

FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO

Time Series Forecasting and Pattern Analysis of Urban and Suburban Highway Traffic for Enhanced Revenue Planning

Ingrid Rodrigues Diniz



Mestrado em Engenharia e Ciência de Dados

Supervisor: João Pedro Mendes Moreira and Vítor Manuel Araújo
Cerqueira

November 17, 2025

Time Series Forecasting and Pattern Analysis of Urban and Suburban Highway Traffic for Enhanced Revenue Planning

Ingrid Rodrigues Diniz

Mestrado em Engenharia e Ciência de Dados

November 17, 2025

Abstract

This dissertation presents a comparative analysis of long-term traffic forecasting models from three distinct families—statistical (ARIMA), machine learning (GBM), and deep learning (LSTM)—each enhanced with prediction intervals to quantify forecast uncertainty. The study focuses on two highway concessions operated by Ascendi: one located in an urban environment and another in a suburban area. Using daily traffic data disaggregated by vehicle class, the models aim to forecast traffic volumes three months ahead, providing reliable inputs for the company’s revenue planning process.

To improve predictive performance, exogenous variables were incorporated, including calendar effects, macroeconomic indicators, and dummy variables signaling external events. Model evaluation combined point accuracy metrics (MASE and RMSSE) and probabilistic metrics (Coverage Rate and normalized Prediction Interval Width, nPIW). The research followed the CRISP-DM methodology, encompassing all stages from business understanding to deployment.

Results indicate that the LGBM achieved the best performance in both urban and suburban environments. Nonetheless, similar behavioral patterns were observed across vehicle classes, and for a particularly stable urban vehicle class, the AutoLSTM model outperformed all others. Overall, the study highlights the potential of data-driven approaches for reliable long-term traffic forecasting, supporting more robust financial and operational planning in highway management.

Keywords: traffic forecasting; urban vs. suburban; ARIMA; Gradient Boosting; LSTM; probabilistic forecasting; prediction intervals; MASE; RMSSE; coverage; nPIW; revenue planning.

Resumo

Esta dissertação apresenta uma análise comparativa do desempenho de modelos de previsão de tráfego de longo prazo pertencentes a três famílias distintas — estatística (ARIMA), machine learning (GBM) e deep learning (LSTM) — todos enriquecidos com intervalos de previsão, de forma a quantificar a incerteza associada às previsões. O estudo incide sobre duas concessões rodoviárias da Ascendi: uma localizada em ambiente urbano e outra em ambiente suburbano. Com base em dados diários de tráfego desagregados por classe de veículo, os modelos têm como objetivo prever volumes de tráfego com um horizonte temporal de três meses, fornecendo estimativas robustas para apoiar o planejamento de receitas da empresa.

Para melhorar o desempenho preditivo, foram incorporadas variáveis exógenas, incluindo efeitos de calendário, indicadores macroeconômicos e variáveis dummy que sinalizam eventos externos. A avaliação dos modelos combinou métricas de precisão pontual (MASE e RMSSE) e métricas probabilísticas (Coverage Rate e largura normalizada do intervalo de previsão – nPIW). A investigação seguiu a metodologia CRISP-DM, abrangendo todas as fases, desde a compreensão do negócio até à implementação dos modelos.

Os resultados indicam que o modelo LGBM apresentou o melhor desempenho em ambos os ambientes, urbano e suburbanp. Ainda assim, observou-se um padrão de comportamento semelhante entre as classes de veículos, e, para uma classe particularmente estável no ambiente urbano, o modelo AutoLSTM superou os restantes. De forma geral, o estudo evidencia o potencial de orientadas por dados para previsões de tráfego de longo prazo fiáveis, capazes de apoiar um planejamento financeiro e operacional mais robusto na gestão de infraestruturas rodoviárias.

Palavras-chave: previsão de tráfego; urbano vs. suburbano; ARIMA; Gradient Boosting; LSTM; previsão probabilística; intervalos de previsão; MASE; RMSSE; cobertura; nPIW; revenue planning.

Acknowledgments

This dissertation would not have been possible without the exceptional guidance and continuous support of my supervisors, Professor João Pedro Mendes Moreira and Professor Vítor Manuel Araújo Cerqueira. Their profound knowledge and thoughtful feedback were truly invaluable, shaping every stage of this work.

I am also immensely thankful to Ana Santos for her unwavering assistance with all bureaucratic processes and for always approaching every task with a welcoming smile.

To my brother and best friend, Igor Diniz, thank you for being a constant source of support throughout this master's journey, from start to finish. Sharing this experience with you made it immeasurably better.

My heartfelt appreciation goes to Ascendi for generously providing the data and allowing me the unique opportunity to develop my thesis within a real-world corporate setting. This practical application imbued my research with deep meaning. A special note of thanks to the entire Business Intelligence team - Sérgio Fernandes, Bruno Machado, Vanessa Boaventura- for their warm reception, willingness to share their expertise, and for facilitating access to the crucial information that underpinned this dissertation.

I extend my gratitude to my classmates and the professors whose shared knowledge significantly enriched my academic journey.

To my incredible parents, Maria Lúcia Diniz and Evandro Diniz, thank you for granting me the opportunity to pursue my studies in Portugal and for celebrating every milestone, big or small. To my boyfriend, Gabriel Meirinho, thank you for your steadfast presence through every challenge. And to my sister-in-law, Laura Porto, who I consider family, your unwavering support means the world to me.

Finally, to all those who have touched my life and contributed, in countless ways, to the construction of this academic chapter—thank you for making this accomplishment a reality.

Ingrid Diniz

*“If I have been able to see further,
it is only because I stood on the shoulders of giants.”*

Isaac Newton

Contents

1	Introduction	1
1.1	Motivation	2
1.2	Goals	3
1.3	Document Structure	4
2	Literature Review	5
2.1	Time-Series Foundations for Highway Traffic Analysis	5
2.1.1	Temporal Structure: Trend, Seasonality, and Cycles	6
2.1.2	Forecast Horizons: Short- vs Long-Term	7
2.1.3	Uncertainty	8
2.1.4	Urban–Suburban Traffic Regimes	10
2.1.5	Contextual Exogenous Variables	11
2.2	Modeling Paradigms	11
2.3	Statistical Models	12
2.4	Machine Learning Models	12
2.5	Deep Learning Models	13
2.6	Deep Description of Architectures Used	14
2.6.1	ARIMA Family	14
2.6.2	LightGBM	15
2.6.3	LSTM	17
2.6.4	Evaluation Protocols and Metrics	19
2.7	Gaps and Contributions	20
3	Case Study	22
3.1	Business Understanding	22
3.1.1	Determine Business Objectives	23
3.1.2	Determine Data Mining Goals	25
3.2	Data Understanding	27
3.2.1	Collect Initial Data	27
3.2.2	Data Description	30
3.2.3	Exploratory Data Analysis	33
3.2.4	Data Quality Verification	37
4	Implementation	48
4.1	Data Preparation	48
4.1.1	Select Data	48

4.1.2	Clean Data	53
4.1.3	Construct Data	54
4.1.4	Integrate Data	55
4.1.5	Format Data	56
4.1.6	Final Dataset	57
4.2	Modeling	57
4.2.1	Selection of Modeling Technique	58
4.2.2	Generate Test Design	60
4.2.3	Build Model	61
4.2.4	Access Model	63
5	Results and Discussion	66
5.1	Evaluation	66
5.1.1	Evaluate Results	67
5.1.2	Feature Importance Results	67
5.1.3	Results	68
5.1.4	Discussion	77
5.2	Deployment	80
6	Conclusions and Future Work	82
6.1	Future Work	83
A	Vehicles Classification	84
A.1	Toll Classes in Portugal	84
B	Pattern Analysis by Vehicle Class and Environment	86
	References	89

List of Figures

1.1	Electronic tolling system and price classes on the A41 motorway in Portugal. The sign shows different toll rates by vehicle class Renascença (2025)	2
2.1	A time series with varying characteristics Erlangung (2023) .	6
2.2	Illustration of probabilistic forecasting outputs: distribution forecast, quantile forecast, prediction interval forecast, and scenario forecast.	8
2.3	Schematic representation of the LightGBM model Yao et al. (2022) .	17
2.4	Recurrent Neural Network architecture.	18
3.1	Detailed Business Understanding, Chapman (2000)	23
3.2	Detailed Data Understanding, Chapman (2000)	28
3.3	Geographical framework of the study areas: (a) urban and (b) suburban concession, Ascendi (2025c) .	29
3.4	Time series plots for each vehicle class in the (A) urban and (B) suburban environments.	34
3.5	Monthly seasonal additive decomposition for small vehicles in BLA	35
3.6	ACF for each vehicle class in the (A) urban and (B) suburban environments.	38
3.7	PACF for each vehicle class in the (A) suburban and (B) urban environments.	39
3.8	Daily traffic volume for Class 1 vehicles in the suburban environment during 2024, with significant holidays highlighted. Each red marker indicates a national or major holiday, annotated with the respective holiday name.	41
3.9	Average percentage change in traffic volume by holiday and vehicle class in the (A) urban and (B) suburban environments.	42
3.10	Average percentage change in weekend traffic volumes adjacent to holidays for each vehicle class in the (A) urban and (B) suburban environments.	44
3.11	Traffic change analysis for Classes 2 and 3 in the urban environment before and after February 10, 2025. Left: Daily time series with seven-day moving average and mean lines for the periods before and after the intervention. Right: Boxplots showing the distribution and means for each period.	45
3.12	Comparison of Class 5 (motorcycle) traffic on Expomoto event days and control days in the suburban environment.	47
3.13	Abrupt changes in traffic volume on specific dates in the suburban environment, with annotation of fire-related incidents.	47
4.1	Detailed Data Preparation, Chapman (2000)	49

4.2	Pearson correlation coefficients between vehicle classes, calendar and macroeconomic indicators, and flagged external events in (A) urban and (B) suburban environments.	51
4.3	Detailed Modeling, Chapman (2000) , Chapman (2000)	58
4.4	Time-series cross-validation for time series. From top to bottom: Split 1, Split 2, Split 3, and the final Full Training + Test.	61
4.5	Unified modeling pipeline: Raw Data → Transformations → Model training (AutoARIMA, LightGBM, AutoLSTM), each validated with cross validation, followed by results comparison across urban and suburban contexts.	62
4.6	Feature Ablation Process	64
5.1	Forecasts of AutoARIMA for all vehicle classes in the urban (GP) environment. Each panel shows observed values (blue), point forecasts (red), 95% prediction intervals (shaded), and ground truth values (green).	72
5.2	Forecasts of AutoARIMA for all vehicle classes in the Suburban (BLA) environment. Each panel shows observed values (blue), point forecasts (red), 95% prediction intervals (shaded), and ground truth values (green).	73
5.3	Forecasts of LGBM for all vehicle classes in the urban (GP) environment. Each panel shows observed values (blue), point forecasts (red), 95% prediction intervals (shaded), and ground truth values (green).	74
5.4	Forecasts of LGBM for all vehicle classes in the Suburban (BLA) environment. Each panel shows observed values (blue), point forecasts (red), 95% prediction intervals (shaded), and ground truth values (green).	75
5.5	Forecasts of AutoLSTM for all vehicle classes in the urban (GP) environment. Each panel shows observed values (blue), point forecasts (red), 95% prediction intervals (shaded), and ground truth values (green).	76
5.6	Forecasts of AutoLSTM for all vehicle classes in the Suburban (BLA) environment. Each panel shows observed values (blue), point forecasts (red), 95% prediction intervals (shaded), and ground truth values (green).	77
B.1	Pattern analysis for class 1 vehicles in the (A) urban and (B) suburban environments.	86
B.2	Pattern analysis for class 2 vehicles in the (A) urban and (B) suburban environments.	87
B.3	Pattern analysis for class 3 vehicles in the (A) urban and (B) suburban environments.	87
B.4	Pattern analysis for class 4 vehicles in the (A) urban and (B) suburban environments.	88
B.5	Pattern analysis for class 5 vehicles in the (A) urban and (B) suburban environments.	88

List of Tables

3.1	Descriptive statistics of daily traffic volumes by vehicle class and environment (Jan 2022–May 2025).	31
3.2	Descriptive statistics for quarterly GDP values in Portugal (2017Q1–2025Q1), in millions of euros.	32
3.3	Descriptive statistics for monthly CPI inflation rates (2017–2025) by region.	32
3.4	ADF Results for Stationarity by Environment and Class	40
3.5	Statistical summary of traffic change before and after February 10, 2025 (urban environment).	46
4.1	Sample of the final, formatted dataset ready for modeling. The full dataset contains 28 columns, including engineered cyclical features, event indicators, and macroeconomic variables. The symbol $\cdots$ indicates additional columns present in the dataset.	57
5.1	Exogenous features used by each model.	68
5.2	Average performance of AutoARIMA, LightGBM and AutoLSTM across three cross-validation windows. Metrics reported separately for urban (GP) and suburban (BLA) highways by vehicle class.	70
5.3	Performance of AutoARIMA, LightGBM and AutoLSTM on the hold-out test set (first month ahead). Metrics reported separately for urban (GP) and suburban (BLA) highways by vehicle class.	71
A.1	Toll Classes in Portugal	85

Abbreviations

ACF	Autocorrelation Function		
ADF	Augmented Dickey–Fuller		
AND	Logical AND (conjunction)		
AR	Autoregressive		
ARIMA	Autoregressive Integrated Moving Average		
ARIMAX	Autoregressive Integrated Moving Average with eXogenous regressors		
AWS	Amazon Web Services		
BLA	Beiras Litoral e Alta		
BU	Business Understanding		
CDF	Cumulative Distribution Function		
CNN	Convolutional Neural Network		
CPI	Consumer Price Index		
CR	Coverage Rate		
CREP	Circular Regional Exterior do Porto		
CRISP-DM	Cross-Industry Standard Process for Data Mining	nPIW	Normalised Prediction Interval Width
DL	Deep Learning	PACF	Partial Autocorrelation Function
DP	Data Preparation	PDF	Probability Density Function
DU	Data Understanding	PI	Prediction Interval
EDA	Exploratory Data Analysis	PICP	Prediction Interval Coverage Probability
EFB	Exclusive Feature Bundling	Q1	First Quartile (25th Percentile)
GARCH	Generalized Autoregressive Conditional Het- eroskedasticity	Q2	Second Quartile (Median)
GBDT	Gradient Boosting Decision Trees	ReLU	Rectified Linear Unit
GDP	Gross Domestic Product	RMSE	Root Mean Squared Error
GNN	Graph Neural Network	RMSSE	Root Mean Squared Scaled Error
GOSS	Gradient-based One-Side Sampling	RNN	Recurrent Neural Network
GP	Grande Porto	S3	Amazon Simple Storage Service
GRU	Gated Recurrent Unit	SAP BW	SAP Business Warehouse
ID	Identifier	SARIMA	Seasonal Autoregressive Integrated Moving Average
IMT	Instituto da Mobilidade e dos Transportes	SARIMAX	Seasonal ARIMA with eXogenous regressors
INE	Instituto Nacional de Estatística	SD	Standard Deviation
IT	Information Technology	SQL	Structured Query Language
ITS	Intelligent Transportation Systems	STARIMA	Space-Time Autoregressive Integrated Mov- ing Average
KARIMA	Kohonen Self-Organizing Map ARIMA	SVM	Support Vector Machine
KNN	k-Nearest Neighbors	SVR	Support Vector Regressor
LGBM	Light Gradient Boosting Machine	VCI	Via de Cintura Interna
LSTM	Long Short-Term Memory		
MA	Moving Average		
MAE	Mean Absolute Error		
MASE	Mean Absolute Scaled Error		
ML	Machine Learning		
MLP	Multilayer Perceptron		

Chapter 1

Introduction

In Portugal, investment in road infrastructure has long been regarded as a fundamental driver of economic growth, regional integration, and social development. Highways and road networks play a crucial role in promoting the movement of goods, facilitating access to essential services, and expanding opportunities for individuals and businesses [Pereira and Andraz \(2011\)](#). The development and modernization of road infrastructure have enabled the efficient circulation of people and products, reducing distances and travel times, and thereby fostering competitiveness and attracting new investment [Pereira and Andraz \(2011\)](#).

Beyond the economic sphere, improved road infrastructure has also contributed to raising standards of living by providing greater access to education, healthcare, and leisure opportunities. These positive impacts underscore the centrality of highways as a pillar of modern society, shaping how people live, work, and move [Pereira and Andraz \(2011\)](#).

Nevertheless, the intensification of motorization in recent decades has brought new challenges, including increased dependence on private automobiles, the emergence of congestion in urban and peri-urban areas, and growing environmental impacts [Pereira and Andraz \(2011\)](#). Despite these concerns, it remains clear that road infrastructure continues to be one of the main pillars of development, profoundly influencing both economic performance and the daily lives of citizens.

Within this context, accurate traffic forecasting becomes essential for managing and optimizing these vital infrastructures. This dissertation focuses on the specific case of Ascendi, a Portuguese company specializing in asset management, road infrastructure operation and maintenance, and toll collection services [Ascendi \(2025b\)](#). Ascendi operates six major road concessions across Portugal and has expanded into France and Spain, overseeing highways that connect metropolitan areas, link coastal and inland regions, and form key routes between northern and southern Portugal and the wider European road

networks [Ascendi \(2025c\)](#).

The company's operations are structured around three main business areas: asset management, which encompasses comprehensive oversight of physical infrastructure; the operation and maintenance of road and ITS, including the implementation of advanced maintenance and traffic management protocols; and toll collection services, which utilize both traditional and electronic systems to generate revenue [Ascendi \(2025a\)](#). This thesis is situated within the domain of road operation and ITS, with a focus on forecasting traffic flow for different vehicle classes and analyzing the models performance between urban and suburban highways.



Figure 1.1: Electronic tolling system and price classes on the A41 motorway in Portugal. The sign shows different toll rates by vehicle class [Renascença \(2025\)](#)

1.1 Motivation

For company's like Ascendi, which operate under fixed toll rates regulated by government authorities, accurate forecasting of traffic volumes is the primary determinant of financial performance and operational efficiency.

Existing forecasting methods used by Ascendi rely on relatively simple deterministic models, which, while interpretable and easy to implement, often fall short in capturing the complex, dynamic, and sometimes nonlinear patterns that characterize modern traffic flows. These limitations become particularly evident over longer forecasting horizons and

in the presence of exogenous factors such as public holidays, policy changes, or major local events, that can substantially influence traffic patterns [Li et al. \(2023\)](#).

Furthermore, the diversity of traffic behavior between urban and suburban highways introduces additional complexity. Urban traffic often exhibit higher variability due to factors like daily commuting, population density, and frequent disruptions, while suburban highways may display more regular patterns but remain sensitive to seasonal fluctuations and special events [Salotti et al. \(2018\)](#).

In light of these challenges, there is a clear need for more sophisticated, data-driven forecasting methodologies that can leverage modern machine learning and deep learning techniques, account for external factors, and provide reliable uncertainty estimates. Enhancing the reliability and robustness of long-term traffic forecasts will empower Ascendi's finance and operations teams to make more informed decisions, improve resource allocation, and better manage financial risk.

1.2 Goals

The primary objective of this dissertation is to develop and evaluate a long-term traffic forecasting framework tailored to the operational needs of highway infrastructure management. The research aims to design predictive models capable of accurately forecasting daily traffic volumes by vehicle class for selected urban and suburban highway sections managed by Ascendi. It seeks to analyze and compare traffic flow patterns and forecasting performance in these different environments, and to understand how these contexts influence both prediction accuracy and uncertainty.

This research also integrates pertinent external variables—such as calendars, public holidays, macroeconomic indicators, and significant local events—into the modeling approach. This integration aims to improve the explanatory capacity and clarity of forecasts while assisting predictors in recognizing irregular traffic patterns. By quantifying forecast uncertainty using PI, the project supports risk-aware planning and decision-making. The ultimate goal is to deliver practical recommendations and tools that can be integrated into Ascendi's business processes, improving the reliability of financial and operational planning.

The specific research questions, enumerated below, that guide this dissertation are directly derived from the business objectives outlined in the Business Objectives sub-subsection [3.1.1.2](#). More details on the rationale behind the selection of these questions are provided on the *Business Understanding* section [3.1](#).

1. *What are the projected daily traffic volumes for each vehicle class over the coming 90 days on both urban and suburban highways, and with what level of uncertainty are these projections made?*
2. *How does traffic evolve over time—including seasonal and cyclical variations—and how do these patterns and predictive performance differ between urban and suburban highways?*
3. *Are there observable trends or abnormalities in traffic volumes that could substantially affect revenue projections, thus impacting budget planning and financial risk management?*

1.3 Document Structure

This dissertation is organized to guide the reader through the theoretical foundations, methodological choices, experimental results, and practical implications of the research.

Chapter 2 provides a literature review covering time series analysis, forecasting models, uncertainty quantification, and the current state of traffic forecasting in intelligent transportation systems.

Chapters 3, 4, and 5 constitutes the methodology and results of this work, structured according to the CRISP-DM framework. Chapter 3 (*Study Case*) covers the initial stages of the CRISP-DM process, namely *Business Understanding* (Section 3.1) and *Data Understanding* (Section 3.2). This chapter details the research objectives, the business context, the available data, and its exploratory analysis.

Chapter 4 (*Implementation*) addresses the subsequent CRISP-DM stages: *Data Preparation* (Section 4.1) and *Modeling* (Section 4.2). In this chapter, the data preprocessing steps, feature engineering, modeling techniques, test design, evaluation metrics, and reliability assessments are presented and discussed.

Chapter 5 (*Results and Discussion*) is dedicated to the final phases of the CRISP-DM methodology, namely *Evaluation* (Section 5.1) and *Deployment* (Section 5.2). In this chapter, the performance of the models is evaluated and the findings is discussed. In the deployment section, a simple plan for implementing the model within the company is mentioned.

Finally, Chapter 6 (*Conclusions and Future Work*) summarizes the main findings, discusses the limitations of the study, and suggests directions for future research, with an emphasis on advancing both the academic and practical aspects of long-term traffic forecasting in highway infrastructure management.

Chapter 2

Literature Review

This chapter provides a comprehensive overview of the theoretical foundations, the principal research streams pertinent to the dissertation, and the architectures of the models employed. The foundations address characteristic patterns of traffic data, distinctions between short- and long-term horizons, approaches to uncertainty quantification, contrasts between urban and suburban traffic regimes, and the role of contextual (exogenous) variables. The review of modeling paradigms covers statistical, machine learning, and deep learning approaches. The chapter then details the specific architectures used in this work—ARIMAX, LightGBM, and LSTM—highlighting their design choices and implementation aspects. Finally, it identifies outstanding research gaps and positions the dissertation’s main contributions within the broader domains of traffic forecasting and infrastructure management.

2.1 Time-Series Foundations for Highway Traffic Analysis

When it comes to traffic, using time series means examining data recorded over time to understand and reproduce patterns of vehicle flow on roadways. This form of analysis makes it easier to identify recurring behaviors, seasonal trends, and variations that occur over the day or hour, allowing future traffic volumes to be forecast with much greater accuracy. Time-series analysis procedures—such as autoregression, machine learning, and neural networks—are employed to capture these regularities and calibrate models that can learn from such patterns, supporting the planning and effective management

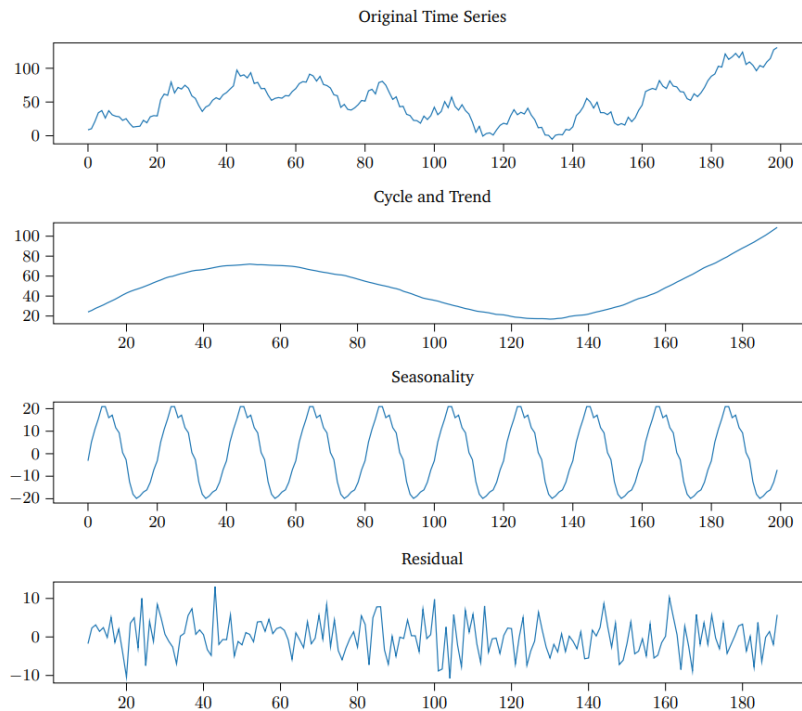


Figure 2.1: A time series with varying characteristics Erlangung (2023).

of transportation systems Lingras et al. (2000). Within the scope of this project, time series played a crucial role in distinguishing traffic particularities across different highway contexts, considering factors such as roadway context and seasonality, and made it possible to compare statistical, machine-learning, and deep learning approaches to projecting future volumes.

The subsections that follow delineate the patterns of traffic data as they pertain to time-series forecasting, define the principal forecasting horizons addressed in the literature for this type of problems, and examine the sources and implications of predictive uncertainty. They further situate traffic prediction within urban and suburban highway contexts and synthesize the external drivers that recent research has identified as influential. Taken together, these topics provide the conceptual foundations for the highway traffic analyses undertaken in this dissertation.

2.1.1 Temporal Structure: Trend, Seasonality, and Cycles

Time series data often exhibit distinct patterns and structures that can be leveraged to improve forecasting and analytical techniques. Among these characteristics, *trend*, *seasonality*, and *cyclic behavior* (see Figure 2.1) play crucial roles in understanding temporal dependencies in the data Hyndman and Athanasopoulos (2021).

