TECHNICAL REPORT DOCUMENTATION PAGE

1. Report No. FMRI-Y5R4- 21	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle MANAGEMENT OF SUPPLY CHAIN DISRUPTION OF FREIGHT NETWORK USING ADVANCED ALGORITHMS		5. Report Date: August, 2022 6. Performing Organization Code:
7. Author(s) Evangelos I. Kaisar, Aline Machado		8. Performing Organization Report No.
9. Performing Organization Name and Address Florida Atlantic University Freight Mobility Research Institute		10. Work Unit No. (TRAIS) 11. Contract or Grant No. 69A3551747120
12. Sponsoring Agency Name and Address Freight Mobility Research Institute (FMRI) Florida Atlantic University 777 Glades Rd., Bldg. 36, Boca Raton, FL 33431		13. Type of Report and Period: Final 14. Sponsoring Agency USDOT
15. Supplementary Notes		

16. Abstract

The growing complexity of global supply chains and the rising frequency of transportation disruptions have underscored the need for predictive tools that enhance freight network reliability. This study explores the integration of big data analytics and machine learning to support short-term travel time prediction and incident detection in freight transportation systems. Using the National Performance Management Research Data Set (NPMRDS) for the state of Florida, a comprehensive dataset was developed, incorporating temporal, spatial, and traffic condition features at 5-minute resolution intervals over a one-year period. Four machine learning models, XGBoost, Random Forest, Decision Tree, and Multilayer Perceptron (MLP), were evaluated for travel time regression and binary incident classification. The results demonstrate that tree-based ensemble models (particularly XGBoost and Random Forest) achieved superior predictive accuracy, with high R^2 scores for travel time prediction and high precision-recall performance for incident classification. Feature engineering techniques, including cyclical time encoding, significantly enhanced model performance. The findings confirm the value of leveraging open transportation datasets and machine learning for improving operational awareness and supporting resilient freight network design. This work contributes to the advancement of data-driven decision-making in transportation logistics, emphasizing the importance of real-time, interpretable, and scalable predictive models in mitigating risk and improving reliability in freight systems.

17. Key Words Supply Chain, Freight Network, Big Data	18. Distribution Statement No restrictions. This document is available to the public through Fmri.fau.edu		
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 41	22. Price

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

FREIGHT MOBILITY RESEARCH INSTITUTE

College of Engineering & Computer Science

Florida Atlantic University

Project ID: Y5R4-21

MANAGEMENT OF SUPPLY CHAIN DISRUPTION OF FREIGHT NETWORK USING ADVANCED ALGORITHMS

FINAL REPORT

by

Evangelos I. Kaisar, Aline Machado

Florida Atlantic University

777 Glades Rd. Boca Raton, FL 33431

E-mail: ekaisar@fau.edu

for

Freight Mobility Research Institute (FMRI)

777 Glades Rd.

Florida Atlantic University

Boca Raton, FL 33431

August, 2022

ACKNOWLEDGEMENTS

This project was funded by the Freight Mobility Research Institute (FMRI), one of the twenty TIER University Transportation Centers that were selected in this nationwide competition, by the Office of the Assistant Secretary for Research and Technology (OST-R), U.S. Department of Transportation (US DOT).

EXECUTIVE SUMMARY

This research investigates the application of big data and machine learning techniques to enhance freight transportation reliability, with a specific focus on travel time prediction and incident detection. As supply chains face increasing complexity and disruption, the ability to anticipate delays and adapt network operations is critical for maintaining efficiency and competitiveness. Using the National Performance Management Research Data Set (NPMRDS) for the state of Florida, the study explores short-term traffic prediction by leveraging both temporal and spatial features. A robust feature engineering process was applied, including cyclical encoding of time-of-day and segment-level traffic patterns. The analysis incorporates both freeway and arterial data and uses a full year (2022) of 5-minute resolution traffic records.

For travel time prediction, four machine learning models, XGBoost, Random Forest, Decision Tree, and Multilayer Perceptron (MLP), were evaluated. All models demonstrated high accuracy, with XGBoost achieving the best overall performance. Random Forest and Decision Tree models also performed well, offering strong accuracy with added interpretability. The MLP showed competitive results, though with slightly higher error metrics. For incident classification, Random Forest again outperformed other models with precision and recall of 0.96, demonstrating its robustness in identifying incidents while minimizing false alarms. XGBoost and MLP achieved moderate success, while the simpler Decision Tree model showed limitations in handling class imbalance.

The study also investigated incident trends by hour of the day, revealing distinct peaks during morning and afternoon rush hours, highlighting the importance of time-aware modeling in freight applications. Ultimately, this research underscores the value of integrating predictive analytics, open transportation datasets, and machine learning models to improve freight network performance and resilience. It also highlights the potential of data-driven decision-making tools to support real-time operations and long-term infrastructure planning in the face of uncertainty.

TABLE OF CONTENTS

EXECUTIVE SUMMARY .... **II**

CHAPTER 1 INTRODUCTION **6**

1.1 Background 6

1.2 Motivation and Problem Statement 7

1.3 Objectives of the Study 7

CHAPTER 2 LITERATURE REVIEW **8**

2.1 Introduction to the Chapter 8

CHAPTER 3 METHODOLOGY **12**

3.1 Dataset and Feature Engineering 12

3.2 Prediction Step and Horizon 14

3.3 Learning Algorithms 14

CHAPTER 4 DATA COLLECTION **15**

4.1 Case Study Area 15

4.2 Data Collection and Integration 17

4.3 NPMRDS – Traffic Performance Data 17

4.4 RITIS – Incident and Event Data 18

4.5 Dataset Construction and Filtering 18

4.6 NPMRDS – Traffic Performance Data 18

4.7 RITIS – Incident and Event Data 19

4.8 Dataset Preparation and Clustering 19

CHAPTER 5 RESULTS **23**

5.1 DESCRIPTIVE STATISTICS 23

5.2 Data Visualization and Insights 24

5.3 High-Risk Segment Identification 29

5.4 Travel Time Prediction Results 31

i. XGBoost 31

ii. Random Forest Regression 31

iii. Decision Tree Regression 32

iv. Neural Network Regression 32

5.5 Predicting Number of Accidents 33

i. XGBoost 33

ii. Random Forest 34

iii. Decision Tree 34
iv. Neural Network (MLP Classifier)..34
CHAPTER 6 DISCUSSION AND CONCLUSION..**36**
REFERENCES..**37**

LIST OF FIGURES

Figure 1: I-95 Segment in Broward County (Street View) 15
Figure 2: Traffic Segment Start and End Points Along Florida’s East Coast 16
Figure 3: Zoomed-in View of Segment Locations in Broward County 17
Figure 4: Elbow Method for Optimal Number of Clusters 20
Figure 5: Geographic Distribution of Clusters by Latitude and Longitude 20
Figure 6. Cyclical Time Encoding: Polar Plot 22
Figure 7: Hourly Distribution of Incident Occurances 24
Figure 8: Relationship Between Vehicle Speed and Travel Time, Colored by Incident Presence 25
Figure 9: Incident Rate by Day of Week 26
Figure 10: Speed Distribution by Incident Occurrence 27
Figure 11: Zoomed-In View of Segment Clusters and Incident Rates 29
Figure 12: Top 10 Segments with Highest Number of Accidents 30

LIST OF TABLES

Table 1. Final dataset structure 13
Table 2. Dataset aggregation 13
Table 3:. Descriptive Statistics of Model Input Features 23
Table 4. Speed Summary 28
Table 5: Travel Time (minutes) 28
Table 6: Top 10 Segments with the Highest Incident Rates 30
Table 7: . Model Performance Metrics for Travel Time Prediction 33
Table 8: Results Sumary 35

CHAPTER 1 INTRODUCTION

1.1 Background

Travel Time Reliability (TTR) refers to the consistency and predictability of travel times across different days and time periods (U.S. Department of Transportation, 2020). It is a foundational metric in transportation planning, used to assess how dependable a route is under normal and varying conditions. For the freight sector, TTR is particularly critical. Unexpected delays, caused by congestion, incidents, or weather, can lead to monetary losses, missed delivery windows, and disruptions in manufacturing or inventory cycles.

Freight shippers and carriers rely heavily on predictable travel times to remain cost-competitive. Unlike passenger travel, which is often more flexible, freight movement is tightly linked to logistics schedules and contractual delivery commitments. A delay of even a few minutes can have ripple effects across the supply chain. Studies consistently show that the freight industry places a high value on travel time reliability, often incorporating it into route selection, pricing strategies, and customer service guarantees (Wang et al., 2016). While travel time reliability is widely used across all transportation modes, its implications differ between freight and passenger vehicles. Freight operations tend to be more sensitive to unreliability due to fixed delivery windows, regulatory constraints (e.g., hours-of-service), and economic penalties for late arrivals. Passenger vehicles, by contrast, often allow for greater flexibility or buffer time (Golob & Recker, 2003).

Furthermore, TTR is closely linked to broader transportation system concepts such as resilience and vulnerability. A resilient freight network can maintain service levels even during disruptions, while a vulnerable one experiences severe performance degradation in the face of disturbances like road closures, accidents, or adverse weather. Poor travel time reliability magnifies a network's vulnerability by increasing uncertainty and reducing the ability to absorb shocks (Theofilatos & Yannis, 2014). Therefore, assessing TTR is essential for identifying supply chain risks and informing investment in mitigation strategies such as alternative routing, dynamic scheduling, and infrastructure improvements.

Accurate short-term prediction of traffic variables, including travel time reliability (TTR), is increasingly recognized as a critical capability for transportation systems. Even minor foresight into near-future traffic conditions can significantly enhance the competitiveness of freight transportation by enabling faster adaptation to network changes. Numerous studies across domains, from transportation to energy and finance, have explored short-term prediction methodologies, highlighting its broad relevance and potential. In freight logistics, predictive insights allow carriers to adjust routing and scheduling in real time, thereby minimizing delays and maintaining service reliability (Li et al., 2021). As data availability and computational capabilities expand, the integration of big data and machine learning (ML) into decision-making processes is rapidly becoming not just feasible but inevitable. Intelligent Transportation Systems (ITS) are at the forefront of this transformation, though many challenges remain before advanced predictive models become standard practice. In particular, the interpretability of ML models is a growing concern, especially when decisions derived from those models influence critical logistics operations. While interpretable models are often favored in operational contexts, they can come at the cost of predictive performance (Baryannis, Validi, et al., 2019).

One of the most significant barriers to widespread ML deployment in freight transportation is data availability. Freight data often resides within private logistics systems and is rarely accessible due to commercial sensitivity. In this context, the National Performance Management Research Data Set (NPMRDS) stands out as a valuable, publicly available resource. Managed by the Federal Highway Administration (FHWA), NPMRDS provides historical travel time data for passenger and freight traffic, aggregated at five-minute intervals across the U.S. national highway system (U.S. Department of Transportation, 2020). This dataset offers a unique opportunity to apply machine learning methods to understand and forecast travel time reliability without relying on proprietary data sources.

1.2 Motivation and Problem Statement

The main objective of this research is to explore the use of the NPMRDS dataset for short-term prediction of travel time reliability in both passenger and freight traffic. We perform a comparative experiment using truck and passenger vehicle data to evaluate differences in model performance across modes. A central focus is on analyzing how time aggregation levels and inclusion of upstream/downstream segments impact the accuracy of predictive models.

A key contribution of this work lies in its balanced attention to both predictive accuracy and model interpretability, a trade-off frequently encountered in machine learning applications for transportation. We evaluated a variety of algorithms, including Decision Tree Learning, which provides high interpretability and transparent decision rules but may sacrifice predictive power. In contrast, we also explored Deep Learning (DL) models, specifically multilayer perceptrons (MLPs), which can model complex, nonlinear relationships in high-dimensional data more effectively. Although DL models often function as "black boxes," their ability to learn hierarchical representations from large-scale spatiotemporal datasets such as the NPMRDS makes them well-suited for tasks like short-term travel time prediction. By comparing these approaches, this study sheds light on the trade-offs between transparency and accuracy, and demonstrates the potential of DL to enhance freight network performance when supported by rich and reliable data.

