VEHICLES AND TRAILERS. [WWW Document], n.d. URL https://codelibrary.amlegal.com/codes/los_angeles/latest/lamc/0-0-0-166594 (accessed 1.31.23).

Segal, M.R., 2004. Machine Learning Benchmarks and Random Forest Regression.

Skerlos, S.J., Winebrake, J.J., 2010. Targeting plug-in hybrid electric vehicle policies to increase social benefits. Energy Policy 38, 705–708. https://doi.org/10.1016/j.enpol.2009.11.014

Song, J., Zhao, Jiwen, Dong, F., Zhao, Jing, Qian, Z., Zhang, Q., 2018. A Novel Regression Modeling Method for PMSLM Structural Design Optimization Using a Distance-Weighted KNN Algorithm. IEEE Trans. Ind. Appl. 54, 4198–4206. https://doi.org/10.1109/TIA.2018.2836953

Su, J.G., Meng, Y.-Y., Chen, X., Molitor, J., Yue, D., Jerrett, M., 2020. Predicting differential improvements in annual pollutant concentrations and exposures for regulatory policy assessment. Environ. Int. 143, 105942. https://doi.org/10.1016/j.envint.2020.105942

Thaker, P., Gokhale, S., 2016. The impact of traffic-flow patterns on air quality in urban street canyons. Environ. Pollut. Barking Essex 1987 208, 161–169. https://doi.org/10.1016/j.envpol.2015.09.004

Truck Electronic Engine Controls [WWW Document], n.d. URL https://www.sae.org/publications/technical-papers/content/820907/ (accessed 2.1.23).

Turner, M.C., Andersen, Z.J., Baccarelli, A., Diver, W.R., Gapstur, S.M., Pope III, C.A., Prada, D., Samet, J., Thurston, G., Cohen, A., 2020. Outdoor air pollution and cancer: An overview of the current evidence and public health recommendations. CA. Cancer J. Clin. 70, 460–479. https://doi.org/10.3322/caac.21632

Vehicle Idle Reduction [WWW Document], 2011. . Sfenvironmentorg - Our Home Our City Our Planet. URL https://sfenvironment.org/article/vehicle-idle-reduction (accessed 1.30.23).

Wali, B., Khattak, A.J., Bozdogan, H., Kamrani, M., 2018. How is driving volatility related to intersection safety? A Bayesian heterogeneity-based analysis of instrumented vehicles data. Transp. Res. Part C Emerg. Technol. 92, 504–524. https://doi.org/10.1016/j.trc.2018.05.017

Węglarczyk, S., 2018. Kernel density estimation and its application. ITM Web Conf. 23, 00037. https://doi.org/10.1051/itmconf/20182300037

Weinberger, R.R., Millard-Ball, A., Hampshire, R.C., 2020. Parking search caused congestion: Where's all the fuss? Transp. Res. Part C Emerg. Technol. 120, 102781. https://doi.org/10.1016/j.trc.2020.102781

Zhang, S., Cheng, D., Deng, Z., Zong, M., Deng, X., 2018. A novel kNN algorithm with data-driven k parameter computation. Pattern Recognit. Lett., Special Issue on Pattern Discovery from Multi-Source Data (PDMSD) 109, 44–54. https://doi.org/10.1016/j.patrec.2017.09.036

Ziegler, T., 2015. Reducing Diesel Emissions, Improving Air Quality, and Promoting Environmental Justice in Southwest Detroit.