A **trend** refers to the long-term progression of a time series, where values tend to increase or decrease over time. Trends may be linear or nonlinear and can fluctuate due to various external influences Hyndman and Athanasopoulos (2021). On traffic analysis, traffic volume on urban roads often exhibits long-term trends due to factors such as population growth, economic development, and changes in transportation infrastructure Wang et al. (2019).

Seasonality denotes patterns that recur at fixed, known intervals, often driven by calendar effects Hyndman and Athanasopoulos (2021). In traffic applications, these periodic fluctuations are evident across many datasets Li et al. (2023), and traffic data are typically characterized by multiple seasonal cycles Mori and Kockelman (2024).

Distinct from seasonality, **cyclic behavior** refers to fluctuations that do not follow a fixed periodicity. These cycles can persist over multiple years and are often linked to broader economic or environmental factors. A well-known example is the business cycle, which consists of alternating periods of economic expansion and contraction Hyndman and Athanasopoulos (2021).

2.1.2 Forecast Horizons: Short- vs Long-Term

The forecasting literature commonly distinguishes traffic prediction into *short-term* and *long-term* categories based on the time span of the projection Manibardo et al. (2021). In practice, however, the boundary between these categories is not consistently defined: while some studies label forecasts made more than one hour ahead as long-term, this convention is not a widely recognized standard.

Most research concentrates on short-term horizons (often ≤ 60 minutes), despite the critical importance of long-term forecasting for infrastructure planning Næss and Strand (2015), logistics planning Lin et al. (2005), management systems, and related decision-making Laña et al. (2018). The literature attributes this imbalance, in part, to challenges in acquiring sufficiently long, high-quality time series and to the greater accumulation of error inherent in longer-range predictions Lingras et al. (2000). Empirical evidence further indicates a systematic degradation in accuracy as the forecast horizon increases Vlahogianni et al. (2004); van Hinsbergen et al. (2007). Whereas many recent studies rely on roughly two years of historical data Lingras et al. (2000), emerging datasets now span five years or more Li et al. (2023), opening opportunities for more rigorous investigations into long-term traffic prediction.

Within this context, the present dissertation adopts a long-term perspective: it produces daily forecasts over a 90-day horizon.

2.1.3 Uncertainty

One of humanity's enduring ambitions is to devise reliable approaches for anticipating uncertain future outcomes. From a theoretical perspective, forecasts are most appropriately formulated in probabilistic terms, describing complete probability distributions over forthcoming quantities or events [Gneiting \(2008\)](#). Nevertheless, the forecasting literature and practice have long focused predominantly on *point forecasts*, which provide only a single numerical estimate. In recent years, however, *probabilistic forecasting* has advanced well beyond its initial applications to binary outcomes and is now increasingly applied across a broader spectrum of predictive tasks [Gneiting \(2008\)](#). Yet, despite this progress—which has opened new avenues for advances in forecasting science—probabilistic predictions remain seldom reported in the traffic forecasting literature.

In contrast to point forecasts, probabilistic methods generate richer forms of output—such as prediction intervals, quantiles, or even full probability distributions—allowing for a more nuanced representation of future uncertainty (see [Figure 2.2](#)). In what follows, it will be provided a more detailed discussion of *prediction intervals*, which serves as a foundation for the present study.

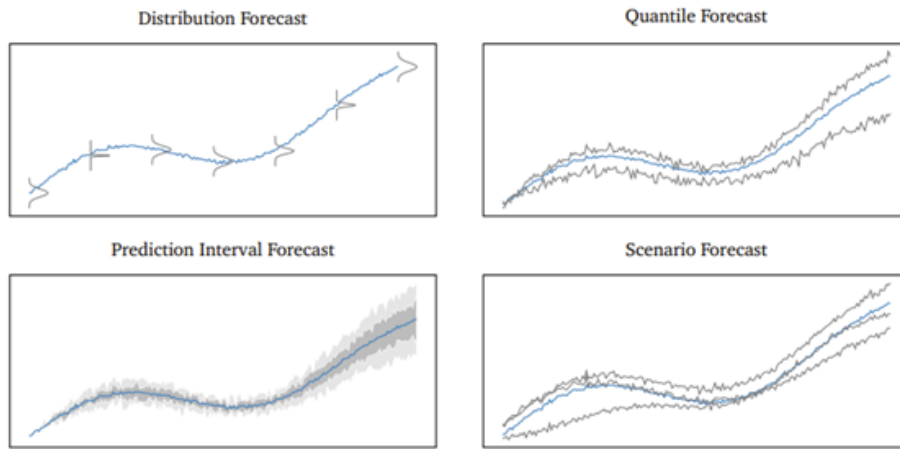


Figure 2.2: Illustration of probabilistic forecasting outputs: distribution forecast, quantile forecast, prediction interval forecast, and scenario forecast.

2.1.3.1 Prediction Intervals

A prediction interval provides an estimated range within which future observations are expected to fall with a specified probability, such as 95% [Nikulchev and Chervyakov \(2023\)](#). Formally, for a given confidence level $1 - \alpha$, the prediction interval at time t is defined as $[l_t, u_t]$ such that:

$$P(l_t \leq y_t \leq u_t) = 1 - \alpha$$

where y_t is the observed traffic value, and l_t and u_t are the lower and upper bounds of the interval, respectively.

Prediction intervals are particularly valuable in traffic forecasting, as they provide stakeholders with a range of possible outcomes, capturing both best- and worst-case scenarios instead of relying solely on single-point estimates. This richer representation of uncertainty supports more informed, risk-aware decisions in revenue management and operational planning. For these reasons, this type of forecasting was selected in the present study.

2.1.3.2 Evaluating Prediction Intervals

In the broader forecasting literature, the boundaries between different notions of uncertainty assessment are not crisply defined [Ho et al. \(2001\)](#); [Zhao et al. \(2008\)](#). Even so, evaluation of PIs commonly emphasizes two complementary properties: *coverage* and *sharpness* [Khosravi et al. \(2011\)](#). Coverage gauges whether the empirical frequency with which observations fall inside the PI matches the nominal confidence level $(1 - \alpha)$, whereas sharpness assesses how concentrated the interval is, typically via its width. For a fixed significance level α , narrower intervals indicate less predictive uncertainty *provided* that coverage remains close to the nominal target [Laña et al. \(2022\)](#).

Consistent with the objectives of this dissertation—namely, comparison of competing forecasting approaches for highway traffic across heterogeneous sites and scales—two simple, model-agnostic metrics are adopted that are widely used in the literature but appear under varying names [Laña et al. \(2022\)](#). For consistency, these are referred to here as *Coverage Rate* and *Normalized Prediction Interval Width*. These criteria enable joint assessment of calibration and sharpness in a manner that is comparable different datasets.

Coverage Rate: The Coverage Rate—also referred to in the literature as PICP—is computed by counting the fraction of realized targets that fall within the constructed prediction intervals (PIs) [Khosravi et al. \(2011\)](#):

$$\text{CR} = \frac{1}{n} \sum_{i=1}^n \mathbb{I}\{l_i \leq y_i \leq u_i\}, \quad (2.1)$$

where l_i and u_i denote the lower and upper PI bounds for observation i , y_i is the realized value, and $\mathbb{I}\{\cdot\}$ is the indicator function. A well-calibrated model yields $\text{CR} \approx 1 - \alpha$ for a nominal level $(1 - \alpha)$ [Khosravi et al. \(2011\)](#).

Normalized Prediction Interval Width: Coverage is intrinsically linked to interval width: a high CR (or PICP) can be attained trivially by widening the intervals [Khosravi et al. \(2011\)](#). To quantify *sharpness*, the average width of the PIs is summarized by the nPIW, which enables comparisons across datasets and sites with different scales—an essential requirement in this study:

$$\text{nPIW} = \frac{1}{n} \sum_{i=1}^n \frac{u_i - l_i}{\bar{y}}, \quad (2.2)$$

where $\bar{y}$ is the sample mean of the observed series. Lower nPIW indicates tighter (sharper) intervals. Because nPIW does not account for calibration, interpretation is made jointly with CR, and models are preferred when a balance between nominal coverage and sharpness is achieved.

2.1.4 Urban–Suburban Traffic Regimes

From a comparative perspective, the literature predominantly relies on a single dataset or on multiple sources of the *same* type (e.g., exclusively highway data or exclusively urban networks), rather than integrating urban and suburban regimes within a unified analytical framework [Manibardo et al. \(2021\)](#). Direct, side-by-side comparisons across regimes remain relatively scarce.

Where this contrast is examined, urban and suburban traffic differ markedly in both volume and data complexity. On highways, segments embedded in urban areas typically carry higher volumes [Lingras et al. \(2000\)](#), whereas in suburban contexts traffic often intensifies during the summer months [Lingras et al. \(2000\)](#). Moreover, urban highways tend to exhibit less pronounced seasonal variability than suburban ones [Lingras et al. \(2000\)](#). Because more stable series are generally easier to predict, these patterns usually translate into higher forecasting accuracy—an expectation carried through this dissertation [Lingras et al. \(2000\)](#).

In this study, a comparative analysis is conducted using two highway datasets representing distinct regimes: (i) a site located at an approach to an urban sphere (treated here as the *urban* dataset) and (ii) a site situated in a *suburban* context. This design enables a controlled assessment of how regime-specific characteristics—volume levels, seasonal structure, and volatility—affect both point forecasts and prediction intervals.

2.1.5 Contextual Exogenous Variables

Contextual—or exogenous—information can materially enhance traffic forecasting by encoding circumstances that, while not direct measurements of the traffic state, nonetheless shape it (e.g., temporal, meteorological, social, and infrastructural factors) [Manibardo et al. \(2021\)](#). A widely used and effective example is calendar metadata [Mori and Kockelman \(2024\)](#); [Manibardo et al. \(2021\)](#); [Li et al. \(2023\)](#), often discretized into categories such as [weekday, Saturday, Sunday] to reflect systematic differences between working days and weekends. The literature further highlights the value of *episodic drivers* (terminology in [Manibardo et al. \(2021\)](#))—such as sporting events, incidents, and roadworks—which help models recognize and learn atypical regimes that deviate from routine patterns.

From a modeling standpoint, these contextual variables are represented as time-indexed inputs aligned with the target series and can be ingested in a similarly by both traditional machine-learning and deep learning approaches. In practice, they are incorporated either as additional covariates (e.g., categorical calendar indicators, continuous weather measurements) or through data-fusion components that learn compact representations of heterogeneous inputs within the model architecture [Manibardo et al. \(2021\)](#). This becomes especially important for extended forecasting horizons where exogenous drivers increasingly govern the dynamics as the prediction window widens.

2.2 Modeling Paradigms

This section synthesizes prior research on traffic forecasting across three paradigms: (i) statistical models, (ii) machine learning models, and (iii) deep learning models. This taxonomy, originally proposed by [Shaygan et al. \(2022a\)](#), has been widely adopted in subsequent studies. Within each group, the most commonly used models are summarized alongside their reported performance, principal advantages, and limitations. The discourse builds upon the latest survey detailed in [Carianni and Gemma \(2025\)](#) and is further supported by additional complementary research. It is important to note that the literature is overwhelmingly concentrated on *short-term* horizons—covering more than 90% of published studies according to [Carianni and Gemma \(2025\)](#)—which complicates a systematic assessment of *long-term* forecasting. To provide a balanced perspective, the section therefore reviews both strands: short-term approaches (reflecting the dominant body of work) and long-term approaches (which remain comparatively underexplored but are central to the objectives of this dissertation).

2.3 Statistical Models

Statistical models rely on mathematically defined structures and explicit assumptions about the data-generating process. Classical autoregressive models—including ARIMA, its seasonal extension SARIMA, and the exogenous-input variant ARIMAX (frequently highlighted in prior studies)—constitute the core of this paradigm [Hamed et al. \(1995\)](#); [Williams and Hoel \(2003\)](#); [Williams et al. \(1998\)](#); [Williams \(2001\)](#). Additional representatives include exponential smoothing families (Holt–Winters) [Williams et al. \(1998\)](#), Gaussian maximum-likelihood state-space models, and Kalman filter formulations [Okutani and Stephanedes \(1984\)](#). The bulk of this work has been developed and evaluated for *short-term* horizons.

Hybrid approaches have also appeared, combining statistical and learning-based components (SARIMA–GARCH for volatility modeling [Guo et al. \(2014\)](#)), with the goal of capturing distinct signal attributes within a unified pipeline. Unlike *long-term* traffic forecasting, which is not extensively tackled, it is predominantly analyzed with ARIMA models for result comparison [Carianni and Gemma \(2025\)](#).

While these statistical models often yield lower accuracy than modern machine-learning and deep-learning methods, they offer several advantages: relatively modest data requirements, transparent structure, and well-understood diagnostics [Carianni and Gemma \(2025\)](#). ARIMA-family models, in particular, explicitly encode assumptions about stationarity (with integration to address nonstationarity), admit seasonal components, and allow exogenous regressors via ARIMAX—features that remain attractive when data are limited or interpretability is paramount [He \(2024\)](#).

2.4 Machine Learning Models

Time series forecasting can be reformulated as a supervised regression problem in which lagged observations are used as input features and the future target as the response [Aufaure and Zimányi \(2013\)](#). Under this view, standard ML techniques learn mappings from past windows to future values, with optional inclusion of exogenous variables (e.g., weather, calendar indicators, incidents) to provide context that improves predictive performance [Aufaure and Zimányi \(2013\)](#).

Recent work documents a steady shift toward ML approaches for traffic forecasting, motivated by their capacity to model complex nonlinear relationships and their adaptability to evolving network conditions [Shaygan et al. \(2022b\)](#). Representative algorithms include KNN, SVM, SVR, and gradient-boosted decision trees such as LightGBM [Ke et al.](#)

(2017b); Liu (2021). These methods capture nonlinear effects and interactions effectively and scale well to large datasets.

Across comparative studies, ML models frequently surpass classical statistical baselines in point forecasting accuracy Shaygan et al. (2022b). This advantage often entails greater computational demands than traditional statistical models, particularly during training and hyperparameter search. Nonetheless, the required resources are generally scalable, and model complexity remains modest relative to deep learning approaches Shaygan et al. (2022b). A persistent limitation, however, is that performance can be *limited and inconsistent* across applications due to reduced representational capacity and a reliance on manual feature engineering, which may fail to capture complex dependencies in traffic regimes.

2.5 Deep Learning Models

Deep learning has become a central paradigm in traffic flow forecasting, with studies spanning MLP, RNNs; notably LSTM and GRU, CNN, GNN, and a variety of hybrid architectures Carianni and Gemma (2025). RNN-based approaches consistently demonstrate strong performance for sequence modeling; LSTM and GRU architectures are frequently reported as state-of-art for short-term forecasting Fu et al. (2016) and remain competitive at longer horizons when sufficient contextual information is available.

Hybrid designs that couple CNNs with LSTM/GRU units leverage spatial feature extraction prior to temporal aggregation, yielding measurable gains in settings with localized spatial structure Zheng et al. (2021). Attention mechanisms—applied on top of recurrent or convolutional backbones—further improve accuracy by adaptively weighting salient time steps or spatial neighborhoods, with several studies reporting enhanced performance in the presence of complex regime shifts and nonstationary patterns Carianni and Gemma (2025).

In general, recent studies consistently report that deep learning models achieve superior predictive performance for traffic flow forecasting Shaygan et al. (2022b). Notwithstanding these gains, several limitations persist. Relative to other machine learning approaches, deep architectures require more complex data representations and substantially greater computational resources for training and tuning. In addition, many DL models operate as *black boxes*, offering limited practical interpretability compared with traditional statistical or simpler ML methods.

2.6 Deep Description of Architectures Used

This section provides a detailed account of the model architectures adopted in this study. It begins with the ARIMA family, representing the statistical paradigm, proceeds to the LightGBM architecture as the machine learning representative, and then presents the LSTM network as the deep learning model. The section concludes by specifying the evaluation metrics used to assess point forecasting performance across all models.

2.6.1 ARIMA Family

The ARIMA model, denoted $\text{ARIMA}(p, d, q)$, consists of three components: autoregressive (AR) terms of order p , integration (I) of order d (differencing), and moving average (MA) terms of order q . Following the notation in [Hyndman and Athanasopoulos \(2021\)](#), a non-seasonal $\text{ARIMA}(p, d, q)$ process is defined by

$$\Phi_p(B) (1 - B)^d y_t = \Theta_q(B) \varepsilon_t, \quad (2.3)$$

where

- B is the *backshift* operator ($By_t = y_{t-1}$);
- $\Phi_p(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$ is the autoregressive (AR) polynomial of order p ;
- $\Theta_q(B) = 1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q$ is the moving-average (MA) polynomial of order q ;
- ε_t is a white-noise innovation sequence.

The differencing operator $(1 - B)^d$ removes low-order polynomial trends, rendering the transformed series stationary before the AR and MA components are applied. Stationarity and invertibility are guaranteed when the roots of $\Phi_p(B)$ and $\Theta_q(B)$ respectively lie outside the unit circle.

2.6.1.1 SARIMA

In contrast, when the data exhibit clear seasonal patterns—such as regular weekly, monthly, or yearly cycles—the Seasonal ARIMA (SARIMA) model is more appropriate [Hyndman and Athanasopoulos \(2021\)](#). SARIMA extends ARIMA by explicitly modeling these repeating seasonal effects, which are really common in traffic data [Mori and Kockelman \(2024\)](#). SARIMA is denoted as $\text{ARIMA}(p, d, q) \times (P, D, Q)_s$, where P , D , and Q are the seasonal AR, differencing, and MA orders, and s is the seasonal period. The SARIMA model is expressed as:

$$\Phi_p(B)\Phi_P(B^s)(1-B)^d(1-B^s)^D y_t = \Theta_q(B)\Theta_Q(B^s)\varepsilon_t, \quad (2.4)$$

where the seasonal polynomials and operators capture the repeating patterns in the data [Hyndman and Athanasopoulos \(2021\)](#).

2.6.1.2 ARIMAX (and SARIMAX)

In complex systems, like traffic forecasting, external factors exert considerable influence [Manibardo et al. \(2021\)](#). The ARIMAX model is appropriate in these scenarios as it incorporates external variables, like holidays in traffic studies, directly into its framework, as explained by [Majka \(2024\)](#). From a mathematical perspective, $x_{t,k}$, where k ranges from 1 to r , denote the current or lagged exogenous variables. The non-seasonal ARIMAX(p, d, q) model is then formulated as follows:

$$\Phi_p(B)(1-B)^d y_t = \sum_{k=1}^r \beta_k x_{t,k} + \Theta_q(B)\varepsilon_t, \quad (2.5)$$

.

If seasonality of period s is also present, the formulation extends to SARIMAX(p, d, q) $\times$ (P, D, Q) $_s$:

$$\Phi_p(B)\Phi_P(B^s)(1-B)^d(1-B^s)^D y_t = \sum_{k=1}^r \beta_k x_{t,k} + \Theta_q(B)\Theta_Q(B^s)\varepsilon_t. \quad (2.6)$$

The coefficients β_k measure the marginal impact of each regressor after the series has been differenced and filtered by the ARMA operators. Stationarity and invertibility requirements mirror those of ARIMA/SARIMA.

2.6.2 LightGBM

LightGBM—originally introduced by [Ke et al. \(2017a\)](#)—is a state-of-the-art implementation of Gradient Boosting Decision Trees designed for high-throughput learning on large-scale and potentially sparse datasets. It preserves the strong predictive performance of traditional GBDT methods while incorporating algorithmic and system-level optimizations that significantly reduce both training time and memory consumption.

2.6.2.1 GBDT Recap

Given a differentiable loss function $\mathcal{L}(y, F_{m-1}(x))$, classical GBDT methods iteratively refine an ensemble predictor according to:

$$F_m(x) = F_{m-1}(x) + \eta h_m(x), \quad h_m(x) = \arg \min_{h \in \mathcal{T}} \sum_{i=1}^N g_i h(x_i),$$

where $g_i = \partial \mathcal{L} / \partial F|_{(y_i, F_{m-1}(x_i))}$ are the negative gradients, $\eta \in (0, 1]$ is the learning rate, and $\mathcal{T}$ represents the space of regression trees.

At each iteration, the algorithm must (i) construct histograms for all candidate features and (ii) search for split points that maximize the expected reduction in loss—operations that dominate computational cost when the dataset contains a large number of samples (N) or features (d).

2.6.2.2 LightGBM Innovations

To accelerate training, LightGBM introduces several key innovations that optimize both computational and memory efficiency. Figure 2.3 provides a schematic overview of the LightGBM model and its key architectural principles.

Gradient-based One-Side Sampling GOSS accelerates training by keeping the top $a\%$ of samples with the largest gradients and uniformly sampling $b\%$ from the rest. The sampled instances are reweighted by $(1-a)/(1-b)$ to preserve unbiased split estimation. Since large gradients correspond to more informative samples, this reduces histogram construction to $O((a+b)N)$ with negligible accuracy loss.

Exclusive Feature Bundling EFB exploits sparsity by grouping mutually exclusive features into single pseudo-features through a greedy graph-coloring algorithm with $O(d \log d)$ complexity. This significantly lowers the effective dimensionality $\tilde{d} \ll d$, cutting both memory use and split-search time.

Leaf-wise Tree Growth Instead of growing trees level-wise, LightGBM expands the leaf offering the greatest loss reduction. This yields deeper, narrower trees that capture stronger patterns with fewer leaves, while overfitting is controlled via `max_depth` or `num_leaves`.

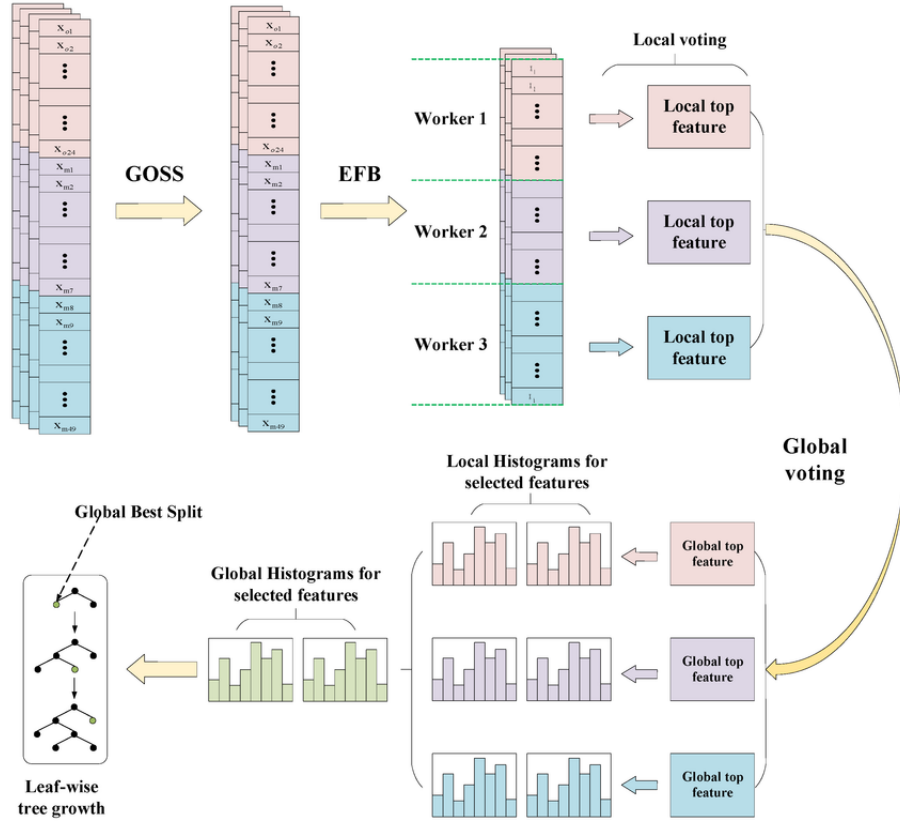


Figure 2.3: Schematic representation of the LightGBM model Yao et al. (2022).

2.6.3 LSTM

LSTM is a type of recurrent neural network (RNN) specifically designed to address the vanishing gradient problem that often affects traditional RNNs. By introducing a memory cell and a set of gating mechanisms, the LSTM enables more stable and effective learning of long-range temporal dependencies in sequential data.

2.6.3.1 RNN Recap

A Recurrent Neural Network (RNN) is a foundational architecture for modeling sequential data in which the order of observations is crucial Yin et al. (2021). RNNs process data across multiple time steps, making them particularly suitable for time series modeling Carianni and Gemma (2025). These networks maintain a hidden state that propagates information over time, allowing them to capture temporal dependencies within the data.

At each time step t , the hidden state h_t is updated according to:

$$h_t = \phi(W_h h_{t-1} + W_x x_t + b), \quad (2.7)$$

where x_t represents the input at time t , h_{t-1} is the previous hidden state, W_h and W_x are learnable weight matrices, b is a bias term, and $\phi(\cdot)$ is a nonlinear activation function, typically tanh or ReLU. The network output y_t is computed as $y_t = W_y h_t + c$, where W_y and c denote the output weights and bias. Figure 2.4 illustrates the standard RNN architecture.

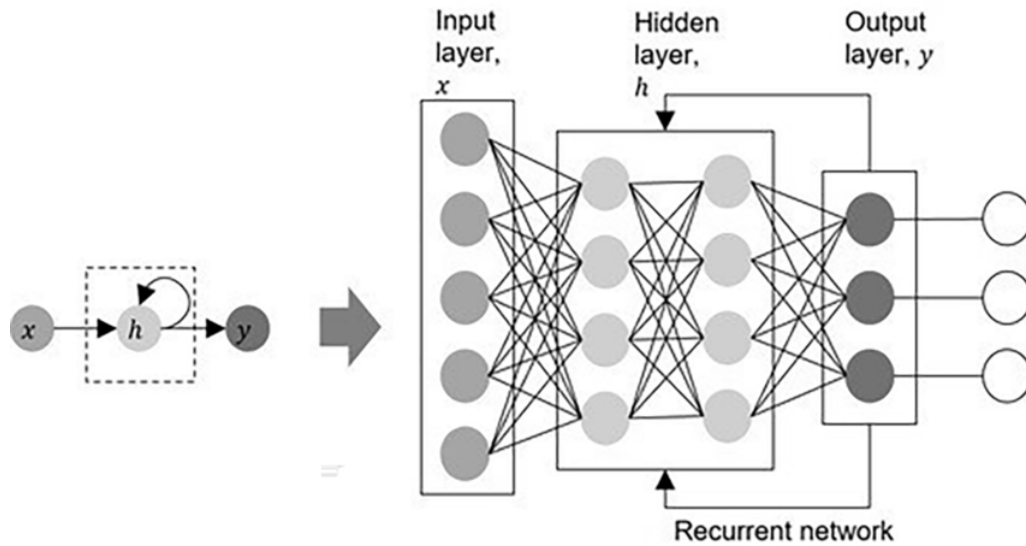


Figure 2.4: Recurrent Neural Network architecture.