1.3 Objectives of the Study

- Correlation Analysis: Calculate correlation coefficients across different combinations of parameters (e.g., speed, travel time, reliability) and time ranges.
- Aggregation Sensitivity: Determine the optimal time aggregation level (e.g., 5-min, 15-min, 30-min) for different predictive modeling applications.
- Spatial Influence: Evaluate whether adding traffic state information from adjacent (upstream and downstream) segments enhances prediction accuracy.
- Transferability: Investigate the ability to train models on data-rich segments and successfully predict on data-scarce segments, assessing model robustness across the network.

CHAPTER 2 LITERATURE REVIEW

2.1 INTRODUCTION TO THE CHAPTER

In recent years, academia and management practitioners have paid close attention to big data analytics (BDA) capability. As computational power has increased, so has interest in the use of big data in supply chain management. Such solutions provide numerous benefits in the business activities of modern enterprises, particularly in the logistics sector. One of the most significant advantages of big data application in logistics is the ability to conduct high-performance analysis. Big data is changing the way supplier networks form, grow, spread into new markets, and mature over time. Numerous methodologies have been used in supply chain risk management (SCRM) research, ranging from qualitative ones like empirical studies and conceptual theories to quantitative ones like mathematical optimization, statistics, and simulation (Ghadge et al., 2012).

The recent experience with BD, according to (Waller & Fawcett, 2013), may help to explain some of the complex phenomena and unresolved issues in logistics and supply chain management. The main goal of this special issue (SI) is to give the logistics and supply chain management community a significant opportunity to influence practice through fundamental research on how organizations can use BDA capability to provide logistics and supply chain insights. Using robust models and case studies, the study by Wamba, Gunasekaran, Papadopoulos, and Ngai 2018, explores how big data can be used to address actual supply chain issues. The study of supply chain risk management (SCRM), which includes risk identification, assessment, mitigation, and monitoring, is a rapidly expanding area of study. There has been a lot of emphasis among researchers on locating, minimizing, and managing risks that have an impact on the supply chain (SC). The SC managers have started to concentrate on making decisions based on a variety of data sources to more accurately predict the uncertainties and create a proactive and predictive intelligent risk management mechanism. Using data-driven AI techniques and relying on the synergy between AI and supply chain specialists, (Baryannis, Dani, et al., 2019)first proposed a framework for supply chain risk prediction. They then implement and apply the framework to the problem of forecasting delivery delays in a real-world multi-tier manufacturing supply chain in order to examine the trade-off between prediction performance and interpretability.

(Baryannis, Validi, et al., 2019) conducted research on the different definitions and categories of supply chain risk and related concepts like uncertainty. Then, they conducted a mapping study to classify the body of literature in accordance with the AI methodology employed, which ranges from mathematical programming to machine learning and big data analytics, as well as the specific SCRM task they address, such as identification, assessment, or response. In a study, (Chandra & Al-Deek, 2009) identified the various AI and ML methods used in various phases of SCRM and presented a systematic and descriptive review of the literature. Additionally, based on the AI technique employed, it examines the various SC risk categories as well as the published articles. This analysis focuses on SCRM-related research articles from three academic databases. The creation of automated screening techniques with a mathematical foundation that can be incorporated into supply chain risk management was proposed by (Zage et al., 2013).

Travel time reliability (TTR) in freight transportation refers to the consistency and predictability of travel times for cargo-carrying vehicles, such as trucks, across different trips, routes, and time periods (Tian Guangqiang, 2019). TTR is a critical performance measure that influences the efficiency, cost-effectiveness, and operational dependability of freight systems. Reliable travel times support timely deliveries, reduce inventory costs, and enhance the overall responsiveness of supply chains. Consequently, researchers have increasingly leveraged historical travel time data to analyze variability patterns and develop predictive models that aid in real-time decision-making, route planning, and schedule optimization.

Short-term traffic prediction is a well-established area of research in transportation, particularly due to its utility in mitigating delay impacts and supporting proactive freight operations. Early studies employed classical statistical approaches to estimate travel time variability. For example, (Figliozzi et al., 2011) used distributional analyses and statistical modeling to understand delay probabilities under recurrent and non-recurrent congestion scenarios. (Cedillo-Campos et al., 2019) applied geographically weighted regression (GWR) to account for spatial heterogeneity in travel time prediction, demonstrating that even non-machine learning methods can provide valuable insights into localized travel behavior.

With the advent of high-resolution probe data and increased computing power, more recent studies have explored the application of advanced machine learning (ML) and deep learning (DL) models to short-term travel time prediction. These methods offer significant advantages in capturing nonlinear relationships and complex spatiotemporal patterns. Johnson et al. (2022) employed Gradient Boosting Machines (GBM)—a powerful ensemble learning approach that includes methods such as bagging, random forests, and boosting. While GBM models have demonstrated high predictive accuracy, enhanced versions can be computationally intensive and less interpretable, posing challenges for real-time deployment in operational freight systems.

Other studies have leveraged non-parametric and data-driven approaches. For instance, (Habtemichael & Cetin, 2016) proposed an improved K-Nearest Neighbor (K-NN) algorithm that identifies similar historical traffic patterns for short-term forecasting. This method is particularly useful when modeling contextually similar but spatially distinct segments, and when the dataset lacks strong parametric structure. A common challenge in transportation modeling is data insufficiency, particularly on lower-volume links such as rural roads or local streets. This issue hinders the training of complex models that require abundant historical data. To address this, (Li et al., 2021) proposed a transfer learning framework that applies knowledge learned from data-rich segments (e.g., freeways) to data-scarce ones. Their approach utilized Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks—specialized forms of Deep Neural Networks (DNNs) capable of learning long-range temporal dependencies in sequential data. LSTMs have proven particularly effective in modeling time-series data with irregular patterns and are increasingly applied in transportation forecasting problems (Ma et al., 2015).

As previously discussed, the concepts of reliability, resilience, and vulnerability in transportation systems are deeply interconnected. Travel time unreliability is often associated with non-recurring events, such as traffic incidents, construction zones, weather disruptions, or large-scale events, that deviate from typical traffic patterns. These unpredictable disruptions introduce significant uncertainty into the network and make short-term prediction even more critical during abnormal operating conditions (Shams et al., 2017).

In such scenarios, traditional single-model predictions may struggle to generalize across the full range of traffic states. To address this, prediction fusion techniques have emerged as a promising approach. These techniques are built on the principle that different models may perform better under different conditions, some during stable traffic flow, others during highly variable or disrupted states. (Guo et al., 2018) proposed a fusion-based prediction framework that combines multiple stand-alone models to improve overall accuracy. Their study tested three fusion strategies—simple averaging, weighted averaging, and k-Nearest Neighbor (k-NN), applied across three machine learning models: Neural Networks (NN), Support Vector Regression (SVR), and Random Forests (RF). The results confirmed that fusion methods can outperform individual models, particularly in dynamic or irregular traffic conditions, by leveraging the complementary strengths of different algorithms (Pournader et al., 2021).

Beyond transportation-specific applications, researchers in the supply chain risk management (SCRM) domain have also explored the integration of AI-driven prediction models for reliability and disruption forecasting. (Chandra & Al-Deek, 2009) highlighted the untapped potential of artificial intelligence (AI) techniques in predicting and mitigating supply chain risks. They proposed a predictive framework that blends data-driven AI methods with domain expertise from supply chain professionals. Also, (Lee & Wong, 2021) applied to a real-world case to forecast delivery delays across multiple tiers of a supply chain. A key insight from this work was the importance of balancing prediction performance with model interpretability, especially when model outputs are used to support high-stakes operational decisions.

The research question addressed in this study aligns closely with the concept of freight fluidity, a framework developed by Transport Canada and increasingly adopted in the United States. Freight fluidity indicators typically measure both total transit time and travel time reliability along freight corridors, offering a comprehensive view of freight network performance. In the U.S., the Federal Highway Administration (FHWA), in collaboration with Texas A&M Transportation Institute, has been working on the National Freight Fluidity Monitoring Program Implementation. Within this context, travel time reliability stands out as one of the most critical components for assessing freight corridor performance and designing efficient multimodal freight systems (U.S. Department of Transportation, 2020).

From a methodological standpoint, (Cedillo-Campos et al., 2019) examined different probability distributions to characterize travel time data, using statistical hypothesis testing to identify the best-fitting models. Their work illustrates the importance of understanding the underlying distributional properties of travel times when modeling reliability.

In the private sector, organizations have also recognized the strategic value of travel time prediction. (Sapankevych & Sankar, 2009; Zeng & Zhang, 2013) developed a machine learning–based framework to forecast air freight transit time delays, enabling proactive mitigation strategies. Their model leveraged 58 internal operational parameters to predict whether the average daily transit time for a given air freight lane would increase or decrease—up to one week in advance. This example underscores the growing use of predictive analytics and machine learning in commercial freight logistics, where timely and reliable deliveries are essential for maintaining service levels and customer satisfaction.

Despite the growing emphasis on predictive modeling and performance monitoring, the National Performance Management Research Data Set (NPMRDS) has not yet seen widespread use in freight-specific academic research. One exception is, whose research (Mishra et al., 2017) explored the application of NPMRDS data for calculating performance metrics. Also, (Shang et al., 2017) developed an optimal classifier to automatically detect congestion across roadway segments, applying a range of machine learning models including Naïve Bayes, K-Nearest Neighbor (K-NN), Decision Trees, and Support Vector Machines (SVM) (Tirkolaee et al., 2021). His work demonstrated the potential of NPMRDS for generating freight reliability indicators, though broader adoption in freight applications remains limited.

In summary, the field of freight travel time reliability is dynamic and continuously evolving. Advances in data availability, statistical modeling, and machine learning are opening new pathways for both academic research and practical implementation. Continued collaboration between researchers, transportation agencies, and industry partners is essential for addressing the complexities of freight movement and supporting the development of more efficient, fluid, and resilient freight transportation systems.

3.1 Dataset and Feature Engineering

This study utilizes data from the National Performance Management Research Data Set (NPMRDS), a comprehensive dataset developed using time and location information collected from probe vehicles. The NPMRDS provides speed and travel time measurements aggregated at 5-minute intervals, enabling fine-grained temporal analysis of traffic performance across the national highway system.

The spatial resolution of the NPMRDS is defined by Traffic Message Channel (TMC) codes. Each TMC represents a unique, directional road segment, with lengths typically ranging from 0.5 to 1 mile in urban and suburban areas, and up to 5 to 10 miles in rural regions. These TMC segments form the foundational spatial units for data aggregation and analysis. The dataset has been made available since 2013 and includes both freeways and arterials, offering a wide spatial and functional coverage (FHWA, 2020).

For this study, a subset of the NPMRDS for the state of Florida was extracted. As shown in the Regional Integrated Transportation Information System (RITIS) platform, data coverage across Florida varies by location and road class. To ensure data quality and consistency, this analysis focuses on Interstates, the Florida Turnpike, and U.S. Routes—road classes with high probe vehicle coverage. Including U.S. Routes adds diversity to the dataset by incorporating non-freeway corridors, thus enhancing the robustness of the analysis.

Given the focus on temporal prediction rather than network connectivity, the geospatial continuity of TMC segments was not considered essential for this analysis. Each dataset—corresponding to a specific road class or corridor—was downloaded and processed independently. Road classification was retained as a new feature in the dataset.

The original NPMRDS records contain the following fields: TMC code, timestamp, instantaneous speed, average speed, reference speed, travel time (in seconds), and data density. The data density attribute, which originally uses a qualitative scale from A to D, was converted into a numerical scale to support compatibility with machine learning algorithms.

In total, the dataset includes 17,820 unique TMC segments across the state. Using the segment lengths available from the NPMRDS metadata, a Travel Time Index (TTI) was calculated for each observation. This index serves as a derived feature to quantify the relative difference between observed travel time and reference (free-flow) conditions, and is particularly useful for identifying recurring and non-recurring congestion patterns.