Although RNNs are theoretically capable of modeling both short- and long-term dependencies, in practice they often suffer from the vanishing and exploding gradient problems during training. These issues hinder the ability of the network to retain information over long time spans. The LSTM architecture [Hochreiter and Schmidhuber \(1997\)](#) was developed to overcome these limitations by incorporating a memory cell and three gating mechanisms—namely, the *input*, *forget*, and *output* gates—which regulate the flow of information and preserve relevant context over time.

The LSTM cell is updated at each time step t according to:

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \quad (2.8)$$

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \quad (2.9)$$

$$\tilde{c}_t = \tanh(W_c[h_{t-1}, x_t] + b_c) \quad (2.10)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \quad (2.11)$$

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \quad (2.12)$$

$$h_t = o_t \odot \tanh(c_t) \quad (2.13)$$

where x_t is the input vector at time t , h_{t-1} and c_{t-1} denote the previous hidden and cell states, f_t , i_t , and o_t are the forget, input, and output gates, $\tilde{c}_t$ is the candidate cell state, W_* and b_* are learnable parameters, $\sigma(\cdot)$ denotes the sigmoid activation function, and $\odot$ represents element-wise multiplication.

Through this gated mechanism, LSTMs effectively mitigate the vanishing gradient problem and enable the modeling of both short- and long-term temporal dependencies [Hochreiter and Schmidhuber \(1997\)](#). They have achieved state-of-the-art performance in various time series forecasting applications, including traffic flow prediction [Fu et al. \(2016\)](#); [Liu et al. \(2017\)](#); [Zheng et al. \(2021\)](#). Moreover, recent research has extended the standard LSTM architecture with additional layers, attention mechanisms, and convolutional components to further enhance its modeling capacity, interpretability, and robustness in complex temporal domains.

2.6.4 Evaluation Protocols and Metrics

This subsection details the metrics used to assess point forecasting performance.

In the traffic flow literature, the most common error measures are the RMSE and the MAE [Shaygan et al. \(2022a\)](#), which quantify, respectively, the root-mean-square discrepancy and the average absolute discrepancy between observed values and model predictions. However, direct use of RMSE/MAE complicates cross-dataset comparisons because these metrics are scale-dependent. Given that this dissertation compares models across distinct highway datasets and regimes, scale-free criteria are adopted. Specifically, the MASE, proposed by [Hyndman and Koehler \(2006\)](#), and the RMSSE, introduced in [Makridakis et al. \(2022\)](#) are employed; both compare model errors to a naïve seasonal benchmark, enabling fair comparison across datasets.

2.6.4.1 Mean Absolute Scaled Error (MASE).

The MASE compares the mean absolute forecast error to the in-sample mean absolute error of a naïve seasonal forecast:

$$\text{MASE} = \frac{\frac{1}{n} \sum_{t=1}^n |y_t - \hat{y}_t|}{\frac{1}{n-m} \sum_{t=m+1}^n |y_t - y_{t-m}|}, \quad (2.14)$$

where y_t is the observed value at time t , $\hat{y}_t$ is the corresponding forecast, m denotes the seasonal period (e.g., $m=7$ for weekly seasonality in daily data), and n is the sample size.

Values $\text{MASE} < 1$ indicate improvement over the naïve seasonal benchmark, while $\text{MASE} > 1$ indicates inferior performance. In nonseasonal settings, $m=1$ reduces the denominator to the mean absolute first difference.

2.6.4.2 Root Mean Squared Scaled Error (RMSSE).

The RMSSE analogously scales the root mean squared forecast error by the root mean squared error of the naïve seasonal benchmark:

$$\text{RMSSE} = \sqrt{\frac{\frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2}{\frac{1}{n-m} \sum_{t=m+1}^n (y_t - y_{t-m})^2}}. \quad (2.15)$$

As with MASE, values below one denote improvement over the benchmark. RMSSE is more sensitive to large deviations than MASE, providing complementary insight when occasional large errors are operationally critical.

Together, MASE and RMSSE offer a balanced, scale-robust evaluation protocol that facilitates comparisons across datasets, horizons, and regimes, and reflects both typical errors and sensitivity to outliers.

2.7 Gaps and Contributions

Despite substantial progress in time series forecasting, notable gaps persist in the domain of *long-term* traffic prediction. The literature remains heavily concentrated on *short-term* horizon [Manibardo et al. \(2021\)](#)—frequently limited to windows under 60 minutes—thereby restricting generalizable insights for multi-week or multi-month planning. The treatment of predictive *uncertainty* is comparatively underdeveloped [Nikulchev and Chervyakov \(2023\)](#): although uncertainty quantification has advanced in adjacent fields, it is less frequently integrated into traffic studies [Nikulchev and Chervyakov \(2023\)](#). Recent work has also emphasized increasingly complex deep learning architectures [Carianni and Gemma \(2025\)](#), often at the expense of interpretable approaches that facilitate interpretability, maintenance, and large-scale deployment. In addition, many studies rely on datasets spanning at most two years [Wickramasuriya et al. \(2018\)](#), which constrains empirical analysis of long-term horizons; only a limited number of emerging datasets extend to five or more years [Li et al. \(2023\)](#). Finally, comparative assessments across operational regimes remain scarce, as most studies use a single source or multiple sources of the same scope [Lingras et al. \(2000\)](#), limiting evidence on cross-regime (e.g., urban versus suburban highways).

The present research addresses these limitations by developing and evaluating a long-horizon forecasting framework that generates *daily* predictions over a three-month (90-day) horizon; by reporting both point forecasts and *prediction intervals* to enable risk-aware planning; by conducting a controlled, side-by-side comparison of urban and suburban highway contexts to clarify how regime-specific characteristics affect accuracy and uncertainty; and by systematically incorporating exogenous information—including calendar metadata, episodic drivers (incidents, special events, planned works), and macroeconomic signals. A further contribution is the deliberate emphasis on deployability and interpretability, privileging modeling choices that balance predictive performance with transparency, data requirements, and computational cost so as to support revenue and operational planning at scale in highway networks.

In sum, the study advances both methodological understanding and practical applicability through a long-horizon, uncertainty-aware, and context-sensitive forecasting solution tailored to real-world traffic management scenarios.

Chapter 3

Case Study

This chapter presents the initial stages of the CRISP-DM methodology, namely *Business Understanding* (Section 3.1) and *Data Understanding* (Section 3.2). These stages are primarily concerned with defining the problem itself and conducting a preliminary exploration of the data, preceding the subsequent phases directly related to model implementation.

In the BU section, the background of the company and the motivation for the development of this project are first outlined. The business objectives and corresponding success criteria are then discussed, followed by the formulation of data mining goals and their respective technical success metrics.

In the DU section, the datasets collected for this study are described, and the rationale for their selection is provided. An exploratory data analysis (EDA) is performed to obtain deeper insights into the data. Furthermore, the quality of the data is assessed, with a detailed analysis of any anomalous behaviors or irregularities identified within the dataset.

Overall, this chapter provides a solid foundation for the subsequent phases of the CRISP-DM process by thoroughly addressing the context, objectives, and data characteristics underpinning this research.

3.1 Business Understanding

This initial phase focuses on understanding the project objectives and requirements from a business perspective and translating this knowledge into a formal data mining problem definition, along with a preliminary plan to achieve these objectives [Chapman \(2000\)](#).

The following section outlines the main activities, techniques, and considerations associated with the Business Understanding phase, as depicted in Figure 3.1. Although,

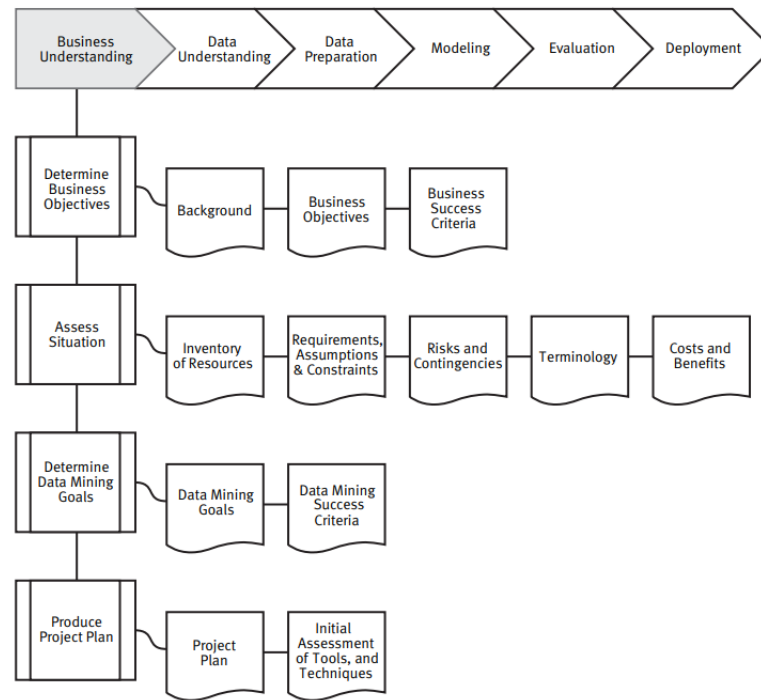


Figure 3.1: Detailed Business Understanding, [Chapman \(2000\)](#)

to maintain the focus of this dissertation, only two steps—*Determine Business Objectives* (subsection 3.1.1) and *Determine Data Mining Goals* (subsection 3.1.2)—will be addressed in detail.

3.1.1 Determine Business Objectives

The main purpose of this subsection is to define what the organization aims to achieve from a business perspective. It begins by outlining the background of the project, followed by a definition of the specific business objectives. Finally, it establishes the criteria by which the success of the project will be evaluated.

3.1.1.1 Background

The primary motivation for this project is to enhance the precision of revenue planning, which is a critical concern for Ascendi's Finance Department. Given that toll fees are regulated, fluctuations in traffic volume are the primary contributors to revenue variability.

Ascendi presently utilizes a regression modeling strategy to predict traffic volumes and toll revenues. This approach depends on historical traffic data broken down by concession, road segment ("*sublanço*"), vehicle category, year, and month. Furthermore, macroeconomic factors like GDP and the CPI, both past and projected, are employed

to formulate various economic scenarios, typically labeled as optimistic or pessimistic. The predictions are produced using linear and logarithmic regression models, applied via forecasting packages in the R programming language.

The current method provides several benefits. It is simple and easy to understand, which makes it approachable for stakeholders throughout the company. Additionally, it includes scenario planning that allows the Finance Department to investigate possible future results under various economic circumstances. Furthermore, the process is well-defined and aligns with the organization's operational and regulatory standards. However, several limitations have been identified. The model's fundamental design impacts its precision in forecasting, particularly in capturing complex or non-linear traffic patterns. It struggle in effectively incorporating temporal dynamics, such as daily and weekly seasonality or abrupt shifts due to external events. With forecasts produced monthly, there's a risk of missing short-term changes critical for tactical decision-making. Moreover, the model depends on a narrow set of input variables, focusing largely on historical traffic data and a few economic indicators, while overlooking other potentially important factors like public holidays or local events. In addition, there is a lack of automated evaluation tools, and the model does not systematically track forecast accuracy over time, requiring significant manual effort for updates and validation.

To address these limitations, Ascendi is implementing an enhanced, data-focused strategy. The goal of this data mining project is to explore and utilize modern forecasting techniques capable of handling increasingly complex datasets, identifying intricate temporal patterns, and improving prediction precision. This initiative aims to provide the Finance Department with more dependable and adaptable tools for revenue planning.

3.1.1.2 Business Objectives

The primary objective of this project is to develop accurate forecasts of traffic volumes, categorized by vehicle class for two highways—one situated in an urban area and the other in a suburban context—according to official classifications (refer to appendix A for additional information). These predictions are specifically designed to assist the financial team, allowing for more accurate toll revenue forecasts. Although toll rates are set by government regulations and are not subject to change, fluctuations in traffic volumes are a crucial element affecting monthly revenue and, consequently, financial planning.

Within this context, the project aims to address the following strategic questions for the Finance Department:

1. *What are the projected daily traffic volumes for each vehicle class over the coming 90 days on both urban and suburban highways, and with what level of uncertainty*

are these projections made?

2. *How does traffic evolve over time—including seasonal and cyclical variations—and how do these patterns and predictive performance differ between urban and suburban highways?*
3. *Are there observable trends or abnormalities in traffic volumes that could substantially affect revenue projections, thus impacting budget planning and financial risk management?*

By answering these questions, the Finance Department can make more informed decisions, such as improving budget accuracy and better managing financial risks.

3.1.1.3 Business Success Criteria

To achieve these goals, the forecasting models need to exhibit strong robustness and reliability, forming a solid foundation for financial decisions over long-term horizons. The forecasts must be detailed enough to distinguish between various vehicle classes and account for each highway's distinct traits. Additionally, the models should have the ability to predict traffic variations that might impact the consistency of revenue forecasts.

This initiative is anticipated to significantly enhance the precision of toll revenue predictions, consequently improving the reliability of financial planning. With a better understanding of traffic trends and their changes over time, the organization will be more adept at crafting realistic budgets and identifying potential financial risks due to deviations in advance.

3.1.2 Determine Data Mining Goals

This subsection specifies the objectives the organization intends to accomplish from a data mining perspective. It starts by detailing the goals of data mining and then describes the criteria for assessing the project's success.

3.1.2.1 Data Mining Goals

To align the business objective of enhancing toll revenue prediction with the technical requirements of data-driven modeling, this project establishes a set of data mining goals. These objectives are organized into three complementary dimensions: forecasting, analytical exploration, and model evaluation.

1. Forecasting:

- Develop predictive models capable of reliably forecasting traffic volumes for each vehicle class on both urban and suburban highways, with a forecasting horizon of up to three months.
- Incorporate probabilistic forecasting methods to generate prediction intervals, thereby quantifying the uncertainty associated with future traffic volumes.

2. Analytical Exploration:

- Apply time series decomposition techniques to capture temporal structures such as seasonal effects and weekly variations.
- Detect anomalies in traffic volumes, potentially linked contextual external events.
- Compare traffic dynamics across urban and suburban contexts to identify systematic differences in travel behavior.

3. Model Evaluation:

- Establish robust evaluation strategies to assess forecast accuracy, employing metrics appropriate for both point forecasts and prediction intervals.
- Assess the implementation time of each model.

In summary, the project addresses core data mining challenges in time series forecasting and multivariate analysis. By combining forecasting, exploratory analysis, and rigorous evaluation, the study not only supports more accurate toll revenue predictions but also provides insights into broader traffic dynamics across distinct environments.

3.1.2.2 Data Mining Success Criteria

The data-mining stage will be assessed using an integrated set of quantitative performance metrics and qualitative evaluations, so that the forecasting models produced are not only methodologically sound but also decision-relevant for Ascendi's financial planning. The assessment criteria are enumerated and discussed below:

1. **Predictive Accuracy:** The models should outperform Ascendi's current forecasting approach. Since no established quantitative benchmarks are currently in use within the company, performance will be evaluated through expert assessment and comparison with practical expectations, complemented whenever possible by standard accuracy metrics.

2. **Horizon Robustness:** A successful model must provide reliable daily forecasts up to three months ahead. Consistent results under varying traffic patterns will indicate robustness and the ability to generalize to real-world conditions.
3. **Uncertainty Quantification and Anomaly Awareness:** Beyond point forecasts, the models should quantify forecast uncertainty through prediction intervals (evaluated using the CR and nPIW metrics). They must also learn traffic anomalies associated with contextualized events, such as holidays or weekends.
4. **Operational Readiness and Integration;** For operational success, the models should be computationally efficient, capable of regular retraining, and easily integrated into Ascendi's existing IT infrastructure.

In summary, success will be judged not only by predictive accuracy but also by the model's capacity to generalize across horizons, to quantify uncertainty credibly, to detect context-driven anomalies, and to deliver outputs that are reliable and operationally usable within Ascendi's workflow.

3.2 Data Understanding

The Data Understanding phase follows the formulation of initial project objectives. Its primary purpose is to acquire familiarity with the data that will be used to address the business problem, assess its quality, and uncover preliminary insights. This phase involves collecting, exploring, and analyzing data from various sources to better comprehend its characteristics, structure, and potential issues.

During Data Understanding process, a series of systematic activities are undertaken, as described in Figure 3.2. These include data collection and description, where the available datasets are examined to identify their contents, formats, and relevance. Exploratory data analysis techniques such as statistical summaries and tests, and visualizations are employed to detect patterns, outliers, missing values, and anomalies. Additionally, this phase involves verifying data quality and uncovering relationships and potential correlations within the data.

The following section details the specific activities, techniques, and considerations involved in the Data Understanding phase.

3.2.1 Collect Initial Data

The initial phase of the project consisted of identifying and acquiring the datasets required to support the development of traffic volume forecasting models. Three main sources of

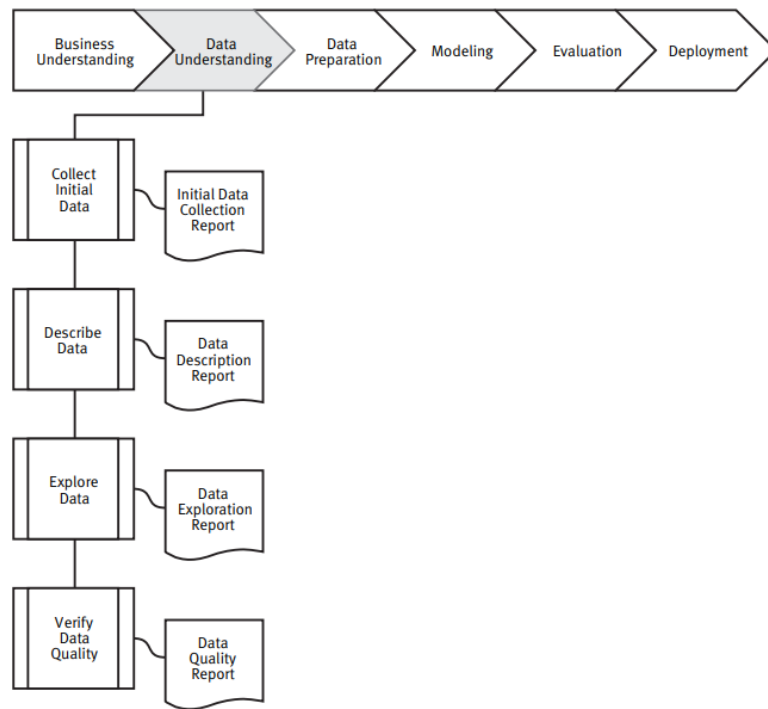


Figure 3.2: Detailed Data Understanding, [Chapman \(2000\)](#)

data were selected: (i) traffic flow data from Ascendi's SAP BW system, (ii) macroeconomic indicators from the *Instituto Nacional de Estatística*, and (iii) holiday calendars obtained from a Python package.

3.2.1.1 Traffic Data

Traffic information was obtained from Ascendi's SAP BW system, which consolidates vehicle passages through toll collection points ("*pontos de cobrança*"). While traffic data are originally collected by physical sensors, the SAP BW platform provides an advantage by reducing the reporting lag from approximately one month to a single day. This timeliness made it the preferred source for the project.

The raw dataset consisted of approximately 1.1 billion records at minute-level granularity, each representing an individual vehicle passage. To ensure consistency and facilitate analysis, redundant records generated by system versioning were removed, and the data were aggregated to a daily frequency. The aggregation was restricted to concessions and road segments relevant to the study, yielding two datasets—one representing an urban environment and the other a suburban environment—with approximately 1,200 daily entries each, covering the period from January 1, 2022, to May 11, 2025.

For the urban environment, the *Freixieiro–Aeroporto* section of the *Grande Porto* concession (A41 highway) was selected due to its high traffic volume and strategic relevance to

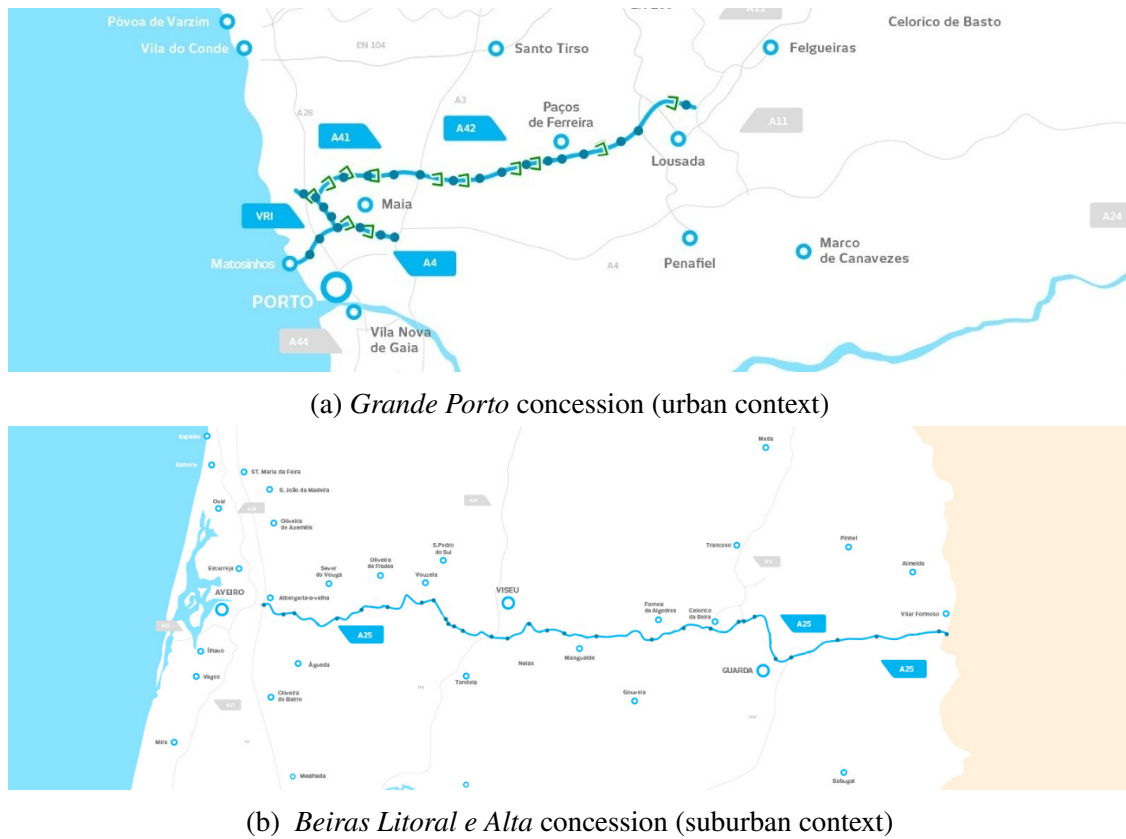


Figure 3.3: Geographical framework of the study areas: (a) urban and (b) suburban concession, Ascendi (2025c).

metropolitan mobility. For the suburban environment, the focus was the *Caçador–Fagilde* section of the *Beiras Litoral e Alta* concession (A25 highway), which similarly registers the highest traffic within its concession. These choices enable a balanced comparison between distinct contexts of demand and usage. The geographical locations of both study areas are illustrated in Figure 3.3.

3.2.1.2 Macroeconomic Data

Complementary macroeconomic indicators were sourced from INE, specifically Gross Domestic Product and the Consumer Price Index. These indicators were included as exogenous variables in the forecasting models to capture the potential impact of broader economic conditions on traffic patterns. The data are up to date as of May 13, 2025, ensuring alignment with the study timeframe.

3.2.1.3 Holiday Data

To account for calendar effects, national holiday dates for Portugal were retrieved using the `holidays` Python library. Movable holidays such as Carnival were computed programmatically to guarantee consistency and completeness of the dataset.

3.2.1.4 Summary and Limitations

The integration of detailed traffic data with macroeconomic indicators and holiday calendars provides a comprehensive foundation for developing accurate and interpretable forecasting models. While traffic data remain the primary endogenous variable, the inclusion of external factors enhances explanatory power, particularly for long-term predictions and the analysis of irregular events.

A limitation of the study, however, is the restricted historical coverage: traffic data are only available from 2022 onwards. This constraint reduces the ability to capture long-term trends, identify rare events, and assess model performance across more diverse conditions. Consequently, while the models are well-suited for recent years, their generalization to pre-2022 or atypical scenarios remains limited.

3.2.2 Data Description

This subsection provides a structured overview of the datasets employed in this research, including their scope, statistical characteristics, and relevance to the forecasting task. The data are grouped into endogenous variables—representing the target of prediction—and exogenous variables, which provide contextual information through economic indicators and calendar effects.

3.2.2.1 Traffic Data

The aggregated traffic dataset contains the following attributes: observation date (`calday`); daily vehicle counts by class (`class_1` to `class_5`); total daily volume (`volume`); and identifiers for the concession (`xbicxzcchcon_id`) and road section (`xbicxzcchroad`). Each record is uniquely defined by the composite key of date, concession ID, and road section.

Vehicle classes follow the official IMT classification (*Appendix A*), with the exception that Ascendi distinguishes motorcycles (Class 5) from private vehicles (Class 1). This separation reflects operational needs, since motorcycles with the Via Verde system receive a 30% toll discount *Via Verde Portugal (2024)*.

The dataset spans from January 1, 2022, to May 11, 2025, and covers two case study environments:

- **Urban:** *Freixieiro–Aeroporto* section of the *Grande Porto* concession (A41 highway).
- **Suburban:** *Caçador–Fagilde* section of the *Beiras Litoral e Alta* concession (A25 highway).

Descriptive Statistics: In the urban area, Class 1 vehicles dominate traffic, with a daily mean of 30,628 vehicles (min: 8,291; max: 48,369). The suburban area also shows Class 1 as the largest segment, averaging 13,232 vehicles per day (min: 1,902; max: 22,580). Total traffic volumes average 36,579 (sd: 7,657) in the urban case and 16,819 (sd: 3,057) in the suburban case.

Table 3.1 summarizes descriptive statistics by class and environment.

	Class 1	Class 2	Class 3	Class 4	Class 5	Total Volume
Urban Environment						
Mean	30,628	3,551	116	2,176	108	36,579
Min	8,291	189	1	30	3	9,163
Max	48,369	8,642	481	4,146	495	58,256
sd	5,250	1,829	83	1,265	70	7,657
Suburban Environment						
Mean	13,232	1,621	38	1,887	41	16,819
Min	1,902	252	0	73	0	2,835
Max	22,580	2,974	90	3,155	231	26,617
sd	2,399	591	21	821	34	3,057

Table 3.1: Descriptive statistics of daily traffic volumes by vehicle class and environment (Jan 2022–May 2025).

3.2.2.2 Gross Domestic Product (GDP)

Quarterly GDP data, covering the period from Q1 2017 to Q1 2025 (33 records), provide a measure of Portugal’s macroeconomic performance. Values range from 45,407 million euros in Q2 2020 (reflecting pandemic-related disruptions) to 73,420 million euros in Q1 2025, with a mean of 57,667 million euros.

These indicators capture the influence of broader economic cycles on traffic demand, particularly relevant for medium- and long-term forecasts. Table 3.2 summarizes the descriptive statistics.

Table 3.2: Descriptive statistics for quarterly GDP values in Portugal (2017Q1–2025Q1), in millions of euros.

Statistic	GDP (millions of euros)
Count	33
Mean	57,667
Std. Dev.	842
Min	45,407
25%	51,359
Median	53,899
75%	65,776
Max	73,420

3.2.2.3 Consumer Price Index (CPI)

Monthly CPI data from INE were collected for the mainland, northern, and central regions of Portugal, covering January 2017 to April 2025 (100 records per region). CPI values are expressed relative to the 2012 base year.

Inflation rates remained modest, with averages close to 0.2% across all regions. The mainland CPI ranged from -1.37% to $+2.52\%$, while the northern and central regions exhibited slightly broader but still moderate variability. Table 3.3 presents the descriptive statistics.

Table 3.3: Descriptive statistics for monthly CPI inflation rates (2017–2025) by region.

	Mainland	North	Center
Count	100	100	100
Mean (%)	0.20	0.22	0.19
Std. Dev. (%)	0.79	0.92	0.72
Min (%)	-1.37	-1.76	-1.15
25%	-0.31	-0.31	-0.26
Median (%)	0.03	0.03	0.04
75%	0.60	0.59	0.46
Max (%)	2.52	2.83	2.67

3.2.2.4 Holiday Data

Holiday data were programmatically retrieved using the `holidays` Python library. The dataset includes national and regional holidays for Portugal between 2022 and 2025, ensuring alignment with the temporal coverage of the traffic and macroeconomic datasets. Movable holidays, such as Carnival, were computed algorithmically to guarantee accuracy.

Together, these datasets provide a comprehensive foundation for forecasting traffic volumes. While traffic data serve as the primary endogenous variable, GDP and CPI account for macroeconomic influences, and holiday calendars incorporate calendar effects. The integration of these heterogeneous sources strengthens the explanatory power of the models, enabling both short- and long-term forecasting.

3.2.3 Exploratory Data Analysis

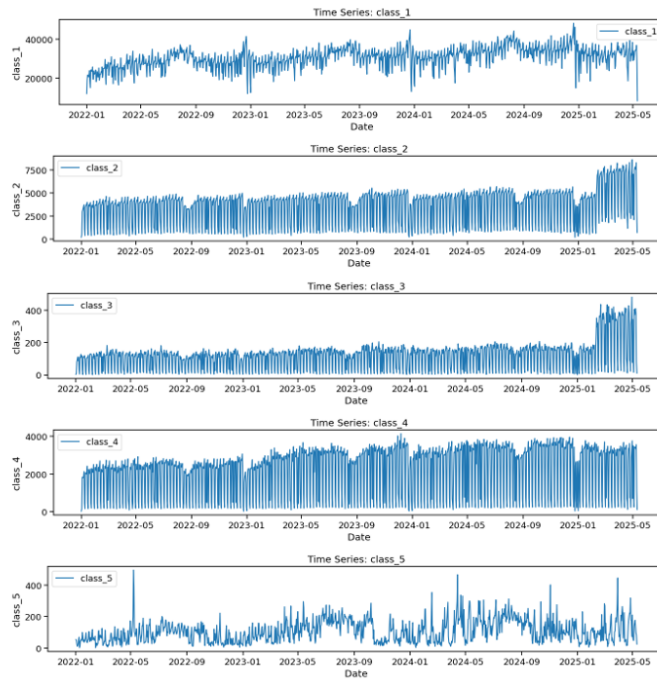
The exploratory analysis was designed to uncover key characteristics and underlying patterns within the traffic dataset, providing critical insights that informed both data preparation and subsequent modeling strategies. A diverse set of exploratory techniques—including time series visualizations, seasonal pattern analysis, statistical tests, and autocorrelation functions—was employed to investigate temporal dynamics, variance structures, and stationarity properties across vehicle classes and contexts.

3.2.3.1 Time Series Visualization

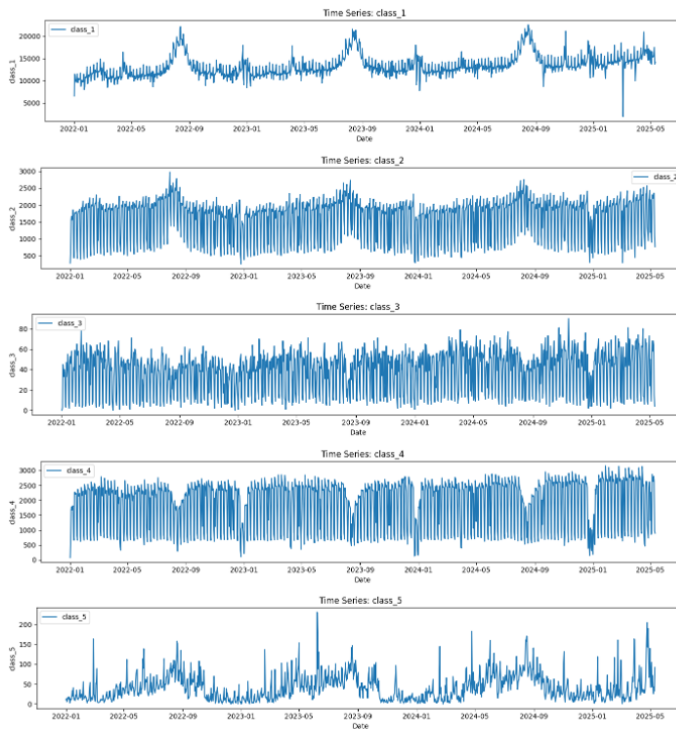
As an initial step, time series plots were generated separately for each vehicle class in both the urban and suburban settings (Figures 3.4a and 3.4b). These visualizations provided a high-level overview of traffic fluctuations over time, serving as an entry point for more detailed temporal analyses. The analysis approach was similar for both environments, allowing for direct comparison between urban and suburban patterns. An initial inspection suggested that Class 1 exhibited unique behavior—appearing more sensitive to vacations periods (August and December)—while Classes 2, 3, and 4 displayed similar trends, with traffic rising and falling, apparently on a weekly basis. Class 5 stood out for its volatility, with pronounced outliers. These initial observations motivated further investigation.

3.2.3.2 Seasonal Pattern Analysis

To further investigate recurring temporal structures within the traffic data, a seasonal decomposition was performed for each time series using the `seasonal_decompose` function from the `statsmodels` library. This method requires selecting a decomposition model—either additive or multiplicative. To guide this choice, the rolling standard deviation was examined across all vehicle classes and environments. Since the magnitude of seasonal fluctuations remained approximately constant regardless of the series level (Hyndman and Athanasopoulos (2021)), the additive model was deemed most appropriate. Consequently, additive decomposition was applied to all series. This approach enabled



(a) Urban environment



(b) Suburban environment

Figure 3.4: Time series plots for each vehicle class in the (A) urban and (B) suburban environments.

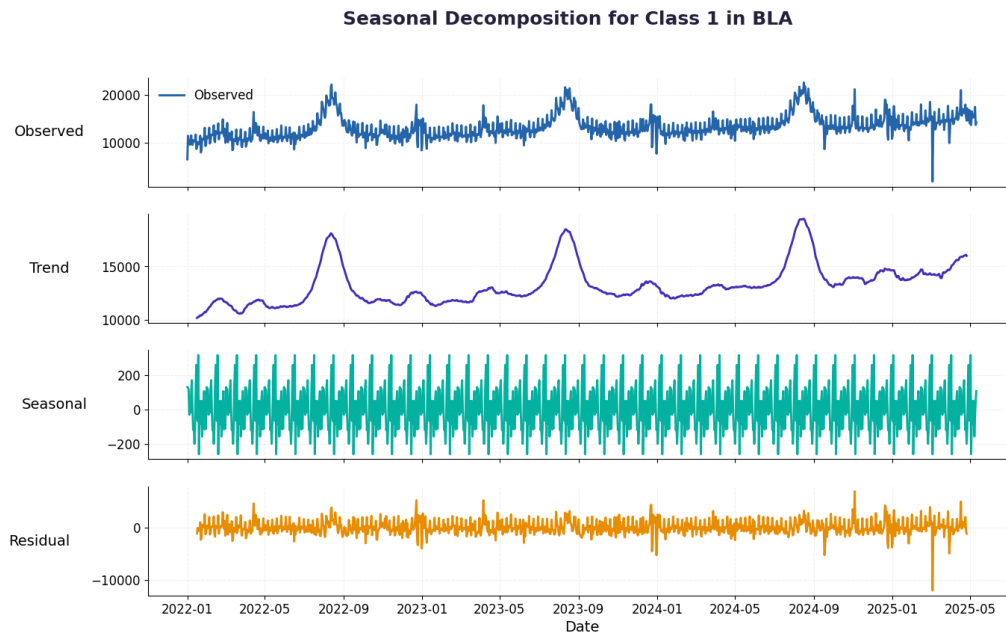


Figure 3.5: Monthly seasonal additive decomposition for small vehicles in BLA

the clear visualization of long-term trends, recurring seasonal patterns, and residual (irregular) components for each vehicle class. The results revealed both pronounced weekly cycles and broader annual patterns across all classes (refer to Figure 3.5 for a sample decomposition).

Upon recognizing seasonal trends within the data, the following three additional analysis was conducted to verify the periodic nature of these patterns. These analyses were executed across each vehicle class and environment to identify key differences and further distinguish the underlying patterns. For individual visual components by vehicle class and context, see Appendix B.

1. **Yearly Seasonality Comparison:** Traffic volumes recorded daily were aggregated by year, and overlaid with line graphs to evaluate the annual patterns over the recorded timeframe.
2. **Weekly Seasonality Examination:** To explore differences within the week, Box-plots for each weekday were created, displaying medians, quartiles, as well as irregularities in traffic patterns within the weekly cycle.
3. **Monthly Average Assessment:** To encapsulate monthly changes, average monthly traffic volumes were computed. Then, bar graphs were created to represent these monthly averages, aiding in recognizing wider seasonal patterns and any potential deviations from anticipated trends.

In terms of findings, the analysis indicates that traffic for class 1 vehicles exhibits similar patterns in both urban and suburban regions, including an annual cycle with slightly lower volumes on weekends and a modest increase on Fridays. Traffic volume typically declines in January but rises during July, August, and December, with the suburban area displaying a particularly notable increase in August, as illustrated in Figure B.1.

For Class 2, traffic patterns have remained consistent over the years, with a marked decline on weekends and a moderate increase on Fridays, especially in suburban areas. July is generally associated with a rise in traffic for both environments, while suburban regions also observe an increase in August. Conversely, there are slight decreases in January and December. Notably, a distinctive pattern emerged in urban areas beginning in February 2025, see Figure B.2.

Classes 3 and 4 display very stable traffic patterns from year to year, characterized by clear annual seasonality. Both classes experience significant drops on weekends and modest decreases in August, December, and January, with similar trends observed across urban and suburban settings, (Figures B.3 and B.4).

Finally, traffic in Class 5 remains highly consistent throughout the year, showing slight increases during weekends and stable levels on weekdays, alongside a pronounced rise during the warmer months. These patterns are comparable in both urban and suburban environments, as illustrated in Figure B.5.

Collectively, these analyses reveal both regular seasonal effects and subtle distinctions across classes and contexts, supporting a comprehensive understanding of temporal dynamics in traffic flow.

3.2.3.3 Autocorrelation Analysis

To better understand the influence of previous days on traffic, an autocorrelation and partial autocorrelation analyses were performed using the `plot_acf` and `plot_pacf` functions from the `statsmodels` library. The resulting ACF and PACF plots were generated for each time series to identify the presence and extent of autocorrelation at different lags, a crucial step for selecting and specifying appropriate time series forecasting models.

For Class 1, both the urban and suburban environments exhibit an exponential decay in the ACF, as shown in Figures 3.6a and 3.6b. The decay rates are only slightly different between environments. The PACF plots (Figures 3.7b and 3.7a) for Class 1 are also very similar, with a strong initial correlation that drops off rapidly.

Classes 2, 3, and 4 display nearly identical periodic or cyclical autocorrelation patterns in both urban and suburban settings, again visible in Figures 3.6a and 3.6b. The oscillation periods, amplitudes, and overall structure are highly consistent across environments. The

PACF for these classes (Figures 3.7b and 3.7a) also shows remarkable agreement between urban and suburban data.

For Class 5, both environments exhibit a similar pattern of rapid decay in the ACF, with less structured correlation compared to other classes, as shown in Figures 3.6a, 3.6b, 3.7b, and 3.7a.

Overall, the fundamental temporal correlation structures are highly similar between the urban and suburban settings. The differences observed are minor variations rather than fundamentally different behaviors. This similarity suggests that the underlying processes driving these traffic patterns are consistent across both environments and that the temporal dynamics are not strongly influenced by whether the setting is urban or suburban. In summary, the autocorrelation analysis confirms that the traffic behaviors across both settings are indeed very similar.

3.2.3.4 Stationarity Testing

To assess the suitability of various time series models, we performed the statistical ADF test on each individual series. This ADF test serves to identify the existence of a unit root within the series, thereby assisting in determining whether the series is stationary. The hypotheses for this test are stated as follows:

- **Null Hypothesis (H_0):** The series has a unit root (i.e., it is non-stationary).
- **Alternative Hypothesis (H_1):** The series does not have a unit root (i.e., it is stationary).

If the null hypothesis cannot be rejected, this provides evidence that the series is non-stationary. In practical terms, these tests helped determine whether the series exhibited a constant mean and variance over time or required transformation (e.g., differencing) prior to modeling.

Table 3.4 presents a summary of the results. Generally, traffic series on highways are stationary in both environments. However, an anomaly is noticed in Classes 2 and 3 within the urban environment, causing non-stationary behavior. This anomaly aligns with prior observations of irregular patterns in the urban context. Aside from this, all classes in both environments display stationary behavior, supporting the notion of uniform underlying temporal patterns.

3.2.4 Data Quality Verification

Before implementation, a thorough evaluation of the dataset's quality and comprehensiveness was conducted. Throughout the analysis period from January 2022 to May 2025, the

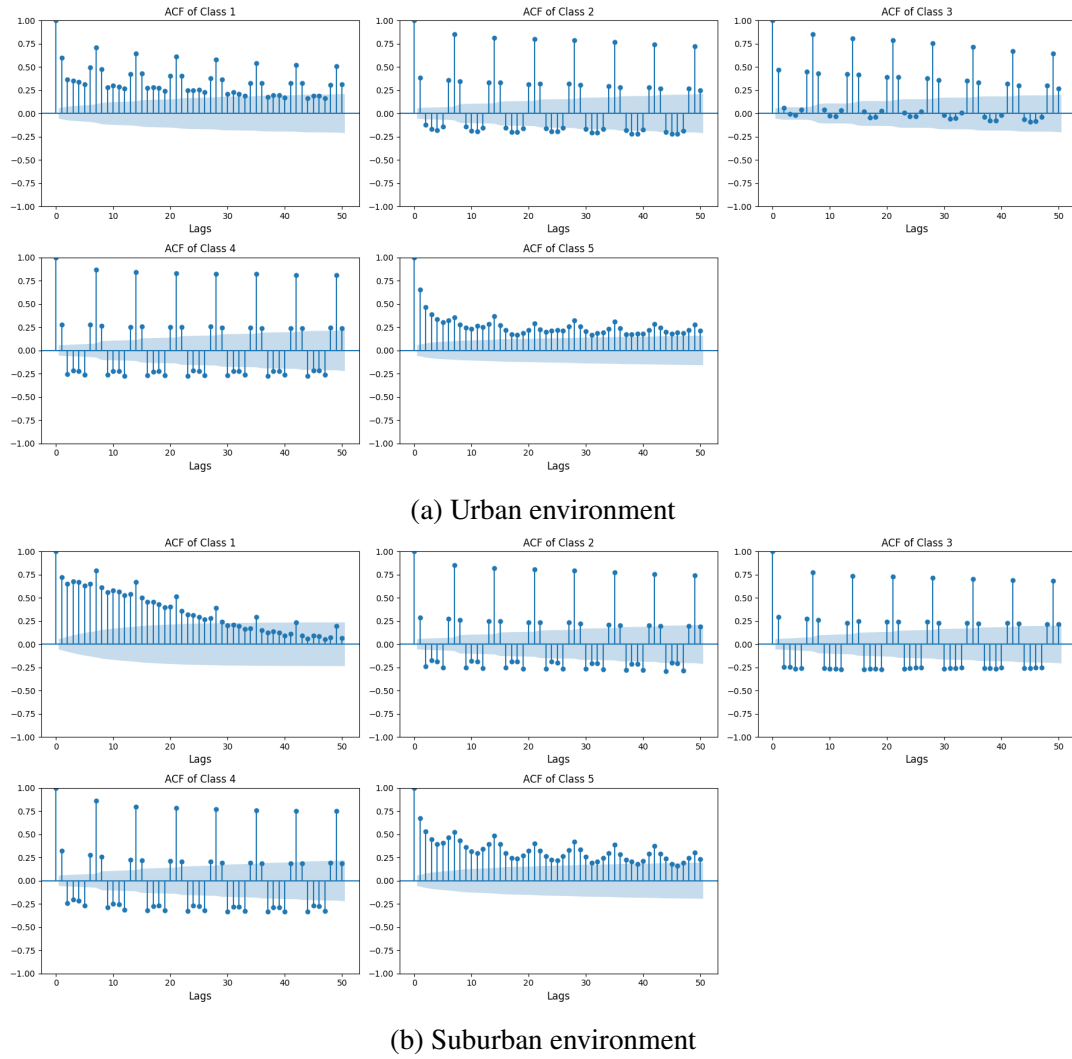
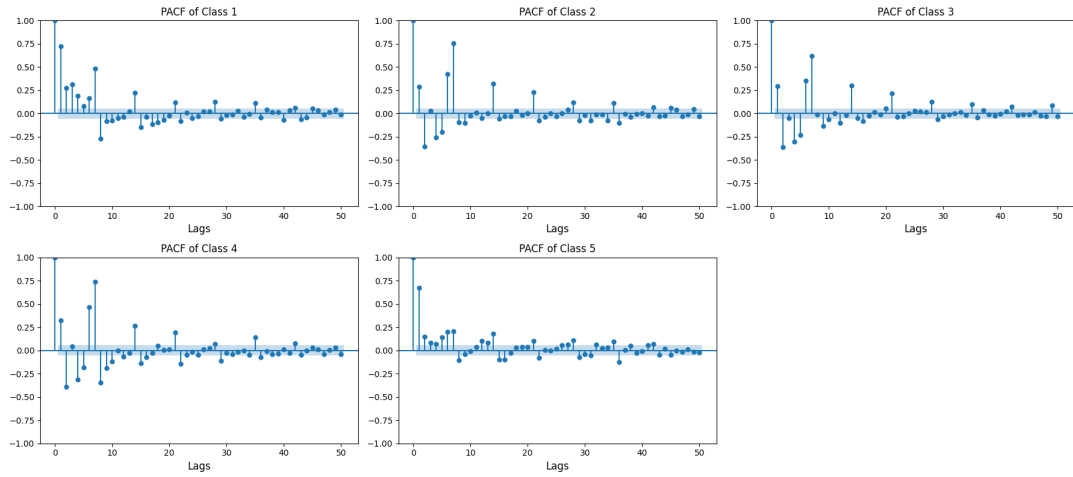
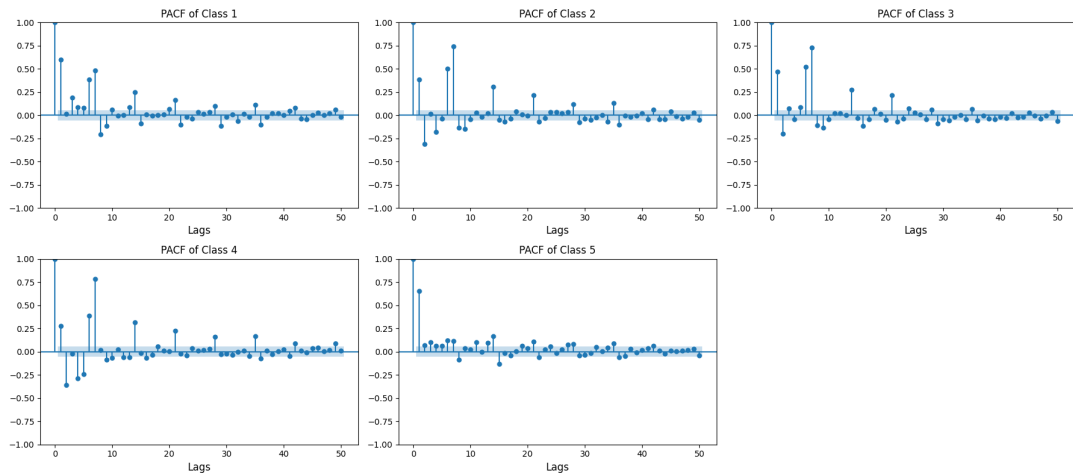


Figure 3.6: ACF for each vehicle class in the (A) urban and (B) suburban environments.



(a) Suburban environment



(b) Urban environment

Figure 3.7: PACF for each vehicle class in the (A) suburban and (B) urban environments.

Environment	Class	ADF Statistic	p-value	Critical Value (5%)	Stationarity
Urban	Class 1	-3.198	0.020	-2.864	Stationary*
	Class 2	-1.851	0.355	-2.864	Non-stationary
	Class 3	-0.084	0.951	-2.864	Non-stationary
	Class 4	-3.193	0.020	-2.864	Stationary*
	Class 5	-3.800	0.003	-2.864	Stationary**
Suburban	Class 1	-4.683	<0.001	-2.864	Stationary***
	Class 2	-3.484	0.008	-2.864	Stationary**
	Class 3	-4.331	<0.001	-2.864	Stationary***
	Class 4	-6.158	<0.001	-2.864	Stationary***
	Class 5	-3.515	0.008	-2.864	Stationary**

Table 3.4: ADF Results for Stationarity by Environment and Class

Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Statistical significance levels indicate rejection of the null hypothesis of non-stationarity.

traffic data are devoid of missing values, ensuring complete representation for each day. Zero values within this dataset signify that on those particular days, vehicles of a certain class were absent, rather than indicating any missing or incorrect data.

During previous analysis, deviations and atypical values were observed within the data. These anomalies were reviewed in consultation with traffic specialists, who confirmed that such deviations are not attributable to measurement error or random noise. Instead, they may reflect noteworthy phenomena in traffic patterns. As a result, these cases have been preserved in the dataset and will be subject to further, targeted investigation in subsequent analyses.

3.2.4.1 Holiday Impact on Traffic Patterns

Initial exploration of the data indicated possible links between national holidays and shifts in traffic patterns among different vehicle classes. Figure 3.8 illustrates this using class 1 vehicles, specifically small vehicles, in a suburban setting. To further explore these relationships, the traffic volumes for each class on holidays were analyzed against the average volumes for non-holidays, categorized by the same weekday. This method accounts for regular weekly patterns, enabling a more detailed evaluation of how holidays might distinctly influence traffic compared to regular days.

Figures 3.9a and 3.9b summarize the results for urban and suburban environments, respectively, showing the average percentage change from the non-holiday baseline for each vehicle class and major holiday.