When applying machine learning, feature engineering plays a critical role in the model's ability to learn meaningful patterns from the data. In this study, the original timestamp field was decomposed into several temporal features to enhance the model's understanding of time-based variation. These derived features include: hour of the day, day of the week, month, day of the year, and a local holiday indicator based on the U.S. federal holiday calendar. These attributes are commonly used in traffic forecasting models to capture cyclical and seasonal trends in travel behavior.

To supplement the raw speed measurements, an additional feature was created: the average speed by hour for each TMC segment. This feature captures the typical speed conditions for each segment during each hour of the day, offering a benchmark for identifying deviations from expected performance. Due to the large size of the dataset, computing this feature required significant processing time.

The shape of the final dataset, including the number of features and records, is summarized in Table 1. The dataset covers the full calendar year of 2022, spanning from January to December, with the most granular data available at 5-minute intervals. To assess the impact of temporal resolution on model performance, the dataset was aggregated at multiple time intervals. The aggregation began at the base level of 5 minutes and was progressively increased by a factor of two for each subsequent level (e.g., 10 minutes, 15 minutes, 30 minutes, etc.).

Table 1. Final dataset structure

Column	Dtype	mean	std	min	max
tmc_code	string	42.0088	18.9866	3	99
speed	float64	40.9005	18.8395	3	99
average_speed	float64	60.0008	11.8108	3	99
reference_speed	float64	103.663	145.881	0.18	16290.9
travel_time_seconds	float64	1.15827	1.74524	0.00345	36.8823
data_density_num_asc	Int64	1.66811	0.788199	0.0997	32.9996
miles	float64	11.857	6.06333	1	3
TTI	float64	2.90181	1.95484	0	6
hour_of_day	int32	4	0	4	4
weekday	int32	105.512	8.51825	91	120
month	int32	42.0088	17.5335	3	94
day_of_year	int32	mean	std	min	max
is_holiday	bool	42.0088	18.9866	3	99
avg_speed_by_hour_segment	float64	40.9005	18.8395	3	99

The number of records corresponding to each aggregation level is presented in Table 2, providing a view of how temporal granularity affects data volume and, by extension, model complexity and training time.

Table 2. Dataset aggregation

Aggregation	Number of records
5 minutes	128572480
10 minutes	64286240
15 minutes	32143120
30 minutes	16071560
1 hour	8035780
2 hours	4017890
4 hours	2008945
8 hours	1004472
16 hours	502236
24 hours	251118

3.2 Prediction Step and Horizon

In this study, the prediction step corresponds to the short-term forecasting of travel time and incident occurrence using minimum 5-minute time intervals, consistent with the temporal resolution of the NPMRDS dataset. The prediction horizon spans a single future time step (i.e., 5 minutes ahead), allowing the models to capture near-real-time variations in traffic conditions. This short-term horizon is particularly relevant for operational freight decision-making, such as dynamic routing, congestion avoidance, and incident response. The focus on immediate prediction ensures that the models remain responsive to rapidly changing traffic patterns while maintaining computational efficiency for real-time applications.

3.3 Learning Algorithms

To evaluate the effectiveness of machine learning in short-term freight travel time and incident prediction, four supervised learning algorithms were selected:

- XGBoost,
- Random Forest,
- Decision Tree,
- Multilayer Perceptron (MLP).

These models represent a mix of ensemble methods, interpretable classifiers, and deep learning architectures, allowing for a balanced comparison between accuracy and explainability. Each algorithm was trained on engineered features derived from the NPMRDS dataset, including temporal, spatial, and traffic-related variables. Hyperparameter tuning was conducted to optimize model performance, and each model was evaluated using standard metrics such as R^2, RMSE, MAE for regression, and precision, recall, and F1-score for classification. This diverse model selection helps identify trade-offs between performance and operational interpretability in real-world freight applications.

4.1 CASE STUDY AREA

For this study, we selected a group of traffic segments located in southeast Florida, focusing specifically on Broward and Miami-Dade Counties. This region was chosen due to its high freight traffic volume, dense roadway infrastructure, and frequent traffic incidents, making it a meaningful area for analyzing both travel time and accident prediction.

We used latitude and longitude coordinates from the NPMRDS dataset to map the start and end points of each traffic segment. Green crosses (+) represent segment start points, and red dots (×) represent end points. These visualizations were essential in confirming the accuracy of our spatial data and ensuring proper alignment with incident records from RITIS.

To provide context for the selected corridor, Figure 1 shows a real-world image of I-95 in Broward County, near Cypress Creek Road. This figure presents a real-world street view of I-95 in Broward County, one of the busiest freight corridors in the region. It highlights physical roadway characteristics such as multiple lanes, concrete barriers, and exit signage. These features help visualize the traffic environment and are important when analyzing accident risk and traffic flow.

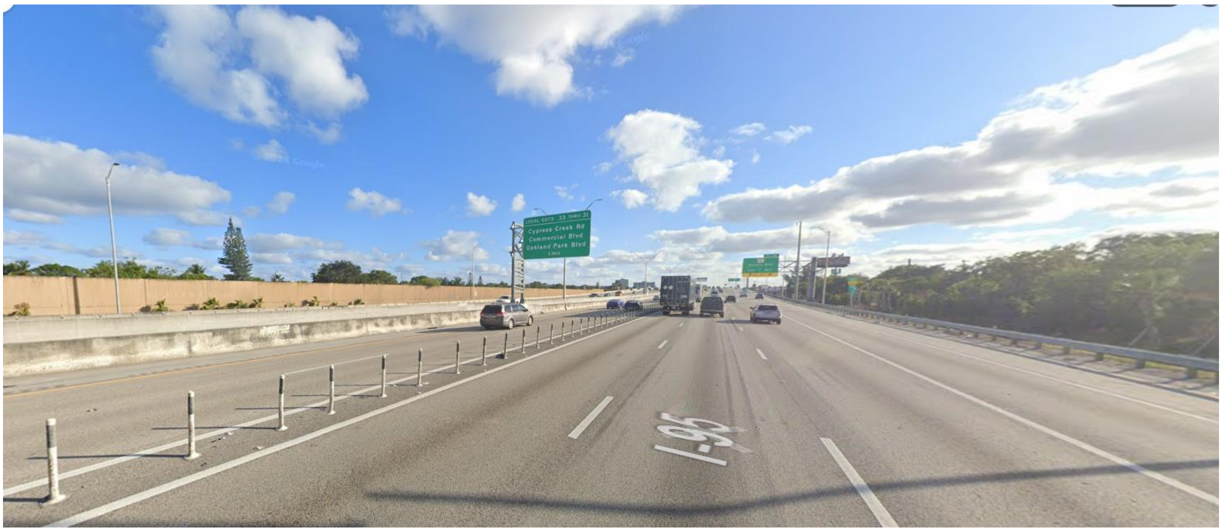

Figure 1: I-95 Segment in Broward County (Street View)

Moving beyond the street view, Figure 2 presents a map of the traffic segment start and end points distributed along Florida's east coast. The mapped segments follow the I-95 corridor and offer a visual overview of the area studied. This figure highlights the general layout and directional coverage of the selected roadways.

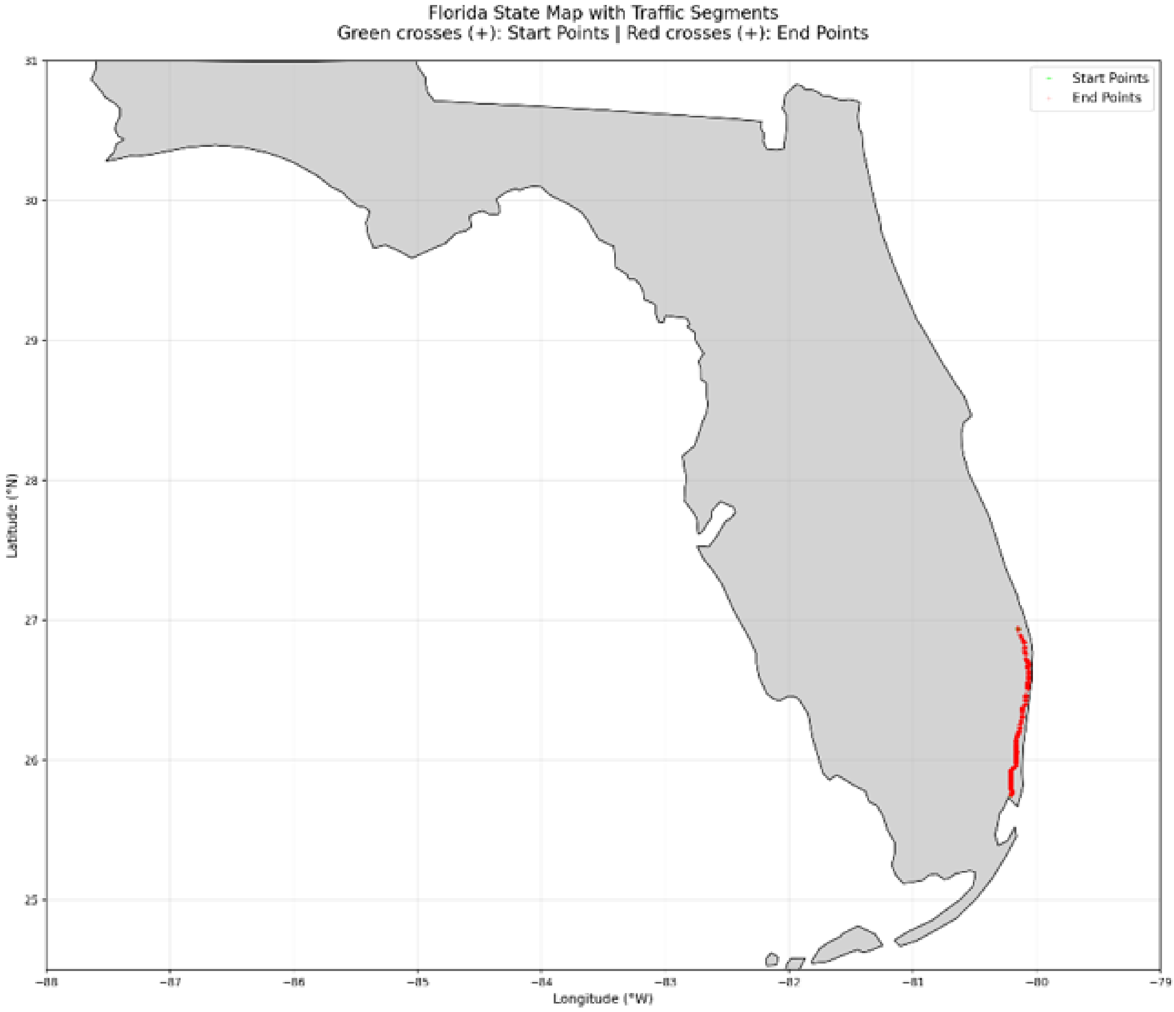

Figure 2: Traffic Segment Start and End Points Along Florida's East Coast

To focus on the core study region, Figure 3 provides a zoomed-in map of Broward County, clearly showing the geographic layout of the traffic segments analyzed. In this figure, start and end points are again marked, and a color scale is applied to show different segment identifiers. This view was particularly useful for verifying the density of observations and organizing the data for machine learning.