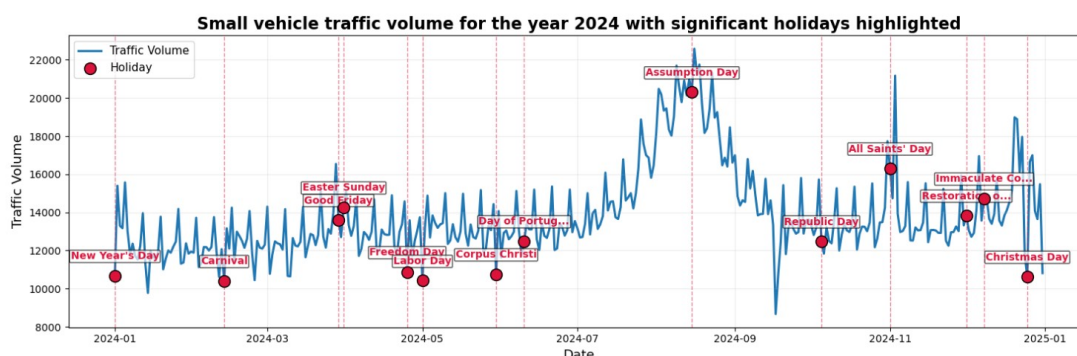


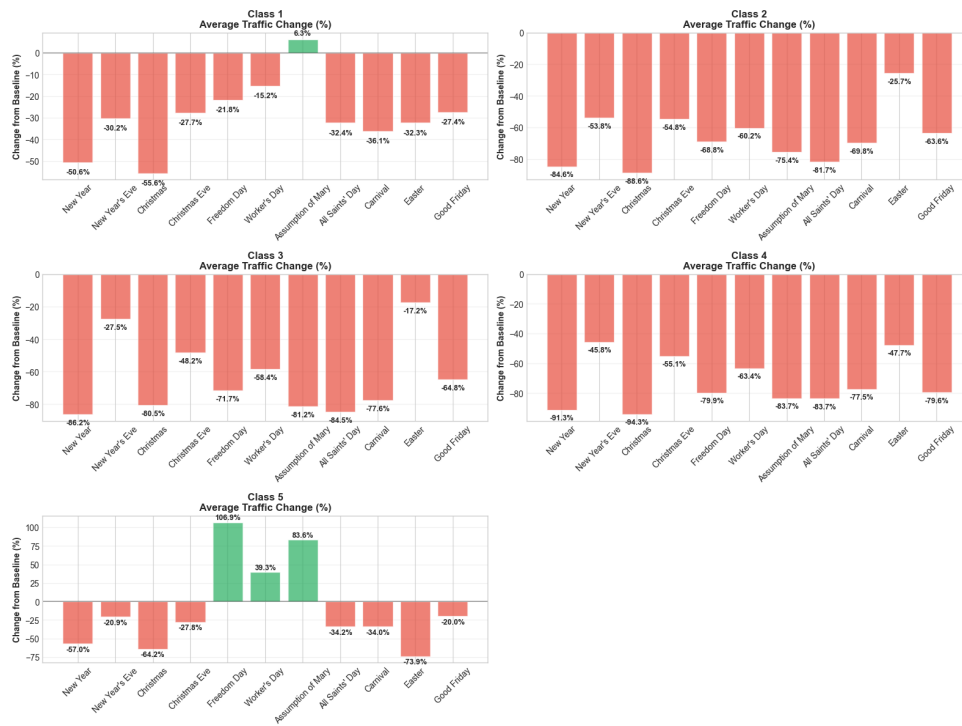
Figure 3.8: Daily traffic volume for Class 1 vehicles in the suburban environment during 2024, with significant holidays highlighted. Each red marker indicates a national or major holiday, annotated with the respective holiday name.

This analysis revealed that Classes 2, 3, and 4, which mainly comprise commercial vehicles, consistently show significant traffic volume reductions during most holidays. Often, these reductions exceed 50% and can even reach or surpass 80% for holidays like New Year and Christmas. This consistent pattern across different holidays and situations indicates that the operation of these vehicle types is especially affected by holiday periods.

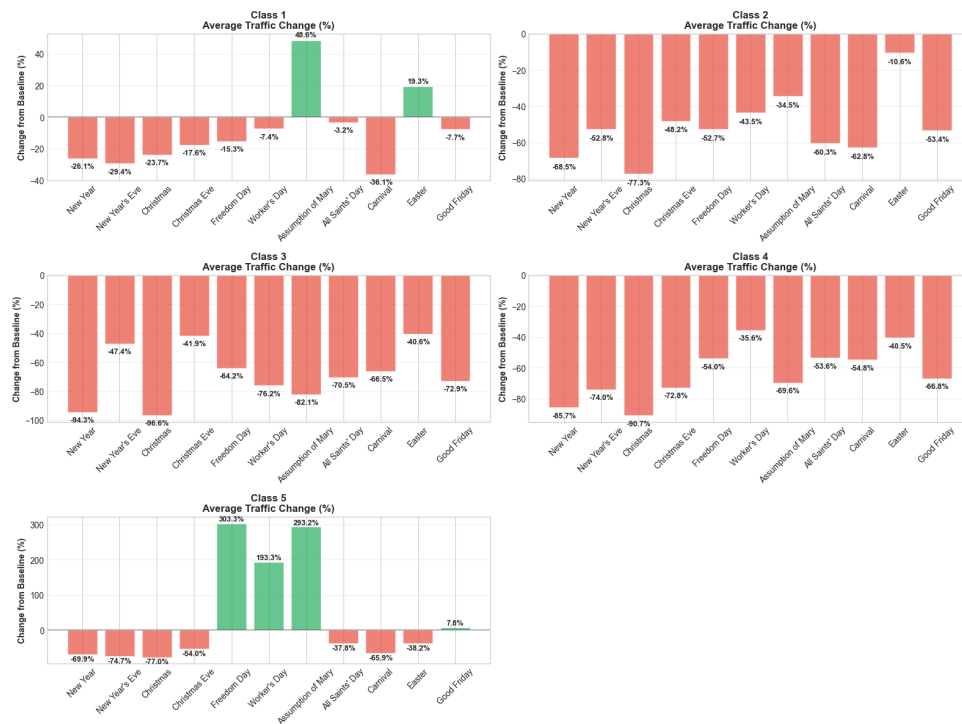
For Class 1 vehicles (private/personal use), the changes are less uniform. Reductions are apparent for several holidays—especially at the end of the year and during civic holidays—yet certain holidays, such as the Assumption of Mary, show positive changes in traffic, with increases of up to 46% in urban areas and 6% in suburban zones. The more accentuated behavior observed in the suburban zone may be influenced by seasons, as traffic volumes typically increase in the summer months. Another notable difference between environments is found during Easter: while urban areas experience a reduction of around 30%, suburban areas tend to show an increase of approximately 19%. These observations indicate that, for Class 1, holiday effects are more varied and may depend on both the season and the specific nature of the holiday.

Class 5 (motorcycles) stands out for its variability and the presence of significant traffic increases during some holidays. Notably, for Freedom Day, Worker's Day, and the Assumption of Mary, traffic volumes for Class 5 rise dramatically, particularly in the urban environment, with increases exceeding 190% and even reaching 300% above baseline. In suburban areas, similar increases are observed, though of lower magnitude. For other holidays, changes in Class 5 traffic are generally negative, with a seasonal influence: increases are more pronounced during warmer months.

Overall, these findings emphasize that holidays can significantly alter traffic patterns. Although, traffic volumes vary among different vehicle categories and settings. Commercial vehicles exhibit the most substantial and consistent decreases, whereas private and



(a) Urban environment



(b) Suburban environment

Figure 3.9: Average percentage change in traffic volume by holiday and vehicle class in the (A) urban and (B) suburban environments.

leisure vehicles (Classes 1 and 5) demonstrate a wider range of reactions, including both decreases and notable increases, contingent on the specific holiday and season.

3.2.4.2 Holiday Impact on Adjacent Weekends

To explore whether holidays influence weekend traffic volumes, we investigated the average percentage change in vehicle volumes across various categories, contrasting weekends next to holidays with typical weekends. Figures 3.10a (urban) and 3.10b (suburban) illustrate the findings, showcasing distinct patterns linked to vehicle type and regional setting.

Findings indicate substantial traffic decline around holidays, particularly in suburban regions, where reductions often exceed 60% for Class 4 vehicles. Classes 3 and 2 also see notable decreases, frequently over 40%. In urban areas, reductions are significant but less pronounced, generally between 20-50%. Some holidays, like Freedom Day, have minimal impact or slight increases in urban traffic.

Class 1 vehicles show varied patterns, with suburban areas experiencing 20-28% reductions during end-of-year holidays but notable increases for some civic holidays (e.g., a 59.7% rise for Worker's Day). In urban areas, reductions are moderate (10-20%) during these holidays, with smaller rises on civic holidays. Class 5 vehicles (motorcycles) are marked by volatility, showing sharp increases on civic holidays in suburban areas (e.g., 106.8% on Worker's Day) but substantial declines during winter holidays. Urban Class 5 vehicles mirror this volatility but with less severe changes.

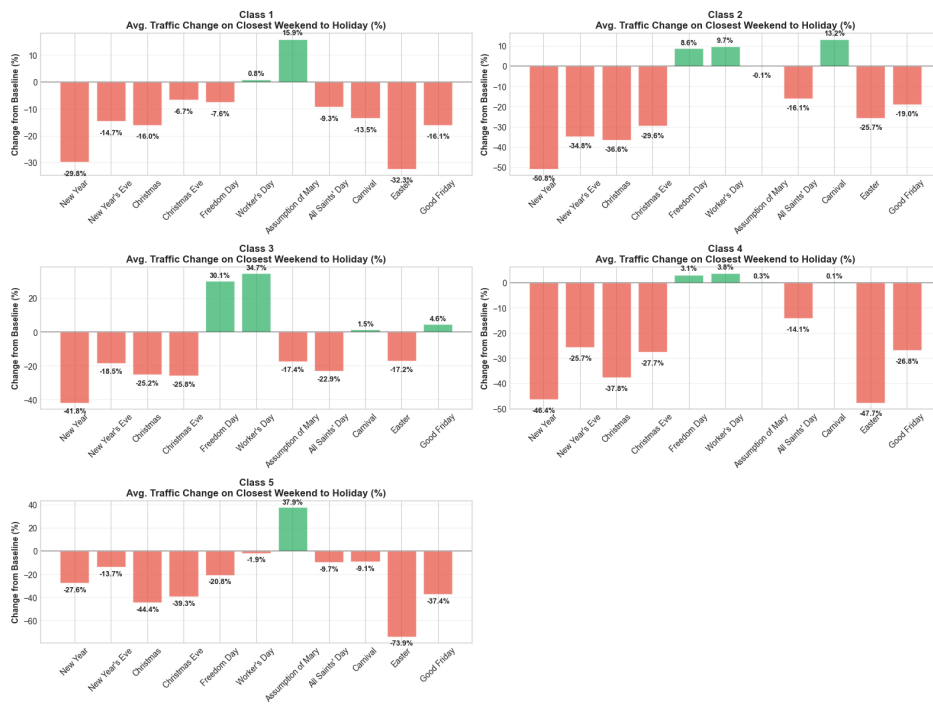
In summary, holiday effects on weekend traffic vary by vehicle type and region, with commercial vehicles most affected in suburbs and passenger vehicles and motorcycles showing diverse responses.

3.2.4.3 Structural Change in Urban Traffic (Feb 2025)

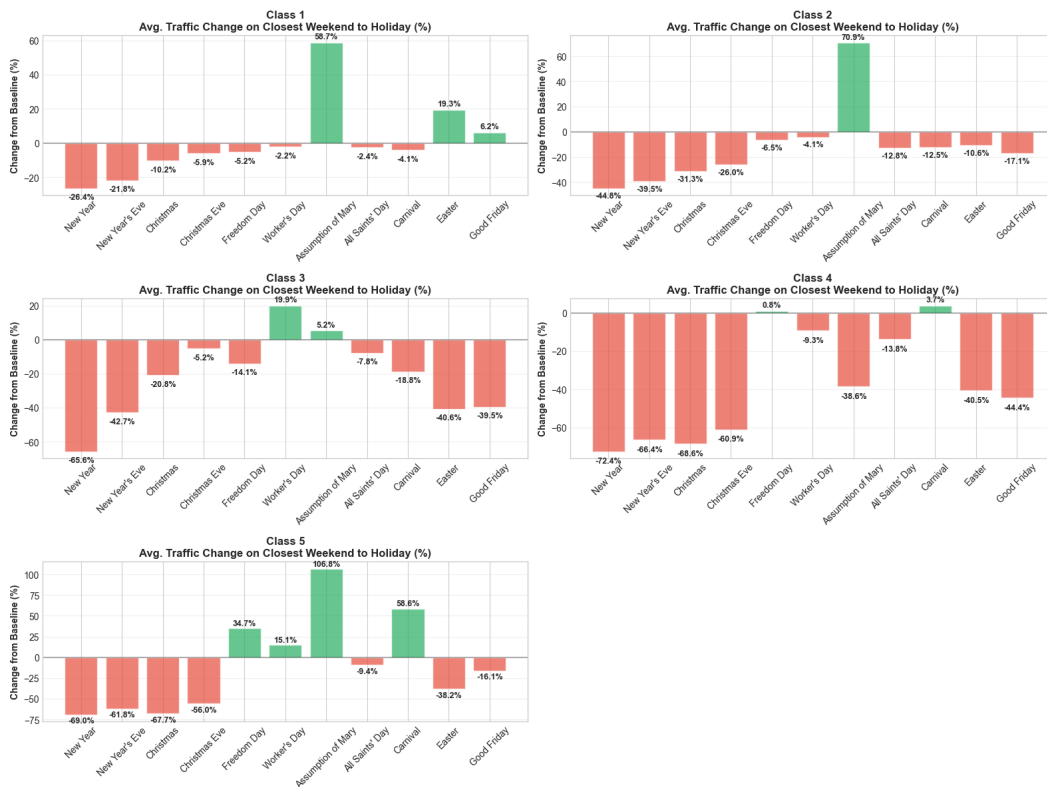
Beyond holiday effects, a notable and prolonged rise in urban traffic volumes for Classes 2 and 3 was observed beginning in mid-February 2025. This trend was validated by traffic experts, who confirmed its authenticity and attributed it to an upcoming policy change: heavy vehicle tolls in the CREP corridor are scheduled to be eliminated starting in 2026. Despite the official implementation date being set for the following year, early anticipation of this measure appears to have already influenced traffic patterns in the region.

To statistically assess the impact of this intervention, a paired t-test was performed to compare traffic volumes before and after February 10, 2025. The following hypotheses were formulated:

- **Null hypothesis (H_0):** There is no significant difference in the mean daily traffic volumes for Classes 2 and 3 before and after February 10, 2025.



(a) Urban environment



(b) Suburban environment

Figure 3.10: Average percentage change in weekend traffic volumes adjacent to holidays for each vehicle class in the (A) urban and (B) suburban environments.

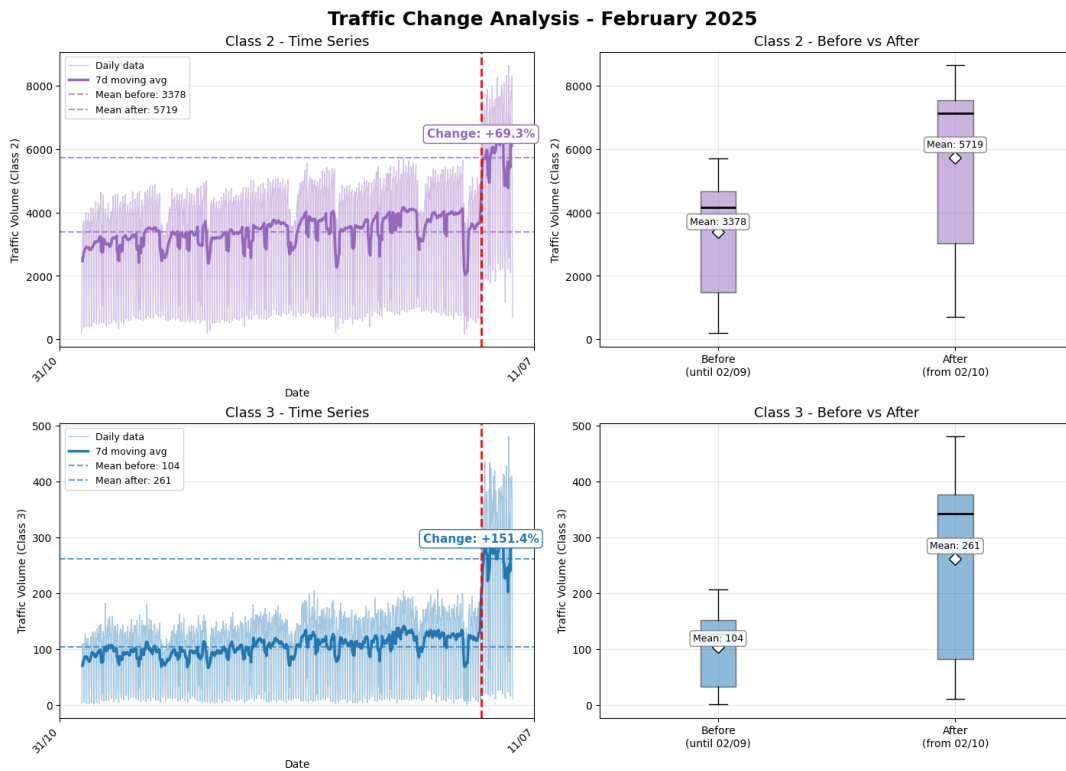


Figure 3.11: Traffic change analysis for Classes 2 and 3 in the urban environment before and after February 10, 2025. Left: Daily time series with seven-day moving average and mean lines for the periods before and after the intervention. Right: Boxplots showing the distribution and means for each period.

- **Alternative hypothesis (H_1):** There is a significant difference in the mean daily traffic volumes for Classes 2 and 3 before and after February 10, 2025.

The test results (see Table 3.5) indicate significant increases in both classes (Class 2: +69.3%; Class 3: +151.4%; both $p < 0.0001$), supporting the rejection of the null hypothesis and confirming the existence of a structural change in traffic patterns. Figure 3.11 visually summarizes the traffic volume change for Classes 2 and 3 before and after the intervention date.

3.2.4.4 Event Impact: Expomoto and Motorcycle Traffic

In the urban environment, a biennial pattern was observed in Class 5 (motorcycle) traffic, coinciding with major motorcycle events such as Expomoto. To investigate this relationship, daily motorcycle traffic on event days was compared to control days (same months, no events). The following hypotheses were tested:

- **Null hypothesis (H_0):** No difference in mean motorcycle traffic between event days and control days.

Table 3.5: Statistical summary of traffic change before and after February 10, 2025 (urban environment).

Class	Period	N	Mean	Median	SD	Change (%)	Significance
Class 2	Before	1,136	3,377.6	4,165.5	1,657.9	+69.3%	Yes***
	After	91	5,719.4	7,144.0	2,398.6		
Class 3	Before	1,136	103.9	128.0	60.6	+151.4%	Yes***
	After	91	261.3	343.0	152.7		

Note: Statistical significance assessed using paired samples t-test. *** indicates $p < 0.001$. The reference date of February 10, 2025, marks the implementation of the new toll system in the urban road segment.

- **Alternative hypothesis (H_1):** There is a significant difference in mean motorcycle traffic between event days and control days.

Figure 3.12 shows the results: the average daily traffic for Class 5 during Expomoto was 248, compared to 126 on control days, an increase of 97.1%. The statistical test confirmed the significance ($p < 0.05$), providing robust evidence that these large-scale events are associated with the mean traffic volume.

3.2.4.5 Additional Anomalies in the Suburban Environment

During the analysis of the suburban environment, three further anomalous values were identified to all vehicle classes. Two of these were linked to fire incidents: on **March 30, 2025**, a fire involving a light vehicle on the A25 highway likely explains the abrupt drop in traffic; on **September 17, 2024**, widespread fires in Aveiro caused multiple closures, severely impacting circulation. The third anomaly, on **March 2, 2025**, showed no clear external cause and is presumed to be a data error.

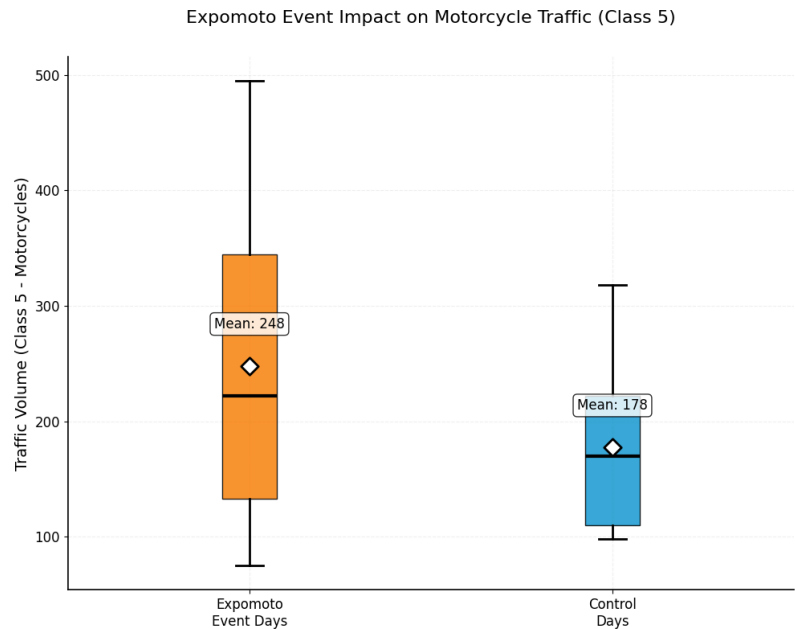


Figure 3.12: Comparison of Class 5 (motorcycle) traffic on Expomoto event days and control days in the suburban environment.

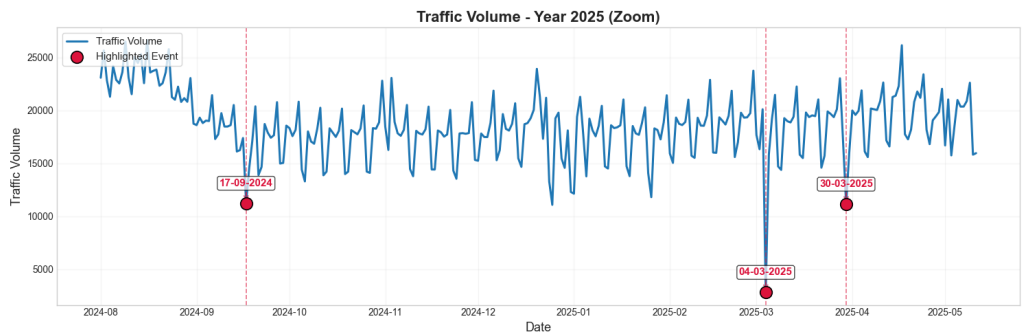


Figure 3.13: Abrupt changes in traffic volume on specific dates in the suburban environment, with annotation of fire-related incidents.

Chapter 4

Implementation

This chapter presents the subsequent stages of the CRISP-DM methodology, namely *Data Preparation* (Section 4.1) and *Modeling* (Section 4.2), which are primarily related to the implementation of the forecasting models. In the DP section, each step of the data engineering process is discussed in detail, from initial data selection to the transformation and formatting required to make the data suitable for modeling. Then, the Modeling section describes the modeling process itself, including model selection, test design, model construction, and model assessment, concluding with the evaluation metrics used.

Overall, this chapter provides a solid foundation for the subsequent phases of the CRISP-DM process by systematically addressing the methodological and practical aspects required for effective model implementation.

4.1 Data Preparation

The Data Preparation process encompasses several tasks, as illustrated in Figure 4.1, including data selection, cleaning, construction, integration, and formatting. Each of these activities addresses specific challenges that may arise from real-world data, such as missing values, duplicates, outliers, and inconsistencies. The process is iterative, requiring multiple passes to refine the dataset and resolve newly identified issues.

This section details the main steps, decisions, and techniques involved in preparing the traffic, economic, and holiday datasets used in this research, and emphasizes their potential influence on further modeling and analysis.

4.1.1 Select Data

Based on the previous EDA (Subsection 3.2.3), features such as holidays, day of the week, weekends preceding holidays, and lagged values were identified as potentially

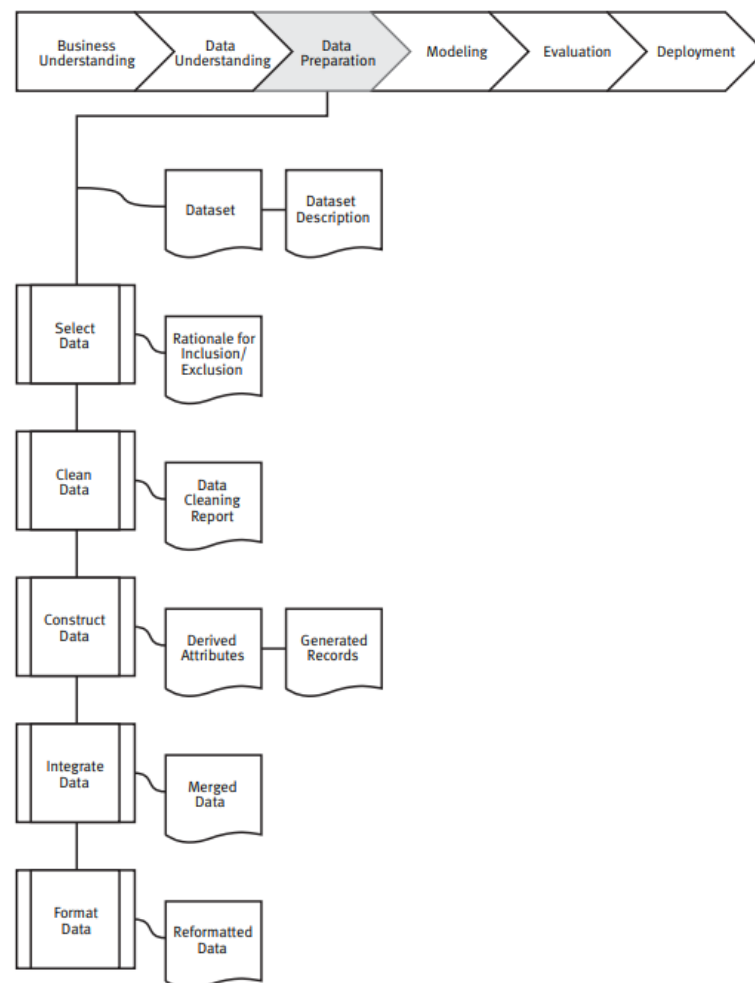


Figure 4.1: Detailed Data Preparation, Chapman (2000)

important predictors. To further investigate the relationships between traffic volumes, vehicle classes, and external factors, and to support the feature selection process, Pearson correlation coefficients were computed using Python's `corr()` method. All datasets were normalized to ensure comparability across variables. The results for both urban and suburban contexts are presented as correlation heatmaps in Figures 4.2a and 4.2b.

4.1.1.1 Correlations Among Vehicle Classes

The correlation matrices reveal distinct groupings among vehicle classes. Classes 2, 3, and 4 (mainly commercial vehicles) exhibit strong mutual correlations in both environments (0.82–0.94), suggesting synchronized traffic dynamics. By contrast, Class 1 (private vehicles) shows weaker correlations with these classes (0.17–0.65), reflecting different underlying travel patterns. Class 5 (motorcycles) presents weak or negative correlations with commercial vehicles, but moderate positive correlations with Class 1, indicating recreational or non-work-related dynamics.

4.1.1.2 Macroeconomic Indicators

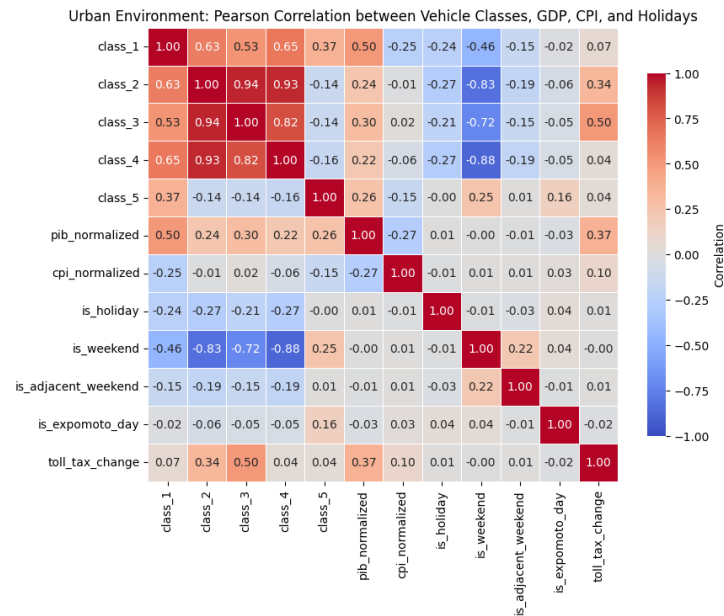
Correlations with GDP are generally weak, except for Class 1, which shows moderate positive correlations (0.41–0.50), suggesting that private vehicle usage is more sensitive to economic growth. In contrast, correlations with the CPI are negative for Class 1 (–0.24 to –0.25) and close to zero for other classes, indicating limited responsiveness to inflation.

4.1.1.3 Calendar Effects

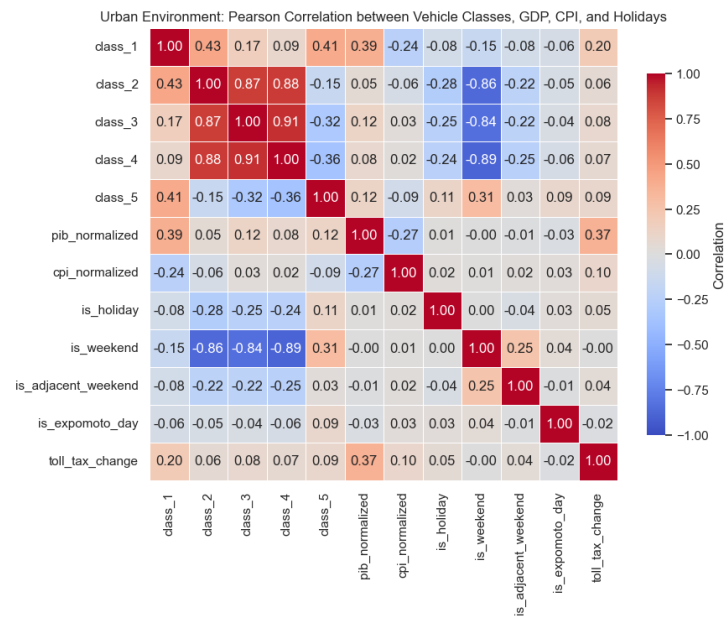
Holidays: Commercial classes (2, 3, and 4) consistently exhibit reduced traffic during holidays (–21% to –28%). Class 1 decreases in urban areas (–24%), while Class 5 shows negligible or slightly positive correlations.

Weekends: Weekends are strongly negatively correlated with traffic in commercial classes (below –0.72 in urban areas and –0.84 in suburban areas). Class 1 also decreases, especially in urban areas (–0.46), while Class 5 increases moderately (0.25–0.31), consistent with recreational use.

Adjacent Holiday Weekends: Traffic decreases moderately for commercial classes (–15% to –25%). For Class 1, the effect is negligible in suburban areas but slightly negative in urban areas (–15%).



(a) Urban environment



(b) Suburban environment

Figure 4.2: Pearson correlation coefficients between vehicle classes, calendar and macroeconomic indicators, and flagged external events in (A) urban and (B) suburban environments.

4.1.1.4 External Events

Expomoto Days: Correlations are negligible for most classes, but slightly positive for motorcycles (0.09–0.16). The limited statistical effect reflects the event’s infrequent occurrence, although earlier analyses (Subsection 3.2.4.4) confirmed a practical impact on motorcycle traffic.

Toll Policy Chang: Classes 2 and 3 in the urban environment were most affected, with positive correlations (0.34 and 0.50, respectively), consistent with earlier findings (Subsection 3.2.4.3). Other classes show minimal sensitivity.

4.1.1.5 Feature Selection Rationale

Drawing on the findings from the correlation analysis and the EDA, the following features were selected for inclusion in the forecasting models:

- **Lagged traffic values:** For each time series, key lagged features (1, 7, 14, 21, and 28) were selected based on the autocorrelation analysis (Subsection 3.2.3.3), in order to capture temporal dependencies.
- **Macroeconomic indicators:** GDP and CPI were included, as they are standard in forecasting practice. Their relevance, however, was found to be significant mainly for Class 1.
- **Calendar effects:** Binary indicators for holidays, weekends, and weekends adjacent to holidays were added as external regressors, given their consistent influence on traffic across several classes.
- **Policy and event indicators:** For Classes 2 and 3 in the urban environment, a binary feature was included to represent the post-toll policy change period. For Class 5 (motorcycles), an indicator for Expomoto event days was added to capture the influence of this large-scale event in both urban and suburban contexts.

For macroeconomic indicators, the data were restricted to January 1, 2022, onward, aligning with the period for which traffic data are available. During preprocessing, no additional external variables were identified that could reasonably be expected to enhance model performance.

4.1.2 Clean Data

To ensure data quality at the level required for reliable forecasting, a systematic data cleaning process was carried out to address the irregularities and anomalous behaviors identified in the *Verify Data Quality Subsection 3.2.4*. The main strategies applied are described below, grouped according to the type of anomaly. For anomalies associated with external factors, the rationale for retaining or excluding data points is also discussed.

4.1.2.1 Outliers Related to Holidays and Weekends

Several outliers detected during seasonal decomposition coincided with holidays or special weekends, as shown in Sub-subsections 3.2.4.1 and 3.2.4.2. Since the forecasting models explicitly incorporate variables for holidays, weekends, and adjacent weekends, these values were retained in the dataset, requiring no further adjustments.

4.1.2.2 Event-Related Variations

Certain extreme values corresponded to known events. For example, spikes in motorcycle traffic were aligned with the Expomoto event (see Sub-subsection 3.2.4.4). Rather than discarding or modifying these records, a binary indicator for Expomoto days (restricted to Class 5) was introduced. A similar approach was applied to account for changes in toll policy, specifically affecting Classes 2 and 3 in the urban environment. In both cases, the original data were preserved to enable the models to learn from these real-world variations.

4.1.2.3 External Incidents and Measurement Errors

As observed in sub-subsection *Additional Anomalies in the Suburban Environment 3.2.4.5*, a limited number of anomalies were attributed to external incidents or confirmed data acquisition errors. For these cases, traffic volumes were imputed using the average value for the corresponding day of the month from previous years. This approach mitigates the impact of anomalies while preserving seasonal and weekly structures.

Overall, the cleaning process prioritized the preservation of genuine traffic variations (e.g., holidays, weekends, and major events), while addressing rare anomalies and technical errors through targeted imputation. Remaining irregularities are considered minor and unlikely to materially affect the modeling results.

4.1.3 Construct Data

Following the data cleaning stage, a set of new features was engineered to enhance the original traffic dataset. This process aimed to incorporate temporal, calendar, and event-related information to improve both predictive performance and interpretability. All procedures were applied consistently to both urban and suburban datasets, ensuring methodological comparability across contexts.

4.1.3.1 Temporal Feature Extraction and Transformation

To capture cyclical and seasonal dynamics, several time-related features were derived from the variable `calday` using `pandas` `datetime` utilities:

- **day_of_year**: Day of the year (1–365).
- **week_number**: Weekday (1=Monday, 7=Sunday).
- **month_number**: Calendar month (1–12).
- **quarter_number**: Quarter of the year (1–4).
- **three_month_number**: Aggregates months into three periods (Jan–Apr, May–Aug, Sep–Dec).
- **semester_number**: Aggregates months into two periods (Jan–Jun, Jul–Dec).

Given the cyclical nature of these features, each was transformed into corresponding sine and cosine components using `numpy`. This method preserves cyclic continuity (e.g., December and January are adjacent in time) and allows models to capture periodic patterns more effectively Hyndman and Athanasopoulos (2021).

4.1.3.2 Weekend and Holiday Indicators

To reflect calendar-driven variations in traffic patterns, three binary indicators were created:

- **is_weekend**: Flags Saturdays and Sundays, computed with `pandas` `datetime` utilities.
- **is_holiday**: Flags national holidays in Portugal (2022–2025), using the `holidays` package. Movable holidays such as Carnival (47 days before Easter) were computed with `dateutil.easter`.
- **is_holiday_weekend**: Flags cases where a holiday coincides with a weekend (logical AND of the two previous indicators).

4.1.3.3 Proximity to Holidays

To capture potential effects of extended weekends, the feature `is_adjacent_weekend` was created. For each weekend day, the minimum distance to the nearest holiday was calculated; weekends adjacent to holidays (distance = 1 day) were flagged. This was efficiently computed using Python `set` operations combined with row-wise applications in `pandas`.

4.1.3.4 Special Event and Policy Features

To incorporate external shocks and policy interventions:

- **is_expomoto**: Flags Expomoto event days, based on the official event calendar.
- **toll_tax_change**: Flags the period after February 10, 2025, when the new toll policy came into effect (binary indicator equal to 1 from this date onwards).

Through these feature engineering steps, the dataset was enriched with attributes capturing temporal cycles, seasonality, holiday effects, and external shocks. This enhanced representation provides a stronger basis for forecasting models, supporting both predictive accuracy and interpretability.

4.1.4 Integrate Data

The integration phase consisted in consolidating and enriching the primary dataset by combining information from multiple sources. Most variables related to traffic, as well as some features representing external events (such as calendar information, holidays, Expomoto, and toll policy changes), were directly incorporated into the main dataset during previous processing steps.

However, it was necessary to integrate external macroeconomic indicators, specifically CPI and GDP, since these originated from separate sources and operated at different temporal resolutions (monthly and quarterly, respectively). The integration process, follow described, was carefully designed to align these variables with the daily frequency of the main dataset.

First, the CPI values were merged into the main dataset by matching the month of each record's date (`calday`) to the corresponding monthly value in the CPI dataset. This operation was performed using the `pandas merge` function, aligning the date to the first day of the month to ensure correct mapping.

Next, GDP values—available on a quarterly basis—were merged in a similar way, with each daily record assigned the appropriate GDP value based on the quarter in which it occurred. This required extracting the start date of each quarter from both datasets and performing a join on this key.

Finally, all integrated columns were renamed for clarity and consistency with the feature naming conventions adopted in the rest of the analysis. This careful alignment and integration of data sources ensured that all relevant information—spanning from granular traffic records to broader economic trends—could be utilized in a unified analytical workflow.

4.1.5 Format Data

The formatting stage focused on syntactic transformations required to ensure compatibility between the traffic dataset and the modeling framework. These adjustments did not alter the underlying meaning of the data but were necessary to meet the input specifications of the *Nixtla* libraries, which provide advanced functionality for time series forecasting.

4.1.5.1 Structural Adjustments

In compliance with the library's requirements, the following modifications were implemented:

- **Unique identifier:** A `unique_id` column was introduced as the first column of the dataset. Each environment was assigned a constant identifier (e.g., "GP" for the urban context and "BLA" for the suburban context), ensuring consistency across records.
- **Date formatting:** The original `calday` variable was renamed to `ds`, in accordance with the expected input format.
- **Target variable:** The dependent variable was renamed to `y`, following the conventions of the modeling framework.

4.1.5.2 Removal of Irrelevant Attributes

To streamline the dataset and reduce potential noise, columns unrelated to the modeling task—such as `volume`, `xbicxzchcon_id`, `xbicxzchroad`, and `Data`—were excluded. This ensured that only relevant features were retained for the forecasting process.

Through these formatting transformations, the dataset was standardized to align with the *Nixtla* libraries' requirements. This step ensured a consistent structure across environments and facilitated seamless integration into the subsequent modeling stage.

4.1.6 Final Dataset

After all the steps of exploratory data analysis, data selection, cleaning, integration, construction, and formatting, the result is a consolidated and fully prepared dataset, ready for modeling. This final dataframe brings together all relevant features, both original and engineered, as well as the integrated macroeconomic indicators and event-based variables.

The table below illustrates a sample of the final dataset, showing the structure and the diversity of information available for subsequent modeling and analysis stages:

unique_id	ds	class_1	class_2	class_3	class_4	class_5	...
GP	2022-01-01	0.0000	0.0001	0.0042	0.0000	0.1037	...
GP	2022-01-02	0.1896	0.0312	0.0146	0.0299	0.0976	...
GP	2022-01-03	0.2650	0.3279	0.1688	0.4361	0.0508	...
GP	2022-01-04	0.2555	0.3477	0.1938	0.4312	0.0183	...
GP	2022-01-05	0.2386	0.3708	0.2271	0.4325	0.0224	...
GP	2022-01-06	0.2604	0.3919	0.1792	0.4473	0.0508	...
GP	2022-01-07	0.3262	0.4188	0.2271	0.5100	0.1057	...
GP	2022-01-08	0.1742	0.1054	0.0625	0.1028	0.1362	...
GP	2022-01-09	0.0859	0.0225	0.0083	0.0328	0.0000	...
GP	2022-01-10	0.3324	0.3898	0.2250	0.4944	0.0203	...

Table 4.1: Sample of the final, formatted dataset ready for modeling. The full dataset contains 28 columns, including engineered cyclical features, event indicators, and macroeconomic variables. The symbol $\dots$ indicates additional columns present in the dataset.

4.2 Modeling

The modeling phase constitutes the application of analytical techniques to the prepared data, aiming to develop predictive models that address the project's objectives. As illustrated in Figure 4.3, this phase encompasses the selection of appropriate modeling techniques, the design of a test strategy, the construction of models, and the evaluation of their predictive performance. The process is inherently iterative: model outcomes frequently prompt refinements in the choice of methods, parameter configurations, or even adjustments to the underlying data.

The definition of parameters and feature sets used throughout this section reflects the best combinations identified after extensive experimentation and multiple rounds of testing.

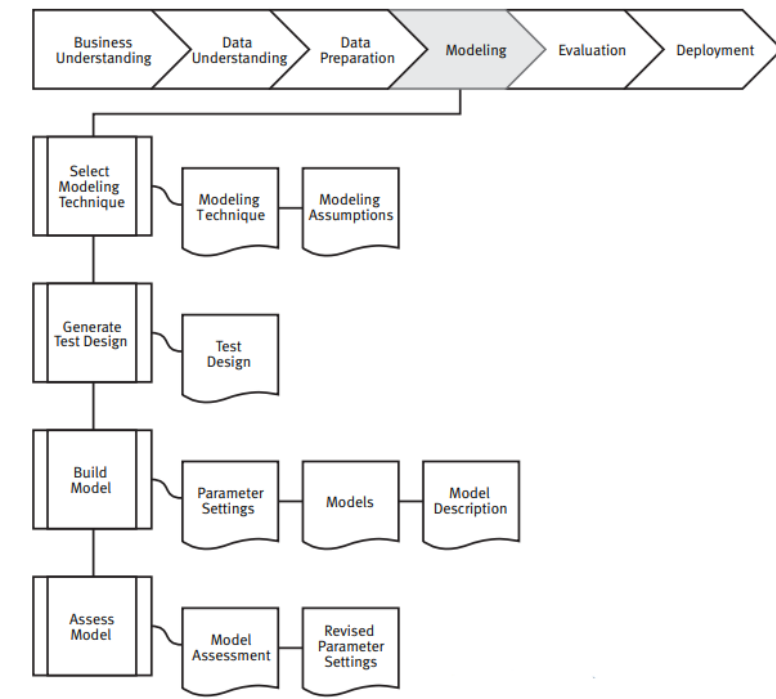


Figure 4.3: Detailed Modeling, Chapman (2000) , [Chapman \(2000\)](#)

In this study, the modeling phase was carefully structured to ensure that model development remained rigorous, transparent, and directly aligned with both business objectives and the specific characteristics of the available data. Detailed descriptions of each step—including technique selection, test design, model construction, and assessment—are presented in the following subsections.

4.2.1 Selection of Modeling Technique

The selection of modeling techniques in this study was informed by the literature review, the requirements identified during the business and data understanding phases, and the characteristics of the prepared dataset, particularly the inclusion of exogenous variables. To provide a robust and comprehensive evaluation, three complementary families of models were considered: statistical, machine learning, and deep learning. Each was chosen for its ability to capture different aspects of the data-generating process and is described in detail in the following subsections.

4.2.1.1 AutoARIMA

The ARIMA model was selected as the principal statistical baseline for this study. To implement this model, the `Nixtla` library was employed, which provides an `AutoARIMA`

algorithm that automatically identifies the optimal model parameters for a given time series (see [Nixtila \(2025\)](#)). The parameter selection process is based on established statistical and optimization techniques, such as the AIC and cross-validation, in order to determine the appropriate orders of parameters.

A key advantage of the AutoARIMA approach is its ability to generate prediction intervals, a functionality not universally available in other time series forecasting methods. This feature is particularly relevant in this study, as it provides a means of quantifying uncertainty around the forecasts, which is essential for robust decision-making in traffic and revenue planning.

Despite its advantages, AutoARIMA is subject to several important assumptions and limitations. First, the model requires that the time series be stationary, or made stationary through differencing. It further assumes linear relationships within the data and residuals that are uncorrelated and normally distributed. Moreover, when exogenous variables are included, the model presumes they are truly exogenous and synchronous with the dependent variable. These assumptions may not always hold in practice, which can affect the reliability of the results.

4.2.1.2 LightGBM

LightGBM was selected primarily due to its suitability for feature-rich forecasting tasks. The dataset developed in this work includes cyclical temporal encodings, event indicators, and macroeconomic regressors, which LightGBM is particularly effective at handling thanks to its ability to capture complex non-linear interactions between heterogeneous variables [Ke et al. \(2017b\)](#).

An additional motivation for its use is the support for quantile regression, which enables the construction of prediction intervals and therefore a direct quantification of forecast uncertainty—an aspect that is crucial for practical planning in this context. Beyond accuracy, LightGBM's efficiency and scalability make it attractive for real-world deployment, where regular retraining and rapid inference are required.

Finally, unlike statistical models such as ARIMA, LightGBM does not rely on assumptions of stationarity or linearity [Ke et al. \(2017b\)](#), making it more flexible for the type of non-stationary traffic data analyzed here.

4.2.1.3 AutoLSTM

The LSTM model constitutes the deep learning component of the modeling strategy. Its inclusion was motivated by its prominence in recent traffic forecasting research, particularly for short-term horizons, and by its ability to incorporate exogenous variables and

produce prediction intervals. These features are directly relevant to the objectives of this dissertation.

To facilitate implementation, the study employed the `AutoLSTM` model available in the `Nixtla` library, which automates hyperparameter tuning and network configuration. This approach simplifies the otherwise complex process of LSTM design while ensuring a fair comparison against statistical and machine learning baselines.

Nonetheless, as highlighted in the literature review, the performance of LSTMs typically benefits from very large datasets, enabling them to capture long-term temporal dependencies. Given the limited historical coverage of the present dataset (from 2022 onward), the ability of the LSTM to exploit such long-term structures may be constrained. Its role in this study is therefore both to provide a competitive benchmark and to assess the extent to which advanced neural architectures can add value under these conditions.

4.2.1.4 Summary

The choice of these three modeling strategies represents a thoughtful equilibrium among interpretability, predictive capability, and alignment with the data and the project’s operational goals. By employing a variety of techniques, the study can evaluate performance across different methodological frameworks, offering Ascendi practical insights for both current use and future methodological advancements.

4.2.2 Generate Test Design

A time-series cross-validation framework [Hyndman and Athanasopoulos \(2021\)](#) was designed to evaluate the models’ predictive capabilities. The dataset (January 1, 2022, to May 11, 2025) was first partitioned into training and test subsets, as illustrated in [Figure 4.4](#). All observations up to one calendar month before the series end were used for training; the final month constituted the test set.

Originally, the last three months were to be held out for testing. However, a structural change in Classes 2 and 3 within the urban context—beginning in mid-February 2025—necessitated revising this plan. Holding out three months would have deprived the models of sufficient exposure to the new regime, limiting their capacity to adapt. Consequently, only the final month was reserved for out-of-sample evaluation, allowing the training data to include the most recent post-change observations.

Within the training portion, cross-validation was implemented to assess performance across distinct historical periods (see [Figure 4.4](#)). Let T_{end} denote the dataset end date and let months be counted in calendar months. Three sequential splits were defined:

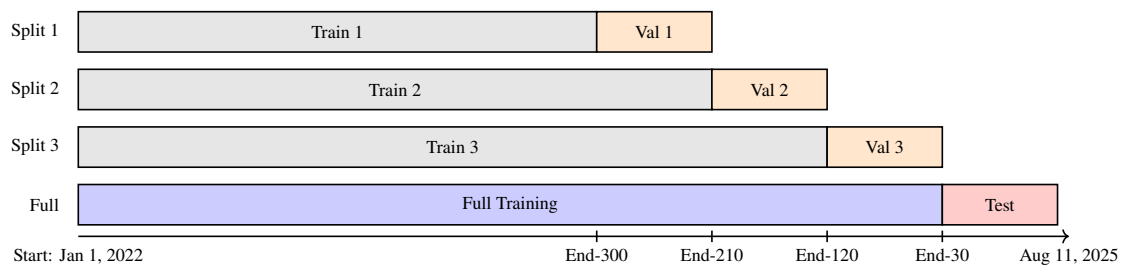


Figure 4.4: Time-series cross-validation for time series. From top to bottom: Split 1, Split 2, Split 3, and the final Full Training + Test.

- **Split 1:** Train from start to $T_{\text{end}} - 10$ months; validate on the subsequent 3 months.
- **Split 2:** Train from start to $T_{\text{end}} - 7$ months; validate on the subsequent 3 months.
- **Split 3:** Train from start to $T_{\text{end}} - 4$ months; validate on the subsequent 3 months.

This windowing logic was implemented explicitly to ensure full transparency and control over the evaluation protocol. After cross-validation, each model was retrained on all data available up to the end of the training period (i.e., excluding only the final test month) and used to produce forecasts up to three months ahead. In the final evaluation, accuracy metrics were computed exclusively on the first month of the held-out test period, ensuring a clean out-of-sample assessment.

This design mirrors the intended operational use case at Ascendi: models are trained on all information available at a given time, evaluated on the immediately following horizon, and then periodically retrained as new data arrive. The approach therefore assesses both average performance across multiple historical contexts and realistic behavior under deployment.

4.2.3 Build Model

The model construction process was conducted using an identical feature selection strategy for both urban and suburban environments, thereby ensuring comparability and robustness throughout the analyses. All models were developed based on the feature-engineered dataset described in subsection 4.1.6, and each modeling approach employed the same rolling window cross-validation strategy detailed in subsection 4.2.2. The following schematic (Figure 4.5) provides a visual overview of the workflow, while the subsequent subsections describe in detail the configuration and parameter selection for each modeling approach.

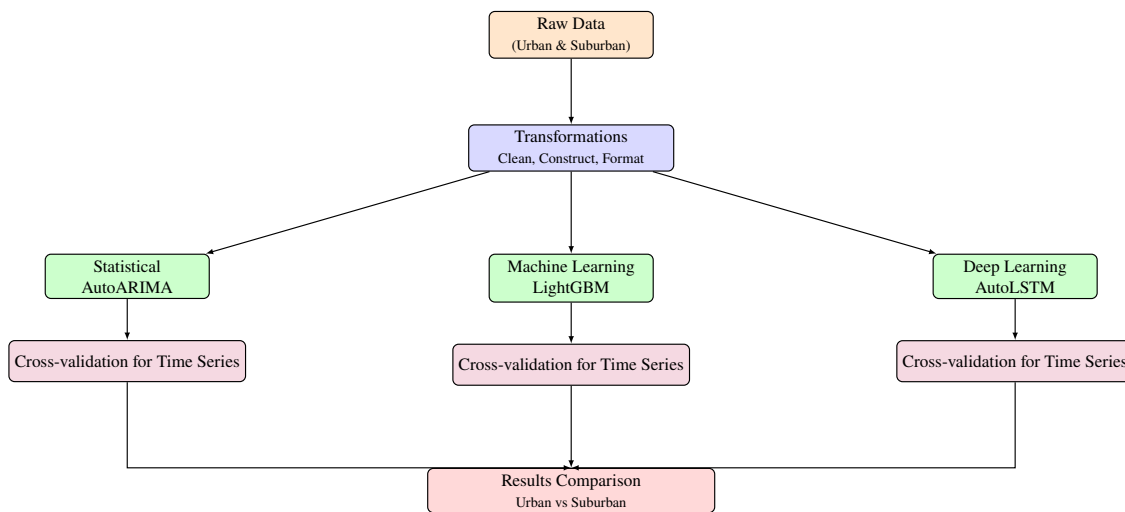


Figure 4.5: Unified modeling pipeline: Raw Data $\rightarrow$ Transformations $\rightarrow$ Model training (AutoARIMA, LightGBM, AutoLSTM), each validated with cross validation, followed by results comparison across urban and suburban contexts.