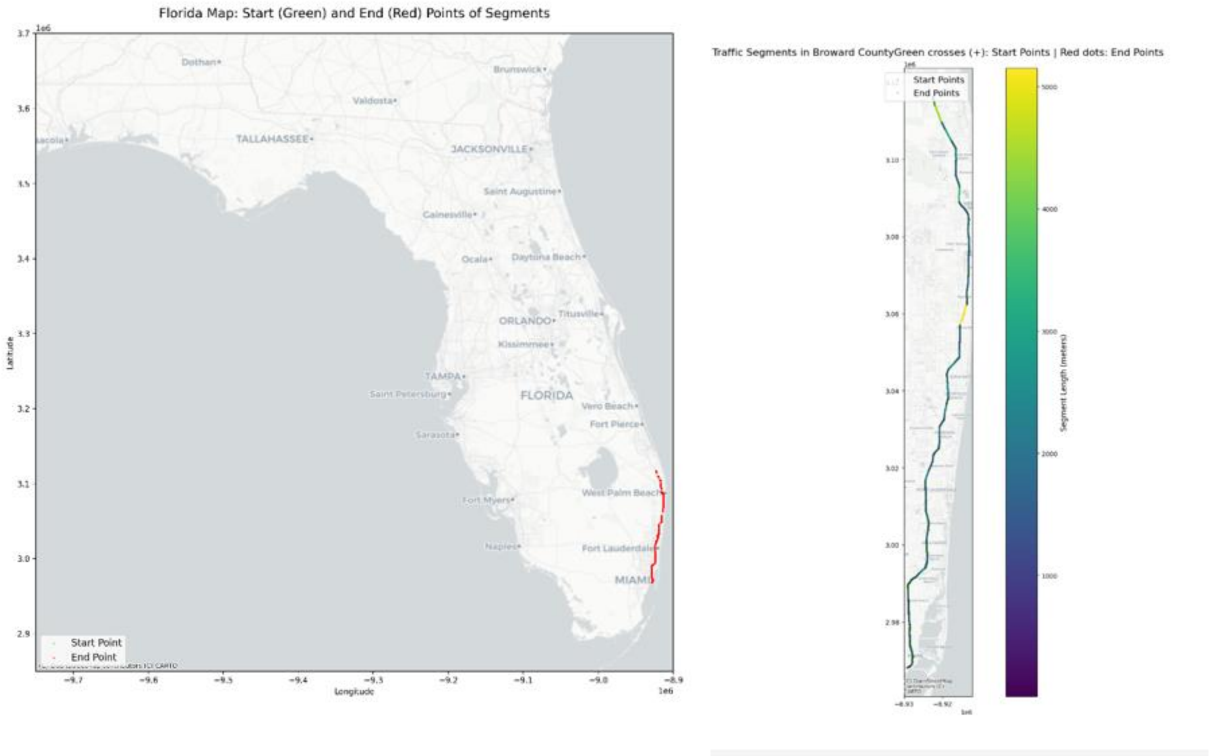

Figure 3: Zoomed-in View of Segment Locations in Broward County

By narrowing our analysis to this specific corridor and focusing on five representative days, we created a clean, structured dataset that is well-suited for developing predictive models. The selected region and visual validation process helped ensure a high level of data quality and relevance for both travel time estimation and incident classification tasks.

4.2 Data Collection and Integration

This study relies on two main data sources: the National Performance Management Research Data Set (NPMRDS) and the Regional Integrated Transportation Information System (RITIS). These sources were combined to form a comprehensive dataset for analyzing and predicting both accident occurrence and travel time performance on freight corridors in Florida.

4.3 NPMRDS – Traffic Performance Data

The NPMRDS is a high-resolution traffic dataset provided by the U.S. Federal Highway Administration. It captures probe-based travel time and speed data reported at 5-minute intervals. For this study, the Florida dataset was selected, with a focus on freight-significant corridors, including the Interstate system, Florida Turnpike, and U.S. Routes.

Each record in the raw NPMRDS data includes:

- tstamp: timestamp of the 5-minute window
- weekday: day of the week
- speed and reference_speed: actual and speed limit
- travel_time_minutes: observed travel time
- confidence_factor: quality of the data
- tmc_direction, and segment coordinates (start and end lat/lon)

These variables were used to understand traffic behavior and construct features for short-term travel time prediction.

4.4 RITIS – Incident and Event Data

To add accident-related outcomes, incident data was collected from RITIS, a powerful transportation data platform developed by the CATT Lab at the University of Maryland. RITIS provides real-time and archived data on roadway events from a variety of sources including state departments of transportation, law enforcement, and traffic management systems.

Using the RITIS Events and Incidents module, the research team downloaded incident records for Florida roadways in 2019. This dataset includes detailed information such as:

- Type of incident (e.g., crash, breakdown, road closure)
- Geographic coordinates of the event
- Start and end time of the event

These records were then joined with the NPMRDS traffic data using Python, based on spatial proximity (matching incident location to segment coordinates) and temporal alignment (matching the incident time window to the 5-minute timestamp).

4.5 Dataset Construction and Filtering

Due to the high volume of data available, the scope of this analysis was narrowed to a focused window: five representative days from 2019. This subset was selected to allow detailed analysis of both accident prediction and travel time modeling, specifically during periods of active freight movement.

The merging process was done programmatically using Python, where:

- Traffic and incident data were joined
- New features were created such as incident (a binary flag indicating presence of an incident)
- Each row was cleaned and structured to be used in machine learning models

The final dataset contained 357,371 rows, each representing a unique 5-minute observation on a roadway segment. This enriched dataset includes both traffic characteristics and incident labels, enabling it to be used for supervised learning tasks like:

- Classifying the likelihood of an accident
- Predicting travel time using environmental and traffic features

This structured integration of two independent sources, one describing normal operating conditions and the other capturing unplanned disruptions, allows for a robust and realistic modeling of real-world transportation behavior.

4.6 NPMRDS – Traffic Performance Data

The NPMRDS is a high-resolution traffic dataset provided by the U.S. Federal Highway Administration. It captures probe-based travel time and speed data reported at **5-minute intervals**. For this study, the Florida dataset was selected, with a focus on **freight-significant corridors**, including the **Interstate system, Florida Turnpike, and U.S. Routes**.

Each record in the raw NPMRDS data includes:

- tstamp: timestamp of the 5-minute window
- weekday: day of the week
- speed and reference_speed: actual and speed limit
- travel_time_minutes: observed travel time
- confidence_factor: quality of the data
- tmc_direction, and segment coordinates (start and end lat/lon)

These variables were used to understand traffic behavior and construct features for short-term travel time prediction.

4.7 RITIS – Incident and Event Data

To add accident-related outcomes, incident data was collected from RITIS, a powerful transportation data platform developed by the CATT Lab at the University of Maryland. RITIS provides real-time and archived data on roadway events from a variety of sources including state departments of transportation, law enforcement, and traffic management systems.

Using the RITIS Events and Incidents module, the research team downloaded incident records for Florida roadways in 2019. This dataset includes detailed information such as:

- Type of incident (e.g., crash, breakdown, road closure)
- Geographic coordinates of the event
- Start and end time of the event

These records were then joined with the NPMRDS traffic data using Python, based on spatial proximity (matching incident location to segment coordinates) and temporal alignment (matching the incident time window to the 5-minute timestamp).

4.8 Dataset Preparation and Clustering

Before applying machine learning models, we first prepared the dataset by organizing the raw data and engineering useful features. The original dataset included tstamp, weekday, speed, reference_speed, travel_time_minutes, confidence_factor, tmc_direction, and geographic coordinates such as tmc_start_latitude, tmc_start_longitude, tmc_end_latitude, and tmc_end_longitude, along with a binary incident indicator.

To enhance the model's understanding of spatial patterns, we analyzed the location of each segment. A total of 247 unique segments were identified in the case study. These were clustered based on their geographic coordinates and incident rates using the K-Means clustering algorithm. To determine the optimal number of clusters (k), we applied the elbow method, which plots the total inertia (within-cluster sum of squares) against different values of k.

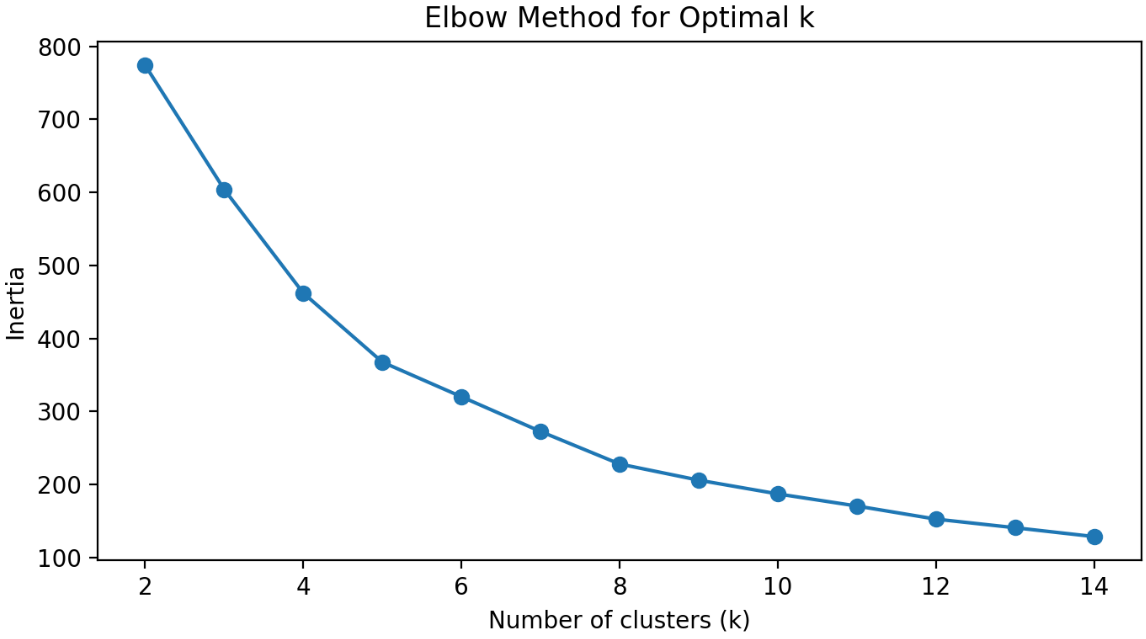

Figure 4: Elbow Method for Optimal Number of Clusters

This figure shows that the drop in inertia starts to level off around k = 6, which we selected as the optimal number of clusters. This choice balances simplicity and segmentation quality.

Once k = 6 was selected, we applied K-Means clustering and assigned each segment a location_accident_cluster label. These clusters grouped segments that were geographically close and had similar incident patterns.

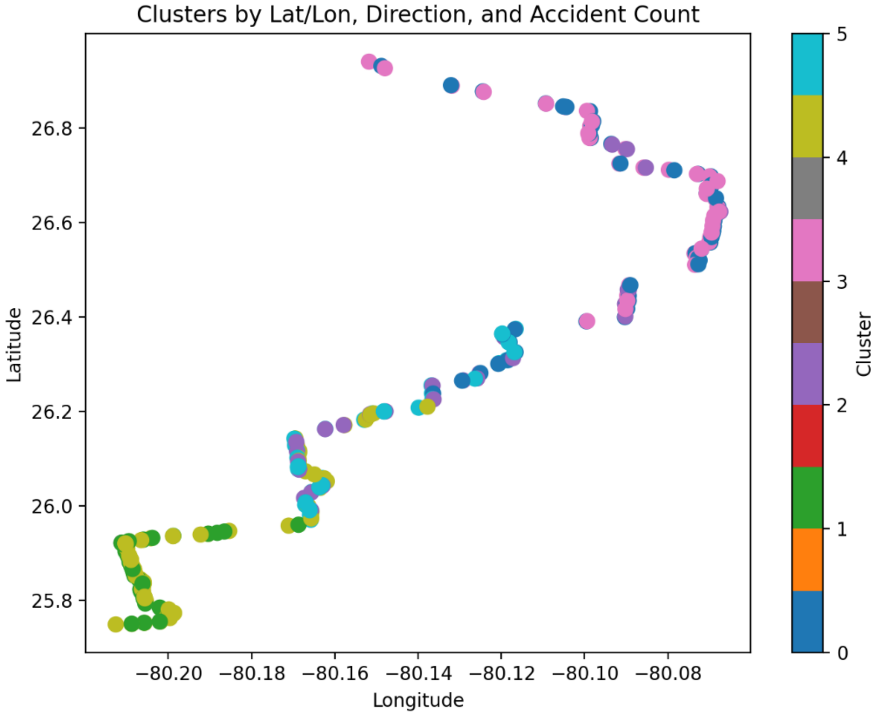

Figure 5: Geographic Distribution of Clusters by Latitude and Longitude

Figure 5 shows the spatial distribution of all segments across the study area, colored by cluster. Each color represents a unique cluster (Cluster 0 to Cluster 5). The size of each point indicates the incident rate in that segment, highlighting the relative risk levels across different locations.

Each cluster was then analyzed to understand its characteristics:

Cluster 0:

- Centered around latitude 26.54, longitude -80.09.
- Contains 51 unique segments, mostly in one direction.
- Incident rate: 5.8%.
- Moderate speeds and travel times.
- Represents a moderately risky, geographically central area.

Cluster 1:

- Centered at latitude 25.86, longitude -80.20.
- 44 unique segments, mostly in one direction.
- Lowest incident rate: 3.9%.
- Lower speeds, shorter travel times.
- Likely a less risky, possibly urban or less trafficked area.

Cluster 2:

- Centered at latitude 26.27, longitude -80.13.
- 44 unique segments, mostly in one direction.
- Highest incident rate: 36.6%.
- Moderate speeds, higher accident risk.
- This is the most accident-prone cluster, possibly due to complex intersections or high traffic.

Cluster 3:

- Centered at latitude 26.64, longitude -80.08.
- 44 unique segments, mostly in one direction.
- Incident rate: 7.1%.
- Moderate speeds and travel times.
- Represents a slightly above-average risk area.

Cluster 4:

- Centered at latitude 25.94, longitude -80.19.
- 44 unique segments, mostly in one direction.
- Incident rate: 6.3%.
- Lower speeds, moderate travel times.
- Slightly above-average risk, possibly suburban or mixed-use.

Cluster 5:

- Centered at latitude 26.18, longitude -80.15.

- 44 unique segments, mostly in one direction.
- Incident rate: 7.7%.
- Moderate speeds, moderate travel times.
- Slightly above-average risk, possibly a transition zone between urban and suburban.

Also, Since the time of day plays a critical role in traffic behavior and accident likelihood, we transformed the raw timestamp into a cyclical format using sine and cosine functions. This approach preserves the periodic nature of time (e.g., 23:59 is close to 00:00), which traditional numeric encoding cannot represent effectively.

We generated two new features:

- time_sin = sin(2π × minutes_since_midnight / 1440)
- time_cos = cos(2π × minutes_since_midnight / 1440)

These variables capture daily traffic patterns and make it easier for machine learning models to learn from time-related trends. For example, in the sample shown, the time is near midnight (0 minutes), resulting in time_sin = 0 and time_cos = 1, which aligns with the expected cyclical behavior.

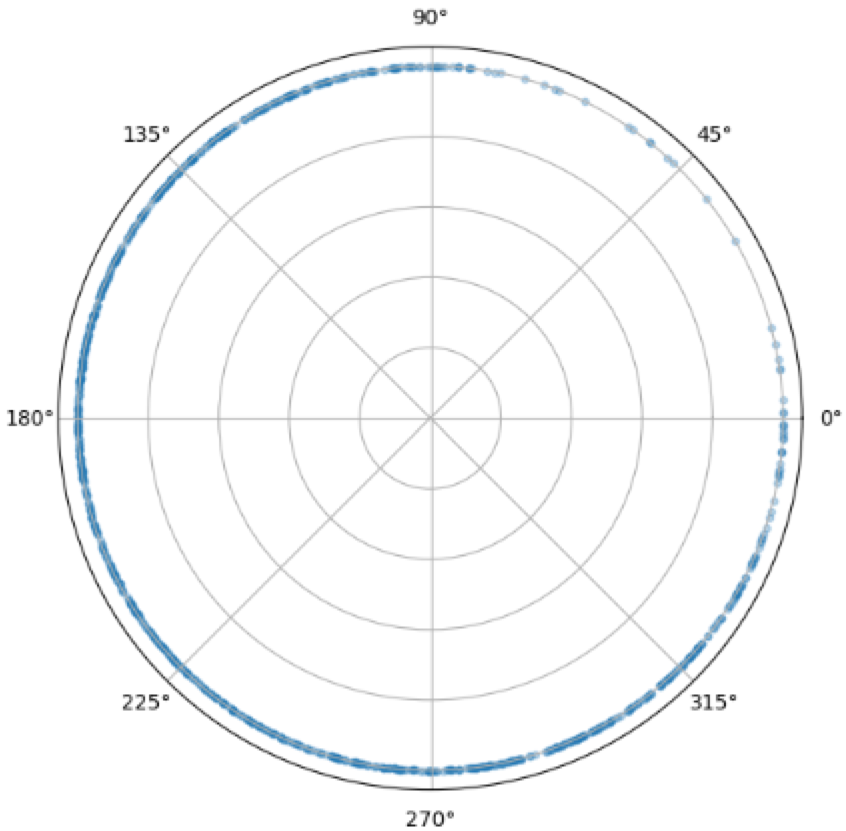

Figure 6. Cyclical Time Encoding: Polar Plot

5.1 DESCRIPTIVE STATISTICS

Table 3 presents a statistical overview of all numerical variables in the dataset (excluding the timestamp). The average vehicle speed across all segments is approximately 58.87 mph, while the reference speed is slightly higher at 60.46 mph, indicating that most traffic moves close to expected flow conditions. The incident rate is relatively low, with a mean of 0.10, suggesting that about 10% of the records are associated with incidents. The travel time ranges from 0.06 to 36.06 minutes, although the majority of observations fall below 1 minute. Time features (time_sin, time_cos) and cluster variables (cluster_0 to cluster_5) are included in binary or cyclical formats, making them suitable for model input. The table confirms that the dataset is clean, well-structured, and statistically sound for further analysis.

Table 3:. Descriptive Statistics of Model Input Features

Feature	Count	Mean	Std	Min	25%	50%	75%	Max
speed	357,371	58.87	15.12	2.00	54.00	63.00	69.00	106.00
reference_speed	357,371	60.46	5.12	52.00	55.00	65.00	65.00	70.00
confidence_factor	357,371	0.98	0.012	0.70	0.98	0.98	0.98	0.98
weekday	357,371	5.12	1.38	3.00	4.00	5.00	6.00	7.00
time_sin	357,371	-0.18	0.70	-1.00	-0.85	-0.37	0.51	1.00
time_cos	357,371	-0.31	0.61	-1.00	-0.85	-0.48	0.11	1.00
cluster_0	357,371	0.21	0.41	0.00	0.00	0.00	0.00	1.00
cluster_1	357,371	0.16	0.37	0.00	0.00	0.00	0.00	1.00
cluster_2	357,371	0.14	0.35	0.00	0.00	0.00	0.00	1.00
cluster_3	357,371	0.14	0.34	0.00	0.00	0.00	0.00	1.00
cluster_4	357,371	0.22	0.41	0.00	0.00	0.00	0.00	1.00
cluster_5	357,371	0.13	0.31	0.00	0.00	0.00	0.00	1.00
incident	357,371	0.10	0.31	0.00	0.00	0.00	0.00	1.00
travel_time_minutes	357,371	0.75	0.63	0.06	0.45	0.63	0.86	36.06

5.2 Data Visualization and Insights

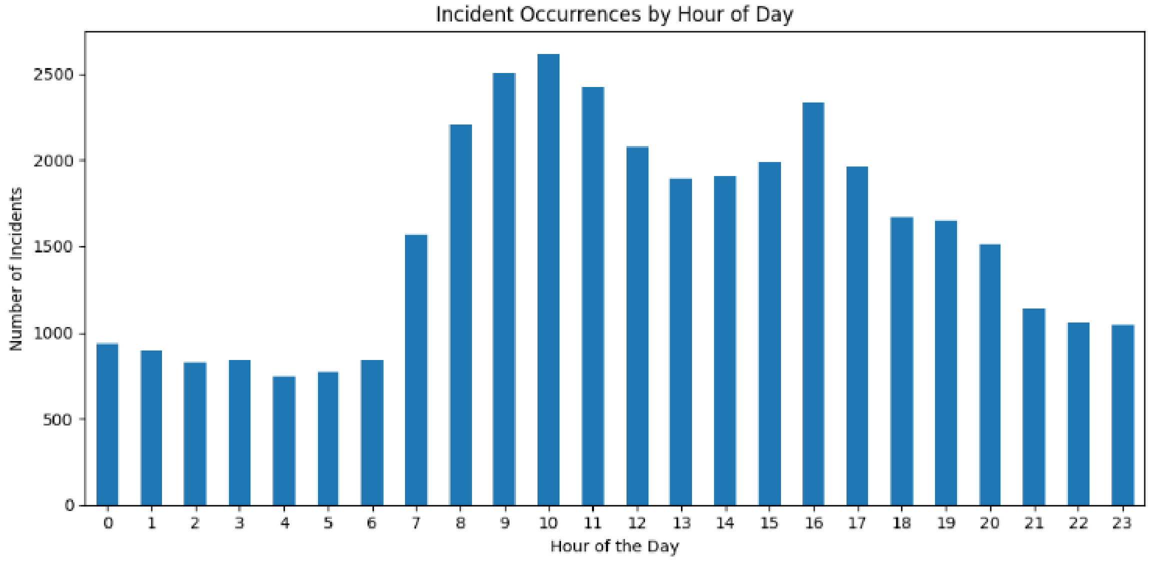

Figure 7: Hourly Distribution of Incident Occurances

Figure 7 presents the distribution of incident occurances by hour of the day, offering valuable insight into when accidents are most likely to occur. The visualization reveals several notable patterns:

- Morning hours (7 AM to 11 AM) show a sharp rise in incident frequency, peaking between 9 AM and 11 AM. This aligns with morning rush hour, when traffic volumes are high and congestion or driver stress may increase the likelihood of incidents.
- A second, though less pronounced peak occurs around 4 PM to 5 PM, corresponding to the afternoon/evening commute period.
- Incident occurrences gradually decline after 6 PM and remain relatively low throughout the late-night and early morning hours (12 AM to 6 AM), likely due to reduced traffic volumes during those periods.
- The lowest frequency of incidents appears between 3 AM and 5 AM, which is consistent with minimal roadway activity.

Overall, the temporal distribution suggests a strong correlation between incident occurrence and peak traffic periods, particularly during the morning commute. These insights can inform time-of-day-specific traffic management strategies or targeted deployment of incident response resources.

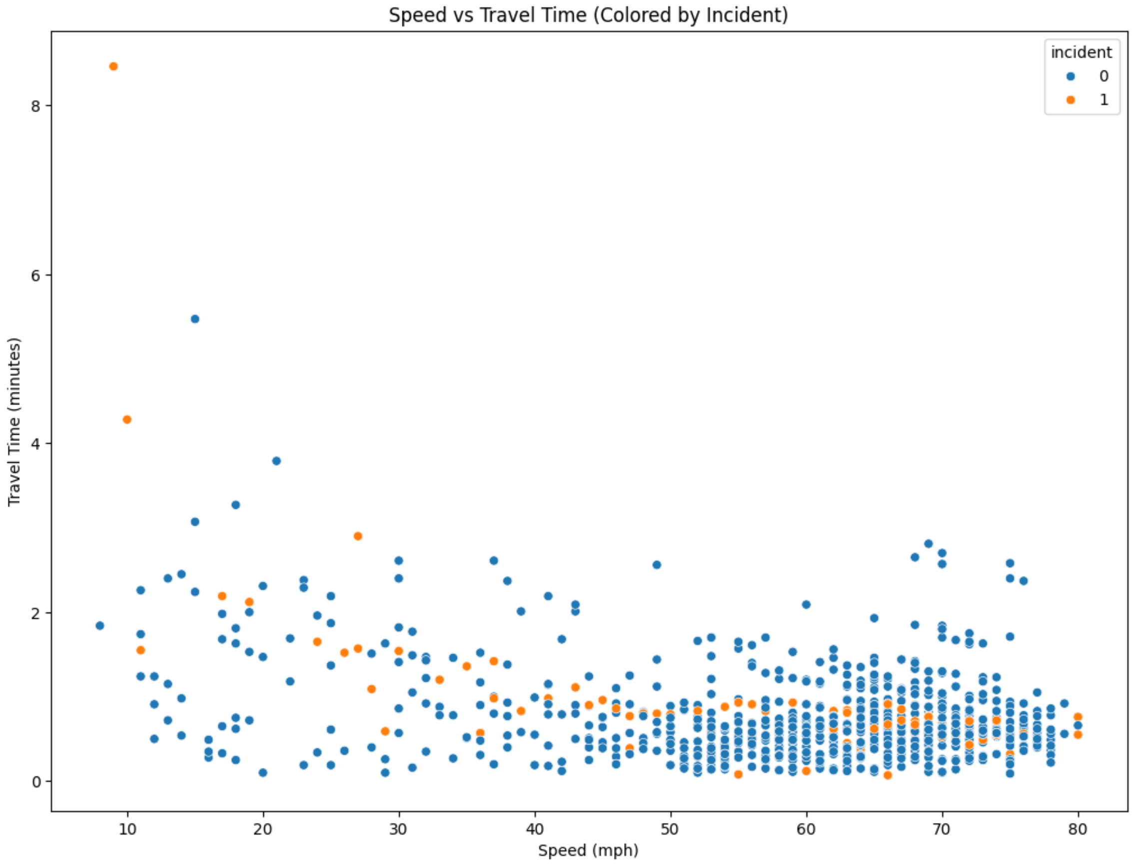

Figure 8: Relationship Between Vehicle Speed and Travel Time, Colored by Incident Presence

Results presented in Figure 8, demostrate the relationship between vehicle speed and travel time, with data points color-coded to indicate whether an incident occurred. The chart shows that as speed increases, travel time generally decreases, which is expected since vehicles traveling at higher speeds cover distances more quickly. Most data points are clustered at higher speeds with shorter travel times, while lower speeds show a wider range of travel times, including several long delays.