4.2.3.1 Model Configuration and Parameter Selection

Statistical Model: The statistical baseline was implemented using the `StatsForecast` library. AutoARIMA automatically selects the optimal autoregressive, differencing, and moving average parameters, with model identification performed independently within each training split. Weekly seasonality (`season_length=7`) was explicitly specified.

Lagged features were also incorporated: lags of 1, 7, 14, 21, and 28 days, together with smoothed rolling averages (3-day average of the 7-day lag and 7-day average of the 14-day lag). Prediction intervals at the 95% level were generated using rolling-window calibration (`level=95`). Variables found to be constant or highly collinear were automatically excluded to ensure valid estimation.

Machine Learning Model: The machine learning approach relied on the `MLForecast` library, employing the same lagged and derived features as in AutoARIMA, supplemented with all exogenous variables engineered during preprocessing. LightGBM was trained using quantile regression to minimize the 95% quantile loss, thereby producing both point forecasts and prediction intervals.

Hyperparameter optimization was conducted with *Optuna*, exploring: learning rate (0.01–0.2, log-uniform), maximum tree depth (2–10), number of estimators (200–1000), and number of leaves (7–63). These ranges follow established practices in gradient boosting to balance predictive accuracy, computational efficiency, and overfitting risk. After optimization within each split, the final model was retrained on the full training data.

Deep Learning Model: The deep learning component was implemented using the `AutoLSTM` module of the `NeuralForecast` library, which automates network configuration and hyperparameter tuning via `Optuna`. The search space included encoder hidden size (64 or 128 units), number of LSTM layers (2–4), learning rate (10^{-4} to 5×10^{-3} , log-uniform), maximum training steps (500 or 1200), and batch sizes (1 or 4). Random seeds (1–42) were also sampled to improve reproducibility.

Calendar features were included as future exogenous inputs (`futr_exog_list`), and macroeconomic indicators as historical exogenous variables (`hist_exog_list`). Unlike the other approaches, explicit lagged features were omitted, as LSTM networks are designed to learn temporal dependencies directly from sequential inputs. Training was again based on minimizing quantile loss at the 95% level, and forecasts were generated for horizons up to three months, ensuring comparability with the statistical and machine learning models across both environments.

4.2.4 Access Model

The evaluation of all candidate models was conducted to ensure consistency with the technical success criteria and performance objectives defined at the outset of the project. Model assessment relied exclusively on unseen validation and test sets, in accordance with the protocol described in Subsection 4.2.2, thereby guaranteeing a fair and robust comparison across approaches while preventing overfitting or information leakage.

The following subsections present the evaluation metrics selected for assessing model performance, together with their corresponding justification. In addition, the feature importance analysis performed after the model evaluation stage is also described, providing further insights into the relative contribution of different predictors.

4.2.4.1 Evaluation Metrics and Justification

For each model—AutoARIMA, LightGBM, and LSTM—the primary evaluation metrics selected were the MASE and the RMSSE. MASE was prioritized due to its interpretability and robustness to anomalies, providing a straightforward measure of average forecast accuracy relative to a naïve benchmark. RMSSE, on the other hand, was chosen for its ability to penalize large deviations more heavily, making it particularly useful in capturing the effect of extreme events such as holidays, which were identified during exploratory analysis.

To assess the quality of prediction intervals, two complementary metrics were employed. The CR quantifies the proportion of actual observations falling within the forecast intervals, thus indicating whether the intervals are well-calibrated. The nPIW was used

to evaluate the average width of prediction intervals relative to the mean of the series, providing a measure of efficiency in addition to reliability.

The choice of these metrics—discussed previously in the literature review—was primarily motivated by the need for scale-independent measures. Since the study involves comparisons across vehicle classes and environments with highly distinct scales, metrics insensitive to magnitude differences were required to ensure meaningful and fair comparisons across all series.

In addition to accuracy and interval-based metrics, the evaluation also considered the processing time of each model. This factor is particularly relevant in operational contexts where forecasting pipelines must run frequently and where computational cost scales with execution time in cloud environments. Thus, processing time served as an additional practical criterion to contextualize the feasibility of each method for deployment.

4.2.4.2 Feature Importance

Each distinct feature’s importance in various modeling strategies was systematically evaluated through an ablation process. In this method, variables were incrementally eliminated, and the subsequent impact on the forecasting mean metrics of the a rolling windows was examined (see Figure 4.6).

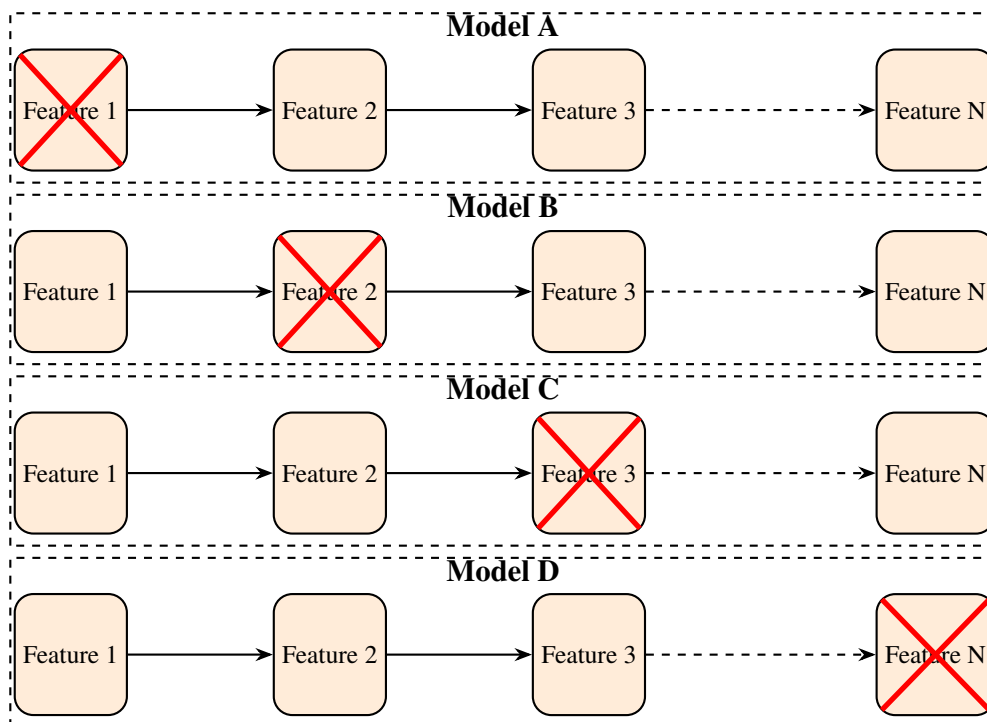


Figure 4.6: Feature Ablation Process

The results obtained from the feature importance analysis, as well as the evaluation outcomes, will be presented in the subsequent chapter, *Results and Discussion* (Chapter 4).

Chapter 5

Results and Discussion

This chapter outlines the concluding phases of the CRISP-DM approach, specifically *Evaluation* (Section 5.1) and *Deployment* (Section 5.2). These stages are primarily concerned with the outcomes and the final steps of the project.

In the *Evaluation* section, quantitative results are compared using the previously defined metrics (see Section 4.2.4.1) and assessed against the data mining objectives established earlier (see Subsection 3.1.2). The subsequent *Deployment* section presents a simplified implementation plan, in recognition of the fact that this dissertation is focused principally on the comparative study of forecasting models rather than on operational deployment.

Overall, this chapter provides a comprehensive overview of the results and their associated discussion, as well as a concise plan for the practical implementation of the most effective models.

5.1 Evaluation

The evaluation phase covers more than just assessing technical performance, in data mining projects, it is crucial to understand that the final outputs include both the predictive models and the supplementary insights gained during the analytical process, Chapman (2000). This perspective can be formalized as:

$$\text{RESULTS} = \text{MODELS} + \text{FINDINGS}$$

Here, MODELS represent the implemented predictive solutions, while FINDINGS encompass all insights and outcomes beyond the models themselves—including discoveries relevant to business strategy, data quality issues identified, or new questions and avenues

for future exploration. This dual focus ensures that the value generated by the project is fully captured, supporting both immediate application and long-term business learning.

Additionally, evaluation stage also requires re-examining the initial evaluation criteria established when the project began, to see if there are any business-related constraints or unmet needs that could make the solution unsuitable for implementation.

5.1.1 Evaluate Results

The evaluation of the developed models was carried out by comparing them against the original business objectives and the data mining success criteria established at the beginning of the project (see subsections 3.1.1.2 and 3.1.1.3). The main aim was to provide precise and dependable predictions of daily traffic volumes for all vehicle classes, for both urban *BLA* and suburban *GP* highway segments, with a forecasting range of up to three months. Additional objectives included the detection of relevant temporal patterns and anomalies, ensuring model generality across the three-month forecasting horizon, and supporting revenue planning and financial risk management at Ascendi by achieving superior performance compared to the currently implemented model.

5.1.2 Feature Importance Results

The ablation study conducted with AutoARIMA, LightGBM and AutoLSTM reveals that a concise set of lagged traffic counts and basic calendar signals already explains most of the variation in both urban and suburban networks. Annual sine-cosine cycles, a weekend dummy and a public-holiday flag emerged as universally indispensable. In AutoARIMA they guided the model's seasonal differencing, in LGBM they drove calendar-based splits, and in AutoLSTM they supplied periodic cues that speeded convergence. For all Classes, these three calendar terms alone were enough for AutoLSTM in both environments.

The statistical and machine-learning implementations (AutoARIMA and LGBM) also required the 1-, 6- and 7-day lags for every vehicle class. Light-vehicle volumes (Class 1), however, showed longer-range dependence. Adding 14-, 21- and 28-day lags markedly improved multi-step accuracy, while for LGBM the inclusion of monthly sine-cosine pairs and higher-order interactions such as “holiday × weekend” and “adjacent-weekend” provided further gains.

The ExpoMoto event indicator failed to improve motorcycle forecasts in both contexts; with no exhibitions occurring from 2024 and none planned until 2026, the variable was omitted. Macroeconomic indicators, GDP and CPI, provided only a slight boost to light

Table 5.1: Exogenous features used by each model.

Model	Exogenous features (excluding lags)
AutoARIMA	is_weekend, is_holiday, day_of_year_sin, day_of_year_cos, month_number_sin, month_number_cos, gdp ¹ , cpi ¹ , is_expomoto ² , toll_tax_change ³
MLForecast	is_holiday_x_weekend, day_of_year_sin, day_of_year_cos, week_number_sin, week_number_cos, month_number_sin, month_number_cos, quarter_number_sin, quarter_number_cos, is_weekend, is_holiday, is_adjacent_weekend, gdp ¹ , cpi ¹ , is_expomoto ² , toll_tax_change ³
AutoLSTM	is_weekend, is_holiday, day_of_year_sin, day_of_year_cos, week_number_sin, week_number_cos, gdp ¹ , cpi ¹ , is_expomoto ² , toll_tax_change ³

¹ Used only for Class 1 (light vehicles). ² Used only for Class 5 (motorcycles, urban and suburban). ³

Used only for Classes 2 and 3 in the urban environment.

vehicles and motorcycles while consistently worsening the performance of commercial flows, which resulted in their removal from the light vehicles final set of features.

Because the cross-validation windows did not span a period in which the toll-tax change could be learned reliably, its incremental value could not be reassessed. Nonetheless, the dummy was retained for Classes 2 and 3 in the suburban AutoARIMA and LGBM models so that its potential effect can be tested in the hold-out evaluation.

Throughout, the rank ordering of the retained variables proved remarkably stable between urban and suburban data, affirming that the dominant drivers of traffic demand are similar across contexts. In short, a trio of calendar signals plus short lags suffice for commercial traffic, whereas light-vehicle and motorcycle series benefit from longer memory. Broader economic indicators play, at most, a secondary role in the forecasting architecture based on AutoARIMA, LGBM and AutoLSTM.

Table 5.1 summarizes the calendar and event-based features retained for each model (lag variables are omitted for brevity).

5.1.3 Results

The results are reported in two separate tables, each summarizing the comparative performance of the evaluated models—AutoARIMA, LightGBM, and AutoLSTM—across both contexts (urban and suburban) and all vehicle classes. Table 5.1.3 presents the average

accuracy and coverage metrics obtained over the three cross-validation validation splits, while Table 5.1.3 reports performance for the first month of the hold-out test horizon.

Separating the results in this manner enables a clear, side-by-side comparison of point- and interval-forecast accuracy for each modeling approach, thereby highlighting the strengths and limitations identified in the preceding analysis. In addition, Figures 5.1–5.6 augment the tabulated hold-out metrics by showcasing the predictive performance of each model visually across various vehicle categories and environments. These charts demonstrate the models’ ability to capture weekly patterns, holidays, and regime changes, highlighting contrasts in forecast uncertainty between urban (GP) and suburban (BLA) highways. The historical data depicted in the images were limited to 6 months to simplify the visual representation. Together, the two tables and the corresponding figures offer the empirical basis for the forthcoming discussion concerning model ranking and selection.

Model	Scenario	Class	MASE	RMSSE	CR	nPIW
AutoARIMA	Urban (GP)	1	1.1713	1.2687	0.6333	0.2440
		2	1.0382	0.9051	0.3000	0.1716
		3	1.3457	1.1720	0.2333	0.2010
		4	0.4373	0.3620	0.5667	0.4184
		5	2.2162	1.8978	0.5000	1.6564
	Suburban (BLA)	1	1.3309	1.2776	0.8963	0.7155
		2	0.3156	0.3251	0.8333	0.3614
		3	0.4538	0.4635	0.7667	0.6674
		4	0.4973	0.4641	0.7630	0.4983
		5	1.0121	0.9498	0.7926	1.6494
LightGBM	Urban (GP)	1	0.6121	0.6291	0.7815	0.1826
		2	0.4369	0.4450	0.6370	0.2221
		3	0.8696	0.8455	0.7074	0.5050
		4	0.1530	0.2225	0.7741	0.2508
		5	1.0122	0.9712	0.8370	1.2301
	Suburban (BLA)	1	0.6787	0.6782	0.8407	0.2055
		2	0.2211	0.2340	0.7815	0.2102
		3	0.4202	0.4174	0.8037	0.6466
		4	0.2367	0.2681	0.6000	0.1879
		5	0.8359	0.8370	0.8667	1.5892
AutoLSTM	Urban (GP)	1	1.3036	1.2192	0.7074	0.2853
		2	0.5515	0.5704	0.7630	0.8619
		3	0.8891	0.9241	0.7630	0.8783
		4	0.5914	0.5846	0.9926	1.2855
		5	1.3068	1.1869	0.9556	2.6090
	Suburban (BLA)	1	1.2753	1.1838	0.7481	0.4551
		2	0.4748	0.4524	0.7370	0.7087
		3	0.5153	0.5052	0.9037	1.1885
		4	0.4286	0.4468	0.8185	0.6345
		5	1.6550	1.3711	0.9704	3.5912

Table 5.2: Average performance of AutoARIMA, LightGBM and AutoLSTM across three cross-validation windows. Metrics reported separately for urban (GP) and suburban (BLA) highways by vehicle class.

Model	Scenario	Class	MASE	RMSSE	CR	nPIW
AutoARIMA	Urban (GP)	1	1.1713	1.2687	0.6333	0.2440
		2	1.0382	0.9051	0.3000	0.1716
		3	1.3457	1.1720	0.2333	0.2010
		4	0.4373	0.3620	0.5667	0.4184
		5	2.2162	1.8978	0.5000	0.7564
	Suburban (BLA)	1	0.9257	0.8469	0.8667	0.9599
		2	0.4361	0.4076	0.6667	0.3763
		3	0.5714	0.5194	0.6000	0.6437
		4	0.4349	0.3496	0.8333	0.4863
		5	2.2063	1.9508	0.6000	0.7210
LightGBM	Urban (GP)	1	0.8147	0.9830	0.7000	0.2259
		2	0.9714	0.7849	0.4667	0.2990
		3	0.9869	1.0269	0.6667	0.4553
		4	0.3281	0.3094	0.5333	0.2004
		5	1.8611	1.6763	0.7333	0.7142
	Suburban (BLA)	1	1.3243	1.2198	0.4333	0.3943
		2	0.5716	0.4485	0.2333	0.2322
		3	0.5061	0.4833	0.7667	0.3794
		4	0.3324	0.2754	0.5000	0.1736
		5	2.2312	2.0706	0.7333	0.4073
AutoLSTM	Urban (GP)	1	1.0336	1.1484	0.7667	0.3425
		2	1.2073	0.9504	1.0000	0.9717
		3	2.2917	1.8130	1.0000	0.8612
		4	0.2716	0.2405	0.8667	0.2775
		5	1.9469	1.6375	0.9000	0.9858
	Suburban (BLA)	1	1.2299	1.1686	0.8000	0.5669
		2	0.3796	0.3198	0.6667	0.3143
		3	0.4058	0.4658	0.9667	0.5517
		4	0.3662	0.2970	0.7667	0.2967
		5	2.2354	1.9188	0.7667	0.5693

Table 5.3: Performance of AutoARIMA, LightGBM and AutoLSTM on the hold-out test set (first month ahead). Metrics reported separately for urban (GP) and suburban (BLA) highways by vehicle class.

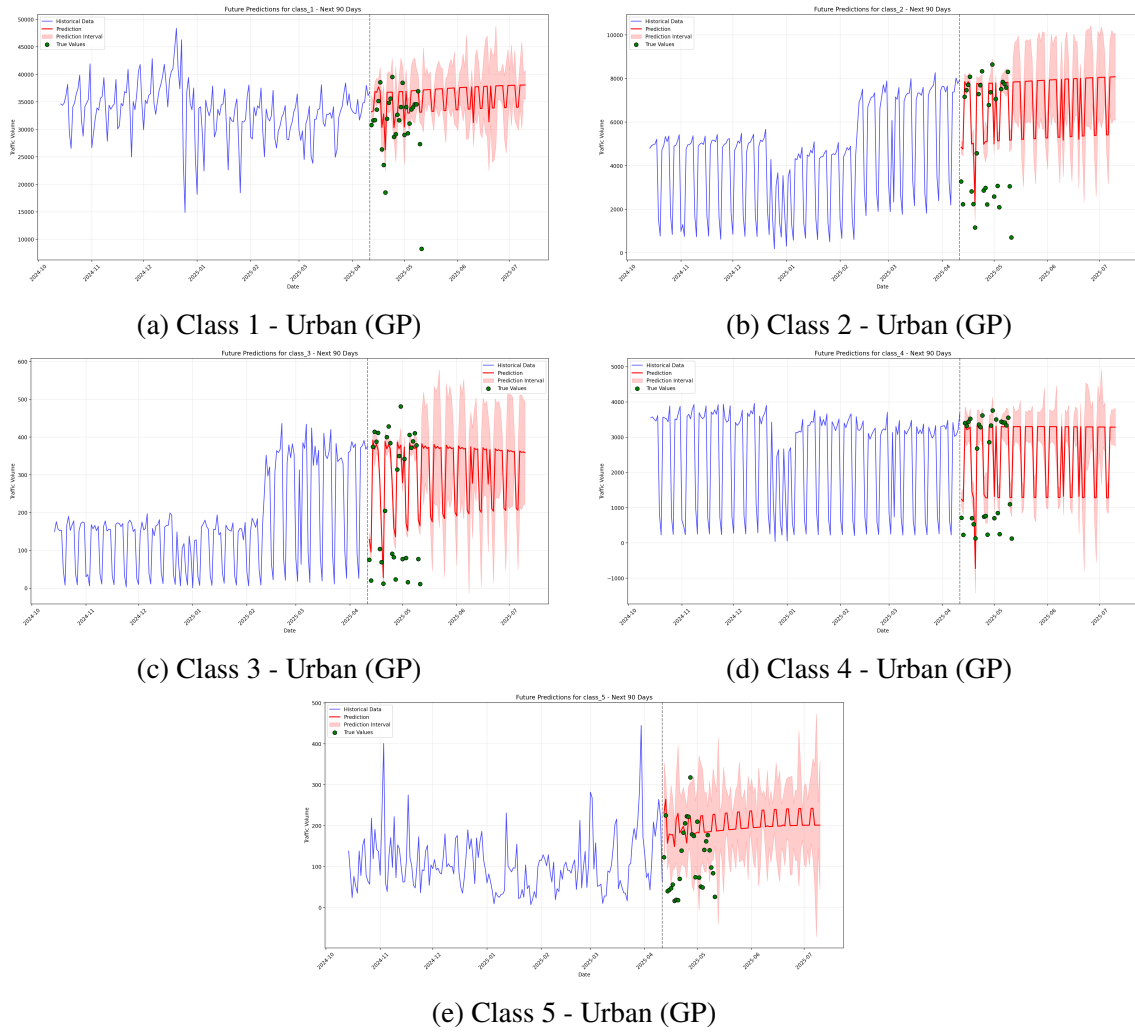


Figure 5.1: Forecasts of AutoARIMA for all vehicle classes in the urban (GP) environment. Each panel shows observed values (blue), point forecasts (red), 95% prediction intervals (shaded), and ground truth values (green).

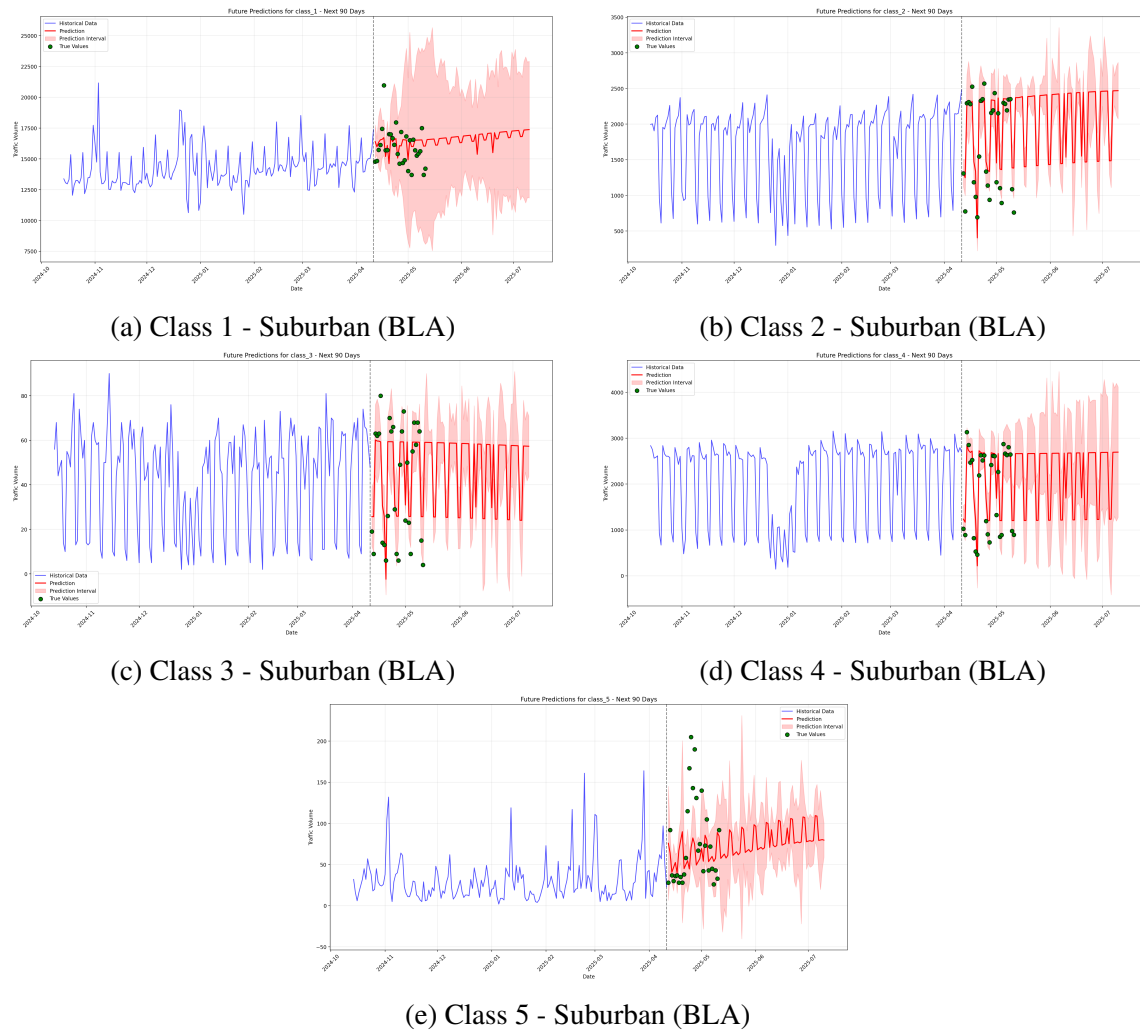


Figure 5.2: Forecasts of AutoARIMA for all vehicle classes in the Suburban (BLA) environment. Each panel shows observed values (blue), point forecasts (red), 95% prediction intervals (shaded), and ground truth values (green).