Notably, the points where incidents occurred (marked in a different color) are scattered mostly among the lower speed ranges and longer travel times. This suggests that incidents are more common when traffic is slower and less predictable, possibly due to congestion, stop-and-go conditions, or other disruptions on the road. At higher speeds, the number of incidents appears to decrease, and travel times remain consistently short.

Overall, the figure highlights a negative relationship between speed and travel time and suggests that lower speed conditions, often linked to congestion, are more prone to incidents. This emphasizes the importance of managing traffic flow to reduce congestion and improve road safety.

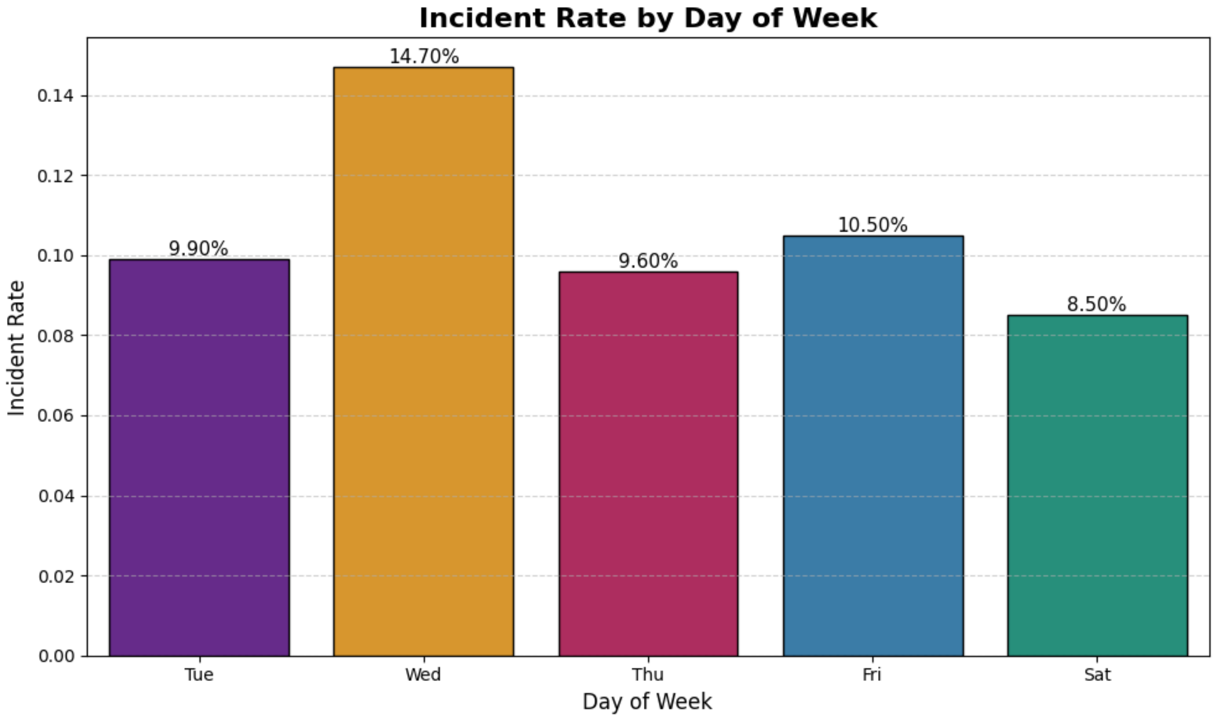

Figure 9: Incident Rate by Day of Week

Results presented in Figure 9, titled "Incident Rate by Day of Week" show how the frequency of incidents varies across the week, using the weekday coding where 1 = Sunday, 2 = Monday, 3 = Tuesday, and so on up to 7 = Saturday. The chart reveals that the highest incident rate occurs on Day 4, which corresponds to Wednesday. This suggests that the middle of the workweek is the period with the greatest risk of incidents, possibly due to increased work-related travel, commuter fatigue, or busy road conditions.

Tuesday (Day 3), Thursday (Day 5), and Friday (Day 6) show moderate incident rates, while Saturday (Day 7) has the lowest rate of all. This pattern indicates that weekends, particularly Saturday, tend to have fewer incidents, likely because there is less commuter traffic and people have more flexibility in their travel times. Interestingly, the slightly higher incident rate on Wednesday may reflect the pressure and workload building up midweek, affecting driver behavior.

Overall, the figure highlights that weekday traffic carries a higher risk of incidents compared to weekends, and that Wednesday is a particularly important day for targeted traffic safety interventions or public awareness efforts.

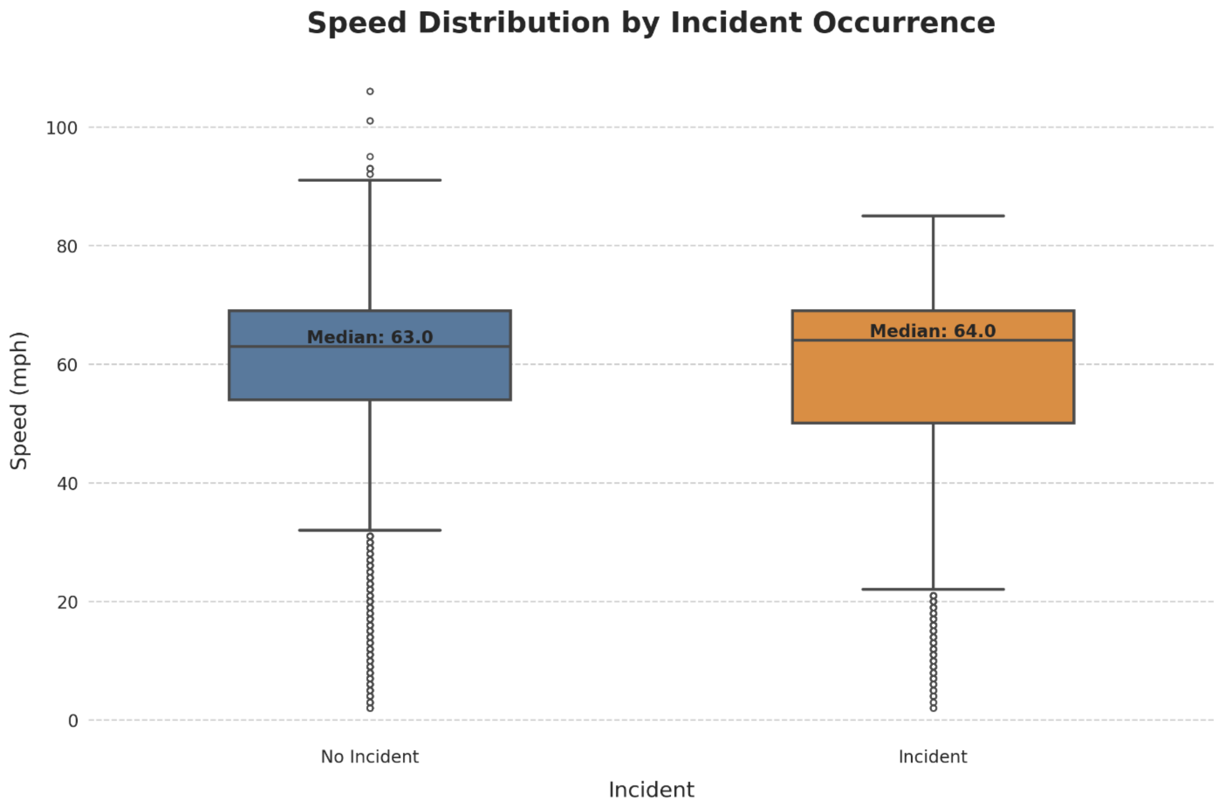

Figure 10: Speed Distribution by Incident Occurrence

Results presented in Figure 10, titled **"**Speed Distribution by Incident Occurrence**"** present a box plot comparing vehicle speed distributions between cases with and without incidents. Specifically, it shows how speeds differ between two groups: 0 (no incident) and 1 (incident).

The analysis shows that incidents tend to occur at slightly lower speeds and display greater speed variability compared to non-incident cases. For the non-incident group (0), speeds are more widely spread, with a slightly lower median and some extreme high-speed outliers, including speeds over 90 mph. By contrast, in the incident group (1), the median speed is slightly higher, but the overall speed range is narrower, the maximum observed speeds are lower, and there are fewer extreme outliers.

These patterns suggest that incidents are more likely to happen within a specific speed range with less variation, while non-incident conditions include both very low and very high speeds. This may indicate that stable, predictable speeds are associated with higher risk, whereas a wider spread of speeds — often seen in less congested or less stressful road conditions — tends to occur when no incidents happen.

Here are the detailed statistics broken down by incident status:

- Incidents (1): slightly higher median speed, narrower range, fewer outliers
- Non-incidents (0): slightly lower median speed, wider range, more high-speed outliers

In the following Table 4 and Table 5, we summarize key descriptive statistics for the variables speed and travel_time_minutes, grouped by whether an incident occurred or not. The statistics reveal meaningful differences between the two groups.

The summary statistics provide a useful comparison between records with no incidents and those with incidents. On average, non-incident cases show a slightly higher mean speed of 59.06 mph compared to 57.19 mph in incident cases, though the incident group exhibits a wider spread of speeds (standard deviation of 17.23 vs. 14.84). The maximum speed recorded without incidents reached 106 mph, while incidents peaked at 85 mph.

For travel time, incident cases had a higher average (0.90 minutes) compared to non-incident records (0.74 minutes), along with greater variability, with some trips lasting up to 36 minutes, whereas the maximum for non-incidents was 27 minutes. These observations indicate that both speed and travel time are moderately influenced by the occurrence of incidents.

Table 4. Speed Summary

Incident Status	Count	Mean (mph)	Std Dev	Min	25%	50% (Median)	75%	Max
No Incident (0)	319,947	59.06	14.84	2	54	63	69	106
Incident (1)	37,424	57.19	17.23	2	50	64	69	85

Table 5: Travel Time (minutes)

Incident Status	Mean	Std Dev	Min	25%	50% (Median)	75%	Max
No Incident (0)	0.74	0.60	0.09	0.43	0.62	0.86	27.25
Incident (1)	0.90	0.83	0.06	0.60	0.70	0.88	36.06

5.3 HIGH-RISK SEGMENT IDENTIFICATION

In the following Figure 11, this map zooms in on the core area of the study, allowing us to better observe how segment clusters and incident severity are geographically distributed. This detailed view is especially useful for understanding localized traffic risks and identifying hotspots along the corridor. By incorporating cluster labels as a new input feature in our machine learning models, we enhance their ability to detect regional and directional accident patterns—ultimately improving both predictive accuracy and model interpretability.

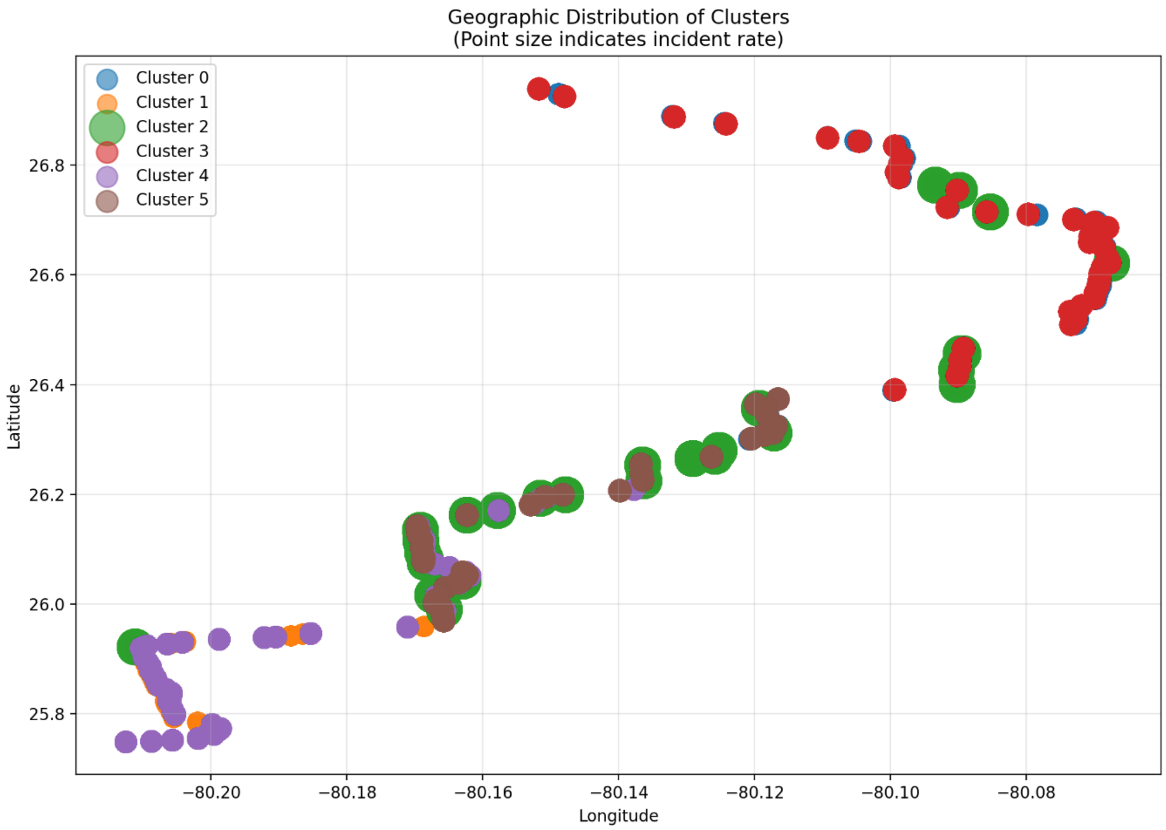

Figure 11: Zoomed-In View of Segment Clusters and Incident Rates

Results presented in Figure 12 show the top 10 segments with the highest number of reported accidents in the study area. The chart highlights that I-95 in Broward County experienced the most incidents, followed closely by segments in West Palm Beach and Middle River. This visualization helps pinpoint high-risk locations where traffic safety measures and infrastructure improvements may be most needed.

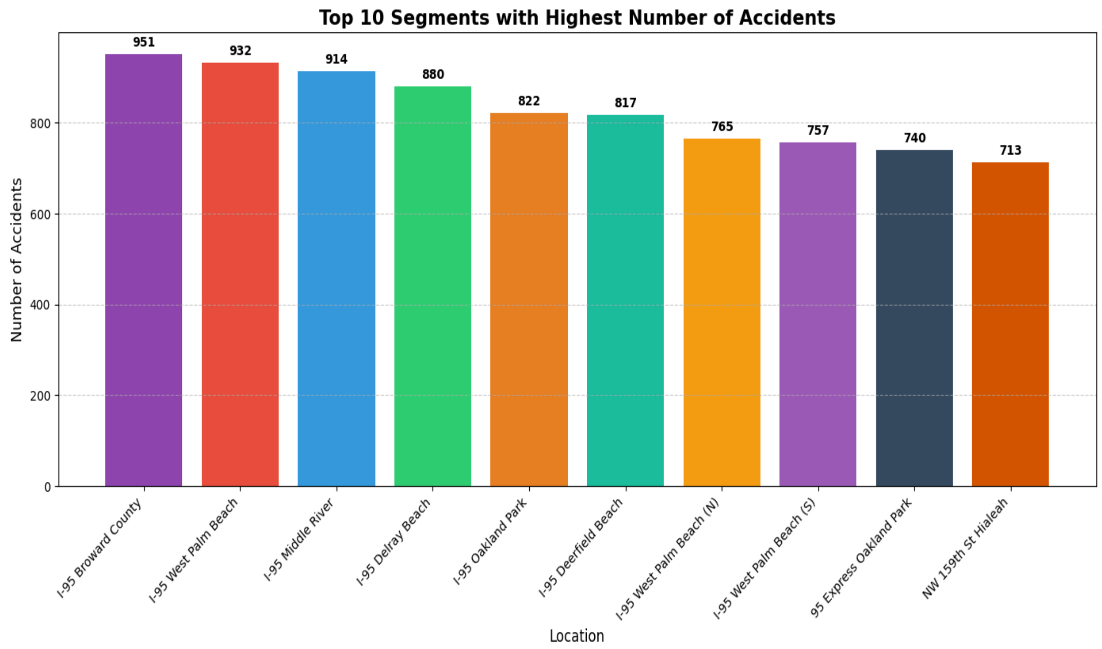

Figure 12: Top 10 Segments with Highest Number of Accidents

Table 6 presents the top 10 traffic segments with the highest incident rates based on geographic coordinates and location names. These segments, primarily located along I-95 in southeastern Florida, demonstrate elevated accident frequency, with the highest rate recorded at 2.61 incidents per segment in Broward County. Identifying such high-risk locations is critical for targeting traffic safety interventions and prioritizing infrastructure improvements.

Table 6: Top 10 Segments with the Highest Incident Rates

segment_id	latitude	longitude	incident_rate	location_name
189	26.07641	-80.1685	2.61	I-95, Broward County, FL
175	26.71536	-80.0853	2.55	I-95, West Palm Beach, FL
63	26.16219	-80.1623	2.5	I-95, Middle River, FL
188	26.45726	-80.0893	2.41	I-95, Delray Beach, FL
44	26.19347	-80.1516	2.25	I-95, Oakland Park, FL
155	26.3124	-80.1172	2.24	I-95, Deerfield Beach, FL
73	26.75447	-80.0898	2.1	I-95, West Palm Beach (N), FL
234	26.76417	-80.0934	2.07	I-95, West Palm Beach (S), FL
49	26.17092	-80.1579	2.03	95 Express, Oakland Park, FL
24	25.92184	-80.2112	1.95	NW 159th St, Hialeah, FL

5.4 Travel Time Prediction Results

To predict travel time for each road segment and time interval, we trained several regression models using traffic, temporal, and spatial features. The input variables included:

- weekday
- speed
- reference_speed
- confidence_factor
- tmc_direction
- time_sin and time_cos (cyclical encoding of time)
- One-hot encoded cluster labels (cluster_0 to cluster_5)
- incident_flag (indicating whether an accident occurred)

The dependent variable was travel_time_minutes.

We evaluated the models using standard regression metrics:

- R^2 Score (explained variance)
- Mean Squared Error (MSE)
- Root Mean Squared Error (RMSE)
- Mean Absolute Error (MAE)

i. XGBoost

The XGBoost model performed exceptionally well in predicting travel time, achieving an R^2 score of 0.9984, indicating that the model explained nearly all the variance in the data. The very low RMSE (0.025 minutes) and MAE (0.006 minutes) suggest highly accurate predictions.

- R^2 Score: 0.9984
- Mean Squared Error (MSE): 0.00063
- Root Mean Squared Error (RMSE): 0.025
- Mean Absolute Error (MAE): 0.0062

ii. Random Forest Regression

The Random Forest model achieved nearly identical performance to XGBoost, also with an R^2 score of 0.9984. Its error metrics remained low, confirming the model's strength in capturing complex relationships in the data.

- R^2 Score: 0.9948
- Mean Squared Error (MSE): 0.00205
- Root Mean Squared Error (RMSE): 0.04531
- Mean Absolute Error (MAE): 0.00073

iii. Decision Tree Regression

The Decision Tree model produced good results, although it was slightly less accurate compared to the ensemble models. The R^2 score dropped to 0.9948, and error values increased slightly, but still remained low overall.

- R^2 Score: 0.9948
- Mean Squared Error (MSE): 0.0020
- Root Mean Squared Error (RMSE): 0.0453
- Mean Absolute Error (MAE): 0 0.0007

iv. Neural Network Regression

A Multilayer Perceptron (MLP) model was trained to predict short-term travel time using features derived from the NPMRDS dataset. To identify the optimal model configuration, a Bayesian hyperparameter tuning procedure was conducted. The objective function minimized the Root Mean Squared Error (RMSE) on a held-out validation set. The tuning process explored a range of architectural and training parameters, including:

- Number of hidden layers: 1 to 3
- Units per layer: 32 to 256
- Dropout rate: 0.0 to 0.5
- Batch size: 128, 256, 512
- Optimizer: Adam and RMSprop
- Learning rate: 1e-4 to 1e-2 (log-scaled)

The best-performing model architecture, consists of:

- 3 hidden layers with 127, 233, and 61 neurons respectively
- ReLU activation functions for all hidden layers
- A dropout rate of 13.2% applied after the second hidden layer
- RMSprop optimizer with a learning rate of 0.000143

Trained using a batch size of 512 over a maximum of 50 epochs, with early stopping applied based on validation loss, we achieved the following final model performance:.

- R^2 Score: 0.9931
- Mean Squared Error (MSE): 0.0028
- Root Mean Squared Error (RMSE): 0.0526
- Mean Absolute Error (MAE): 0 0.0117

These results indicate that the optimized MLP model achieved high predictive accuracy, capturing both recurring traffic patterns and localized variability with minimal error.

Table 7: . Model Performance Metrics for Travel Time Prediction

Model	R^2 Score	MSE	RMSE	MAE
XGBoost	0.9984	0.00063	0.0250	0.0062
Random Forest	0.9948	0.00205	0.0453	0.00073
Decision Tree	0.9948	0.00200	0.0453	0.0007
Multilayer Perceptron (MLP)	0.9931	0.00280	0.0526	0.0117

The performance results indicate that all four models achieved high predictive accuracy for travel time estimation, with R^2 scores exceeding 0.99 across the board. XGBoost outperformed the other models, delivering the lowest error metrics (RMSE: 0.0250, MAE: 0.0062), highlighting its strength in capturing complex nonlinear relationships. Both Random Forest and Decision Tree models achieved nearly identical results, demonstrating strong accuracy with excellent interpretability and minimal tuning. The Multilayer Perceptron (MLP) also performed well, though it showed slightly higher error values, reflecting the need for more data or fine-tuned regularization to match the precision of tree-based models. Overall, tree-based ensemble methods proved most effective for this short-term travel time prediction task.

5.5 Predicting Number of Accidents

To evaluate the performance of different machine learning models for accident prediction, we applied four advanced classification algorithms: XGBoost, Random Forest, Decision Tree, and a Neural Network (MLPClassifier). These models are widely used in predictive analytics due to their ability to handle complex, non-linear relationships in structured data.

Each model was trained on the same dataset, using a combination of engineered traffic features (such as speed, reference speed, and confidence factor) and spatial-temporal variables (such as time of day, day of the week, and location-based clusters). The target variable was binary, indicating the presence of an incident (incident = 1) or its absence (incident = 0).

The dataset was split into training and testing sets, and the models were evaluated using standard classification metrics such as accuracy, precision, recall, and F1-score. This allowed for a consistent comparison across models, focusing particularly on how well each model performed in identifying actual incidents, which are relatively rare but highly impactful.

i. XGBoost

The XGBoost model achieved an overall accuracy of 96.30%, correctly classifying most non-incident cases. It showed high precision (81%) for predicting incidents and relatively moderate recall (84%), meaning it missed few actual accident cases. The model leaned toward conservatively predicting non-incidents, resulting in a high true negative rate but a weaker true positive rate.

- Precision (incident): 0.81
- Recall (incident): 0.84
- F1-score (incident): 0.82
- Accuracy: 96.30%

This model is suitable if minimizing false alarms is more important than catching all incidents, but may not be ideal for safety-critical applications where catching every accident is essential.

ii. Random Forest

The Random Forest classifier significantly improved both precision and recall compared to XGBoost, achieving an overall accuracy of 97.9%. It was able to detect a larger number of actual incidents, with recall rising to 95% and a solid F1-score of 0.96D for the positive class.

- Precision (incident): 0.96
- Recall (incident): 0.95
- F1-score (incident): 0.96
- Accuracy: 97.9%

This balance between true positive and true negative detection makes Random Forest a strong candidate for real-time or near-real-time crash prediction systems.

iii. Decision Tree

The Decision Tree classifier achieved an overall accuracy of 93.87% as well. It achieved slightly low recall (0.49) and high precision (85%) for the incident class, resulting in an F1-score of 0.62.

- Precision (incident): 0.81
- Recall (incident): 0.84
- F1-score (incident): 0.82
- Accuracy: 93.87%

Despite its relatively simple structure, the Decision Tree model produced strong predictive performance, demonstrating its potential value as a practical and interpretable tool in freight travel time prediction. Its transparent, rule-based decision-making process allows for straightforward interpretation of model behavior and variable importance—an advantage in transportation applications where explainability is critical for stakeholder trust and operational deployment. As such, Decision Trees may be particularly useful in scenarios where model transparency, speed of inference, and ease of implementation are prioritized over marginal gains in predictive accuracy.

iv. Neural Network (MLP Classifier)

The Neural Network model achieved a slightly lower overall accuracy of 94.85%, but its performance in detecting incidents was somewhat less robust. While it maintained high precision

(78%) for the incident class, its recall dropped to 70%, indicating more missed accidents compared to the tree-based models.

- Precision (incident): 0.78
- Recall (incident): 0.70
- F1-score (incident): 0.74
- Accuracy: 94.85%

Neural Networks can model more complex patterns, but in this case, the simpler models outperformed it in incident detection, possibly due to the structured nature of the data.

Table 8: Results Sumary

Model	Accuracy	Precision (Class 1)	Recall (Class 1)	F1-Score (Class 1)
Decision Tree	93.87%	0.85	0.49	0.62
Random Forest	97.90%	0.96	0.96	0.95
XGBoost	96.30%	0.81	0.84	0.82
Multilayer Perceptron (MLP)	94.85%	0.78	0.70	0.74

Among the models evaluated for incident prediction, the Random Forest classifier delivered the most robust performance, with high precision and recall indicating its strong ability to correctly identify incidents while minimizing false alarms. The XGBoost model also performed well and offers the advantage of interpretability, though with slightly lower accuracy. The Multilayer Perceptron (MLP) showed moderate performance, with a tendency to overpredict incidents, likely due to lower precision. Decision Tree, due to its inherent simple structure, struggled in this case, compared to the other more advanced classifiers, capturing fewer actual incidents, suggesting limitations in handling class imbalance within this specific dataset.

CHAPTER 6 DISCUSSION AND CONCLUSION

This research explored the intersection of big data analytics and machine learning to enhance freight network design, particularly under the pressures of disruption and uncertainty. The first step involved identifying and leveraging diverse data sources—both open and proprietary—to understand how data is being used globally to solve transportation challenges. This includes crowdsourced probe data, real-time traffic feeds, and freight movement datasets like the NPMRDS, which offer significant potential for predictive modeling.

In the context of resilient supply chain design, the study emphasized the need for robust, flexible infrastructures supported by accurate risk detection and disruption forecasting. Predictive risk assessment models that incorporate global data—from weather and economic indicators to social media and news—can help supply chain stakeholders anticipate and respond to emergent threats. This approach requires handling large volumes of structured and unstructured data, where big data technologies play a vital role in filtering, aggregating, and analyzing inputs in real time.

The second phase of the research focused on constructing a comprehensive, multisource dataset to support the development of a freight network design framework resilient to disruptions. This effort highlights the need for scalable modeling tools that integrate geographic information systems (GIS), cloud computing, and advanced analytics, enabling stakeholders to design networks that are not only efficient but also adaptive to complex, dynamic conditions.

On the predictive modeling front, four ML algorithms, XGBoost, Random Forest, Decision Tree, and Multilayer Perceptron (MLP), were evaluated for travel time forecasting. All models achieved high accuracy ($R^2 > 0.99$), with XGBoost delivering the most precise predictions, followed closely by Random Forest and Decision Tree models. MLP performed well but required further regularization. For incident prediction, Random Forest emerged as the most reliable classifier, achieving strong precision and recall, and effectively minimizing false positives and false negatives. XGBoost also performed well, while MLP showed moderate effectiveness. The Decision Tree model, although interpretable, struggled with class imbalance and underperformed relative to ensemble and neural models.

In conclusion, this research demonstrates that integrating big data and machine learning into freight system design can significantly enhance predictive capabilities and operational resilience. With the right data and computational tools, stakeholders can move toward smarter, more adaptive freight networks capable of withstanding disruption and uncertainty.

REFERENCES

Baryannis, G., Dani, S., Validi, S., & Antoniou, G. (2019). *Decision Support Systems and Artificial Intelligence in Supply Chain Risk Management* (pp. 53–71). https://doi.org/10.1007/978-3-030-03813-7_4

Baryannis, G., Validi, S., Dani, S., & Antoniou, G. (2019). Supply chain risk management and artificial intelligence: state of the art and future research directions. *International Journal of Production Research*, *57*(7), 2179–2202. https://doi.org/10.1080/00207543.2018.1530476

Cedillo-Campos, M. G., Pérez-González, C. M., Piña-Barcena, J., & Moreno-Quintero, E. (2019). Measurement of travel time reliability of road transportation using GPS data: A freight fluidity approach. *Transportation Research Part A: Policy and Practice*, *130*, 240–288. https://doi.org/10.1016/j.tra.2019.09.018

Chandra, S. R., & Al-Deek, H. (2009). Predictions of Freeway Traffic Speeds and Volumes Using Vector Autoregressive Models. *Journal of Intelligent Transportation Systems*, *13*(2), 53–72. https://doi.org/10.1080/15472450902858368

Figliozzi, M. A., Wheeler, N., Albright, E., Walker, L., Sarkar, S., & Rice, D. (2011). Algorithms for Studying the Impact of Travel Time Reliability along Multisegment Trucking Freight Corridors. *Transportation Research Record: Journal of the Transportation Research Board*, *2224*(1), 26–34. https://doi.org/10.3141/2224-04

Ghadge, A., Dani, S., & Kalawsky, R. (2012). Supply chain risk management: present and future scope. *The International Journal of Logistics Management*, *23*(3), 313–339. https://doi.org/10.1108/09574091211289200

Golob, T. F., & Recker, W. W. (2003). Relationships Among Urban Freeway Accidents, Traffic Flow, Weather, and Lighting Conditions. *Journal of Transportation Engineering*, *129*(4), 342–353. https://doi.org/10.1061/(ASCE)0733-947X(2003)129:4(342)

Guo, F., Polak, J. W., & Krishnan, R. (2018). Predictor fusion for short-term traffic forecasting. *Transportation Research Part C: Emerging Technologies*, *92*, 90–100. https://doi.org/10.1016/j.trc.2018.04.025

Habtemichael, F. G., & Cetin, M. (2016). Short-term traffic flow rate forecasting based on identifying similar traffic patterns. *Transportation Research Part C: Emerging Technologies*, *66*, 61–78. https://doi.org/10.1016/j.trc.2015.08.017

Lee, J. M., & Wong, E. Y. (2021). Suez Canal blockage: an analysis of legal impact, risks and liabilities to the global supply chain. *MATEC Web of Conferences*, *339*, 01019. https://doi.org/10.1051/matecconf/202133901019

Li, J., Guo, F., Sivakumar, A., Dong, Y., & Krishnan, R. (2021). Transferability improvement in short-term traffic prediction using stacked LSTM network. *Transportation Research Part C: Emerging Technologies*, *124*, 102977. https://doi.org/10.1016/j.trc.2021.102977

Ma, X., Tao, Z., Wang, Y., Yu, H., & Wang, Y. (2015). Long short-term memory neural network for traffic speed prediction using remote microwave sensor data. *Transportation Research Part C: Emerging Technologies*, *54*, 187–197. https://doi.org/10.1016/j.trc.2015.03.014

Mishra, N., Soni, H. K., Sharma, S., & Upadhyay, A. K. (2017). A Comprehensive Survey of Data Mining Techniques on Time Series Data for Rainfall Prediction. *Journal of ICT Research and Applications*, *11*(2), 168. https://doi.org/10.5614/itbj.ict.res.appl.2017.11.2.4

Pournader, M., Ghaderi, H., Hassanzadegan, A., & Fahimnia, B. (2021). Artificial intelligence applications in supply chain management. *International Journal of Production Economics*, *241*, 108250. https://doi.org/10.1016/j.ijpe.2021.108250

Sapankevych, N., & Sankar, R. (2009). Time Series Prediction Using Support Vector Machines: A Survey. *IEEE Computational Intelligence Magazine*, *4*(2), 24–38. https://doi.org/10.1109/MCI.2009.932254

Shams, K., Asgari, H., & Jin, X. (2017). Valuation of travel time reliability in freight transportation: A review and meta-analysis of stated preference studies. *Transportation Research Part A: Policy and Practice*, *102*, 228–243. https://doi.org/10.1016/j.tra.2016.08.001

Shang, Y., Dunson, D., & Song, J.-S. (2017). Exploiting Big Data in Logistics Risk Assessment via Bayesian Nonparametrics. *Operations Research*, *65*(6), 1574–1588. https://doi.org/10.1287/opre.2017.1612

Theofilatos, A., & Yannis, G. (2014). A review of the effect of traffic and weather characteristics on road safety. *Accident Analysis & Prevention*, *72*, 244–256. https://doi.org/10.1016/j.aap.2014.06.017

Tian Guangqiang. (2019). *Routledge Handbook of the Belt and Road* (Fang Cai & Nolan Peter, Eds.; 1st ed., Vol. 642).

Tirkolaee, E. B., Sadeghi, S., Mooseloo, F. M., Vandchali, H. R., & Aeini, S. (2021). Application of Machine Learning in Supply Chain Management: A Comprehensive Overview of the Main Areas. *Mathematical Problems in Engineering*, *2021*, 1–14. https://doi.org/10.1155/2021/1476043

U.S. Department of Transportation. (2020). *The National Performance Management Research Data Set (NPMRDS) and Application for Work Zone Performance Measurement*.

Waller, M. A., & Fawcett, S. E. (2013). Big Data, Predictive Analytics, and Theory Development in the Era of a Maker Movement Supply Chain. *Journal of Business Logistics*, *34*(4), 249–252. https://doi.org/10.1111/jbl.12024

Wang, Z., Goodchild, A., & McCormack, E. (2016). Measuring Truck Travel Time Reliability Using Truck Probe GPS Data. *Journal of Intelligent Transportation Systems*, *20*(2), 103–112. https://doi.org/10.1080/15472450.2014.1000455

Zage, D., Glass, K., & Colbaugh, R. (2013). Improving supply chain security using big data. *2013 IEEE International Conference on Intelligence and Security Informatics*, 254–259. https://doi.org/10.1109/ISI.2013.6578830

Zeng, X., & Zhang, Y. (2013). Development of Recurrent Neural Network Considering Temporal-Spatial Input Dynamics for Freeway Travel Time Modeling. *Computer-Aided Civil and Infrastructure Engineering*, *28*(5), 359–371. https://doi.org/10.1111/mice.12000