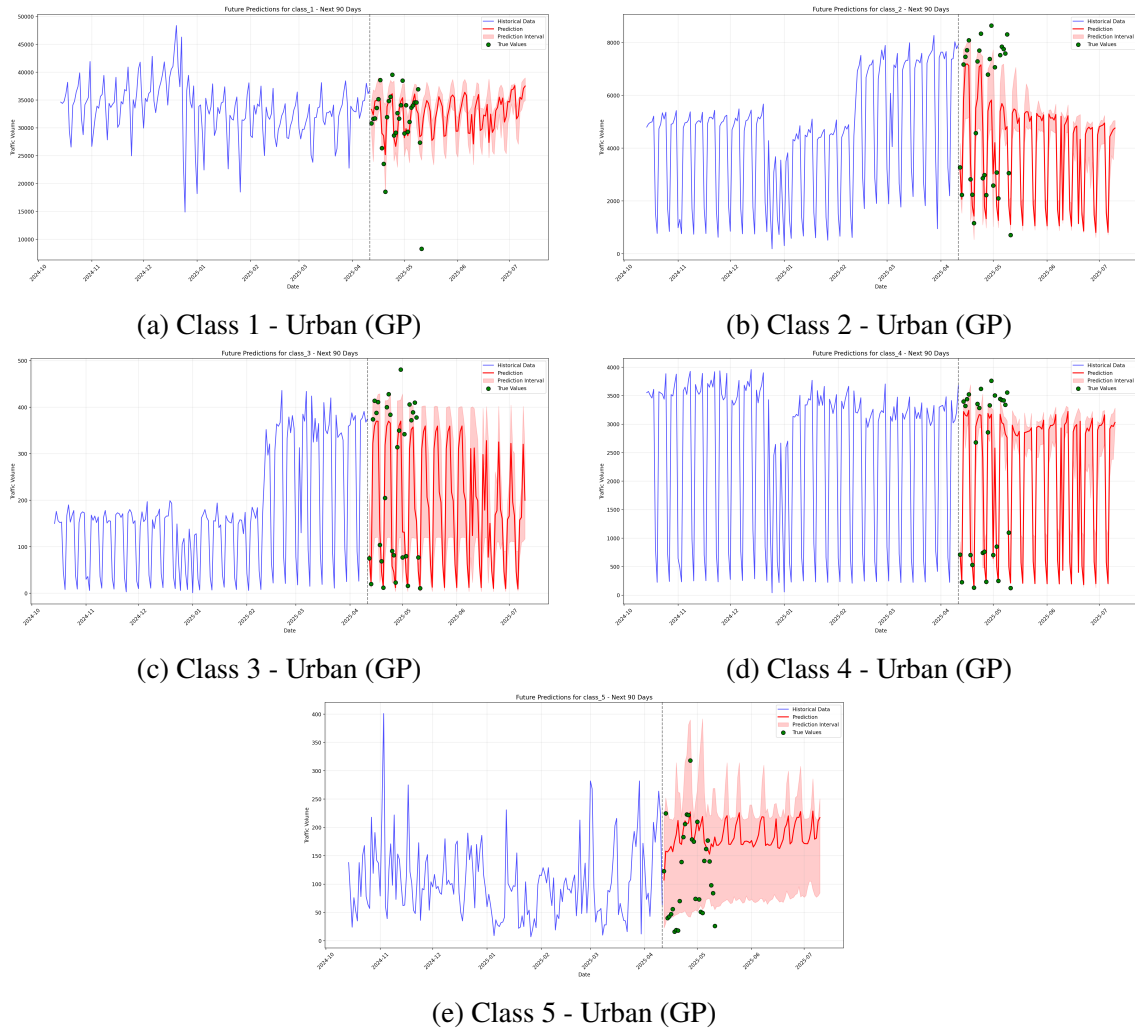


Figure 5.3: Forecasts of LGBM for all vehicle classes in the urban (GP) environment. Each panel shows observed values (blue), point forecasts (red), 95% prediction intervals (shaded), and ground truth values (green).

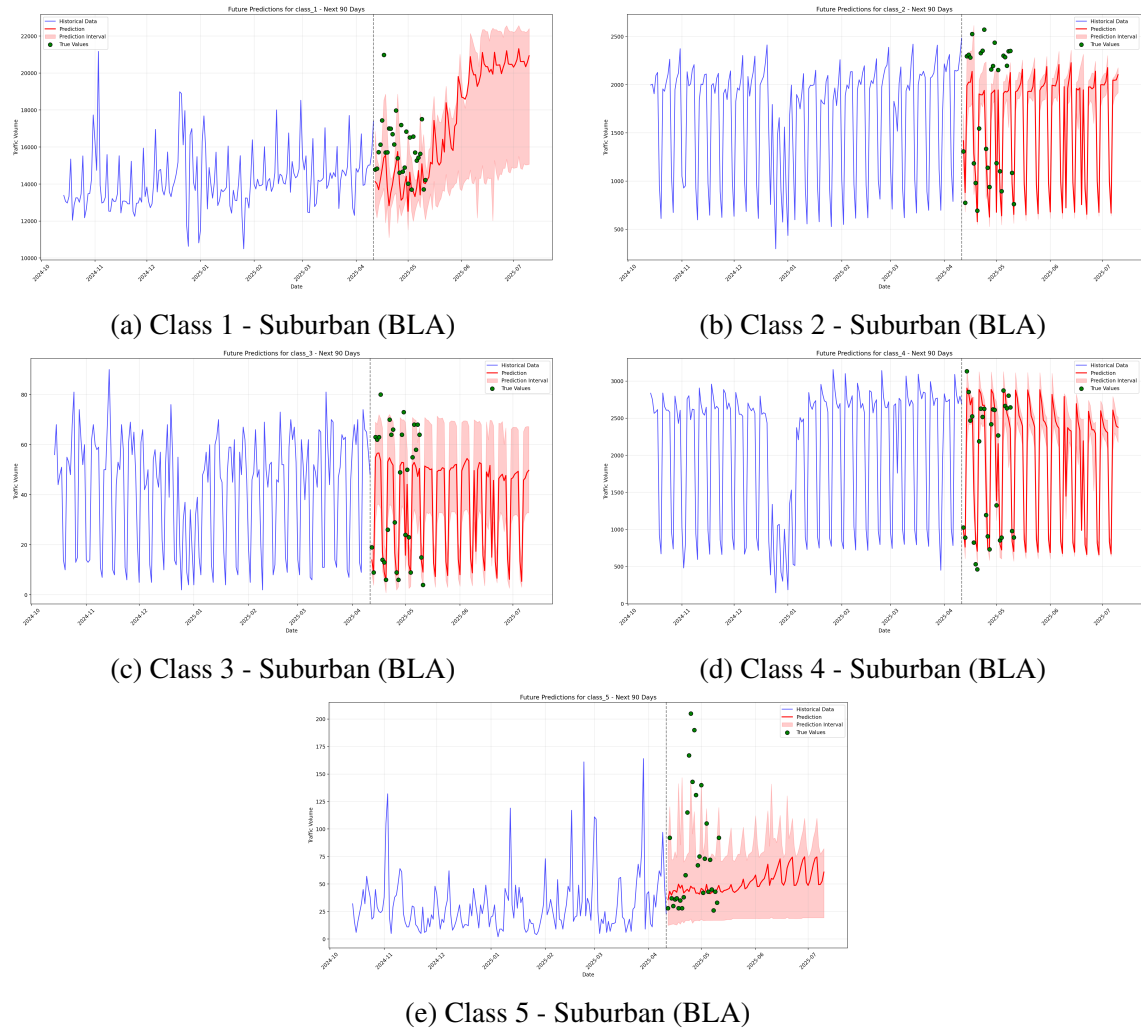


Figure 5.4: Forecasts of LGBM for all vehicle classes in the Suburban (BLA) environment. Each panel shows observed values (blue), point forecasts (red), 95% prediction intervals (shaded), and ground truth values (green).

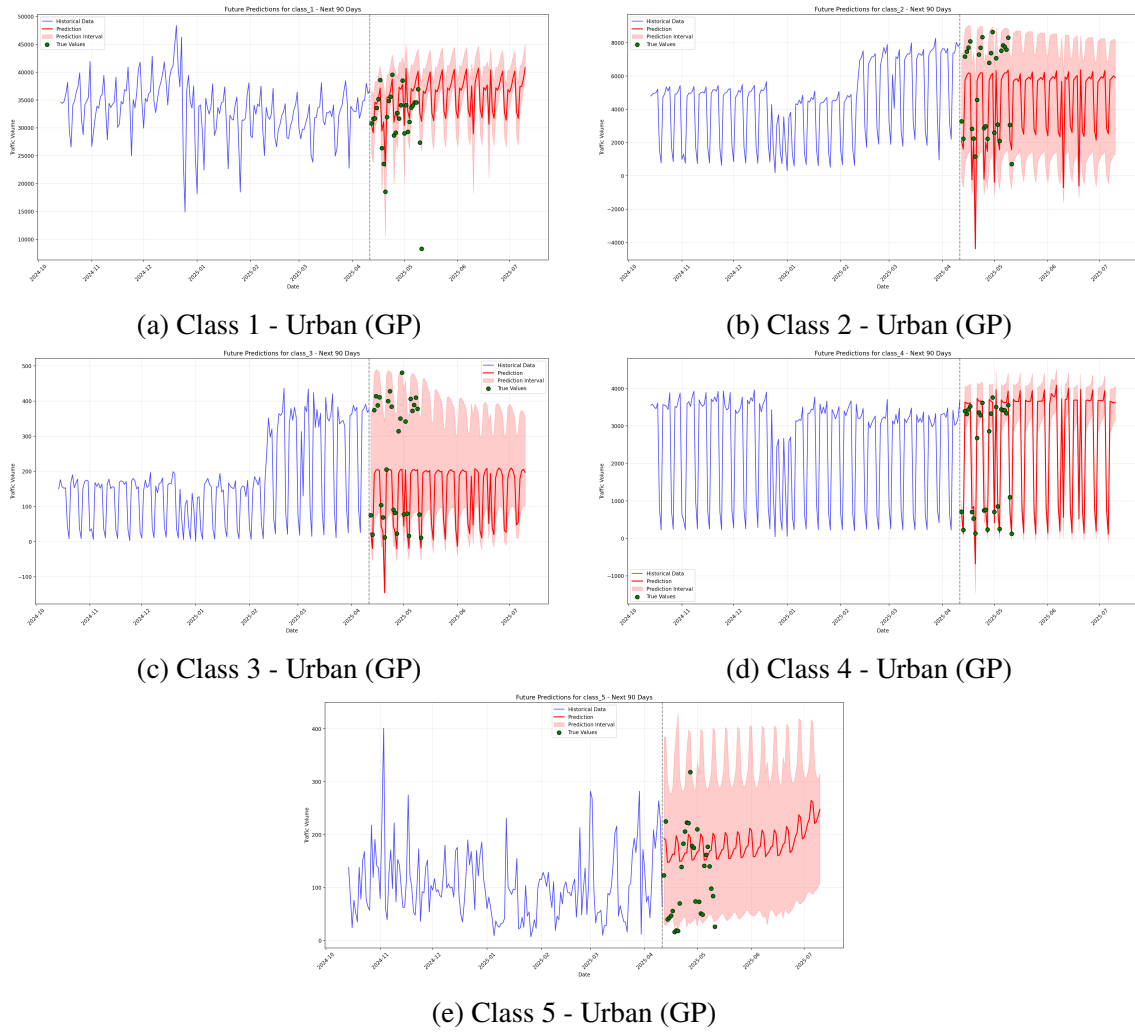


Figure 5.5: Forecasts of AutoLSTM for all vehicle classes in the urban (GP) environment. Each panel shows observed values (blue), point forecasts (red), 95% prediction intervals (shaded), and ground truth values (green).

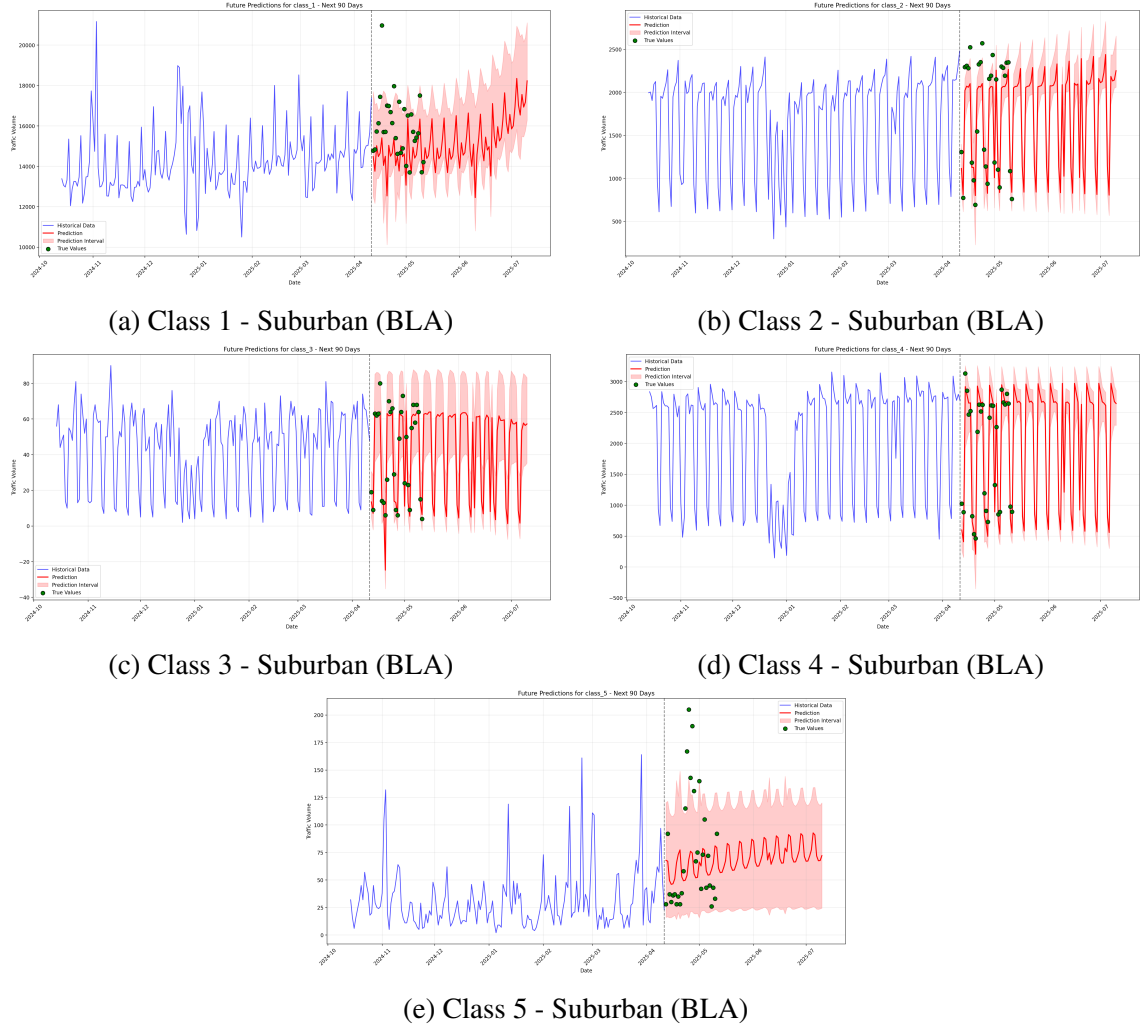


Figure 5.6: Forecasts of AutoLSTM for all vehicle classes in the Suburban (BLA) environment. Each panel shows observed values (blue), point forecasts (red), 95% prediction intervals (shaded), and ground truth values (green).

5.1.4 Discussion

This subsection discusses the empirical findings from both the rolling cross-validation and the hold-out evaluation. It first contrasts model performance across *urban* and *suburban* settings, using MASE/RMSSE for point-forecast accuracy and CR/nPIW to assess the sharpness and calibration of prediction intervals. The hold-out analysis is then reinforced with the accompanying figures, which illustrate salient patterns and help diagnose where each model succeeds or fails. The section concludes with a synthesis of the main insights and a discussion of practical implications for model selection and deployment—highlighting trade-offs between accuracy and uncertainty, the sensitivity to data

drift, and the suitability of each approach (AutoARIMA, LightGBM, AutoLSTM) for real-world traffic and revenue-planning contexts.

5.1.4.1 Mean Rolling Windows Performance

The mean rolling-window results indicate that the suburban setting is consistently easier to forecast than the urban setting. For **AutoARIMA**, suburban classes 2–4 improve markedly on the seasonal-naïve baseline ($\text{MASE} \approx 0.32\text{--}0.50$; $\text{RMSSE} \approx 0.33\text{--}0.46$), whereas the corresponding urban classes are at or above baseline (e.g., urban class 3: $\text{MASE} = 1.35$, $\text{RMSSE} = 1.17$). **LightGBM** outperforms the baseline across all suburban classes and nearly all urban classes, observing only one marginal instance in the urban classification 5 ($\text{MASE} \approx 1.01$). **AutoLSTM** benefits from the suburban context for classes 2–4 but degrades on classes 1 and 5 in both settings (e.g., suburban class 5: $\text{MASE} \approx 1.66$; urban class 1: $\text{MASE} \approx 1.30$). RMSSE mirrors MASE, suggesting that these patterns are not driven by isolated outliers.

Uncertainty metrics reinforce these conclusions. In the urban setting, **LightGBM** maintains reasonable coverage with comparatively sharp intervals ($\text{CR} \sim 0.64\text{--}0.84$, $\text{nPIW} \sim 0.18\text{--}0.51$), except for class 5, where wider intervals ($\text{nPIW} \sim 1.23$) are required to sustain coverage ($\text{CR} \sim 0.84$). **AutoARIMA** shows modest urban coverage in mid classes (e.g., class 2: $\text{CR} \sim 0.30$) and very wide yet only moderately covering intervals for class 5 ($\text{nPIW} \sim 1.66$, $\text{CR} \sim 0.50$). **AutoLSTM** attains very high coverage in urban classes 4–5 ($\text{CR} \sim 0.96\text{--}0.99$) primarily by inflating interval width ($\text{nPIW} \sim 1.29\text{--}2.61$).

In the suburban setting, all models report higher coverage for class 5 but at the cost of very wide intervals (e.g., AutoLSTM $\text{nPIW} \sim 3.59$, $\text{CR} \sim 0.97$). For the easier classes 2–4, **LightGBM** balances coverage and sharpness effectively (e.g., $\text{nPIW} \sim 0.19\text{--}0.21$, $\text{CR} \sim 0.60\text{--}0.78$), while **AutoARIMA** provides adequate coverage with moderate widths ($\text{CR} \sim 0.76\text{--}0.83$, $\text{nPIW} \sim 0.36\text{--}0.67$). **AutoLSTM** again tends to achieve higher coverage through broader prediction intervals.

In general, **LightGBM** offers the most favorable accuracy–uncertainty trade-off across scenarios and series that demonstrate lower volatility—commonly linked with commercial vehicle types—exhibit greater predictability across different scenarios. In contrast, series associated with leisure-focused and more volatile vehicle classes pose challenges for modeling accuracy, necessitating broader intervals to achieve similar levels of coverage, results that align with those highlighted in the article [Lingras et al. \(2000\)](#).

5.1.4.2 Hold-out Performance

The hold-out evaluation reinforces a clear environmental effect—*suburban* traffic is easier to forecast than *urban* traffic—largely because the urban series exhibit stronger distributional drift. Across both environments, the models captured the Good Friday (18/04) holiday pattern for all vehicle classes *except* class 5.

Regarding the *urban* time series, **AutoARIMA** improves on the baseline only for class 4 (MASE = 0.44), while classes 1–3 remain at or above baseline and class 5 performs poorly (MASE = 2.22). Visual inspection indicates limited adaptation to changing patterns, although the model does register drift in classes 2–3. **LightGBM** outperforms the baseline in all urban classes except class 5, with a strong gain in class 4 (MASE = 0.33); it captures early drift but does not sustain this advantage at longer horizons. **AutoLSTM** is most competitive in class 4 (MASE = 0.27) and reflects core traffic regularities, but struggles with the pronounced drift in classes 2–3—consistent with higher data requirements. Class 5 remains challenging across models.

In the *suburban* environment, **AutoARIMA** is competitive for classes 2–4 (e.g., class 4 MASE = 0.43) and deteriorates for class 5 (MASE = 2.21); the summer increase in class 1 is captured mainly via wider intervals. **LightGBM** improves on the baseline for classes 2–4 (class 4 MASE = 0.33) and remains competitive for commercial-vehicle classes, but performs weakly on class 5. **AutoLSTM** mirrors these gains for classes 2–4 (MASE $\approx$ 0.38–0.41) and generally surpasses LightGBM in this setting, while still underperforming on class 5.

Uncertainty behavior follows accuracy. In urban traffic, **LightGBM** typically attains moderate coverage with sharp intervals (class 4: CR = 0.53, nPIW = 0.20), whereas class 5 requires noticeably wider bands (nPIW = 0.71) to sustain coverage (CR = 0.73). **AutoARIMA** under-covers in urban mid classes (e.g., class 2: CR = 0.30, nPIW = 0.17) and still struggles with class 5 despite widening (CR = 0.50, nPIW = 0.76). **AutoLSTM** often achieves high urban coverage in classes 2–3 (CR = 1.00) primarily by broadening intervals (nPIW $\approx$ 0.86–0.97). In suburban traffic, **AutoARIMA** balances coverage and sharpness for classes 2–4 (CR $\approx$ 0.60–0.83), **LightGBM** keeps intervals narrow for class 4 (nPIW = 0.17) but can under-cover in class 2 (CR = 0.23), and **AutoLSTM** secures higher coverage for stable classes (e.g., class 3: CR = 0.97, nPIW = 0.55), with class 5 remaining the most uncertain.

5.1.4.3 Synthesis and practical implications

Overall, **LightGBM** proved strongest in the *urban* context, while **AutoLSTM** performed best in the *suburban* context. All models captured salient traffic regularities—more

cautiously in the case of AutoARIMA and more assertively for LightGBM and AutoLSTM—indicating an ability to represent the main demand drivers despite distributional drift. In broader terms, LightGBM stood out for its adaptability to real-world conditions, delivering the most consistent results in the rolling cross-validation and remaining competitive on the hold-out test. Its lower data requirements likely contributed to this robustness, in contrast to the LSTM’s greater data needs, which can limit adaptation under abrupt shifts. Although AutoARIMA maintained structural changes over longer horizons, its point accuracy and prediction intervals were generally less competitive than the machine- and deep-learning models.

As expected—and in line with prior work—more stable series (commercial-vehicle classes) were easier to forecast, whereas the most volatile series (motorcycles, class 5) were consistently challenging in both environments. Even so, the relative ranking is informative: in the urban setting LightGBM remained comparatively strongest for class 5, while in the suburban setting AutoLSTM led. Despite higher error levels for this class, these patterns are valuable for risk signaling and budgeting decisions in revenue planning.

In terms of processing time, the AutoARIMA model exhibited the quickest duration, closely followed by LightGBM, with both completing training and forecasting in a few seconds per time series. In contrast, the LSTM model required substantially more computational effort, taking on average approximately *1.5 minutes* per time series. All experiments were executed on an `m1.c5.xlarge` AWS instance, highlighting the relative efficiency advantages of AutoARIMA and LightGBM for operational deployment.

5.2 Deployment

The deployment phase proposed in this study is intentionally presented as a simplified implementation plan, as the primary focus of this dissertation lies in the comparative analysis of forecasting models rather than in full-scale operational integration. Nonetheless, the empirical evidence obtained offers clear guidance on how such a system could be effectively transitioned into production within Ascendi’s existing cloud infrastructure.

Considering the combined factors of model performance, robustness under distributional changes, and computational efficiency, the **LightGBM** model emerges as the most suitable candidate for deployment across both motorway segments. Its consistently strong results in cross-validation, competitive performance on the hold-out test, and ability to adapt to covariate shifts—particularly within the urban network—highlight its reliability in real-world conditions.

From an operational perspective, LightGBM also provides meaningful economic advantages. Its short processing time, measured in only a few seconds per time series, aligns well with cloud-based execution models where costs scale with compute time. Deploying a forecasting pipeline built around LightGBM would therefore limit resource consumption and reduce ongoing expenditures associated with AWS usage, making the approach financially viable for continuous operation.

A minimal deployment plan would involve recommending the model for production use and establishing a **weekly retraining** routine. This frequency balances the need for maintaining model adaptability in the presence of evolving traffic patterns with the practical constraints of computational cost. Coupled with routine monitoring of prediction accuracy and interval calibration, this would ensure early detection of performance drifts and help maintain forecasting reliability over time.

Importantly, implementing this forecasting framework would enable Ascendi to transition from its current **monthly** forecasting cycle to a **daily, automated** forecasting workflow. This shift toward higher temporal resolution would grant the company near-real-time visibility into expected traffic demand, supporting more granular and responsive revenue planning. Given that toll tariffs are externally regulated and cannot be adjusted internally, improved forecasting precision becomes a key strategic asset for financial planning and resource allocation across the network.

Overall, although simplified, the proposed deployment plan highlights the practical feasibility and economic attractiveness of adopting LightGBM as the operational forecasting model. Its integration would not only enhance predictive accuracy but would also contribute positively to cost efficiency and planning capabilities within the organization.

Chapter 6

Conclusions and Future Work

This dissertation compared daily traffic-flow forecasting performance across two highway contexts—an *urban* (GP) segment and a *suburban* (BLA) segment—over forecasting horizons of up to three months. Three modeling families were assessed within a unified, time-series-aware evaluation pipeline: ARIMA, Gradient Boosting Machines (LightGBM), and LSTM. All models successfully generated three-month forecasts, improving upon Ascendi’s manual baseline by delivering reproducible results, consistent treatment of calendar effects, and explicit uncertainty quantification through prediction intervals.

LightGBM emerged as the most robust and operationally suitable approach across both contexts. It offered a balanced combination of point accuracy, processing time and informative uncertainty estimates and, importantly, adapted most effectively to the shift observed in the urban dataset. This adaptability is particularly valuable in production settings, where external events can abruptly alter traffic behavior. Furthermore, LightGBM achieved the best overall performance when evaluated through the cross-validation rolling-window framework.

The cross-context comparison revealed that the underlying data structures were highly similar, differing mainly in traffic volume and the strength of seasonal influences. Across both segments, the models consistently captured weekly seasonality and holiday effects, with forecast intervals widening appropriately during such periods—providing clear risk indicators for operational planning. In contrast, macroeconomic indicators did not yield meaningful improvements at the daily resolution or within the three-month horizon considered, reinforcing that calendar structure and recent temporal dynamics remain the dominant explanatory factors for this task. Traffic profiles were found to be strongly influenced by vehicle usage patterns, and motorcycle traffic, in particular, exhibited the highest volatility and lowest predictability across both contexts. Consequently, motorcycle flows should be down-weighted when used for revenue forecasting or long-term planning.

6.1 Future Work

Future work should first focus on bringing the proposed forecasting framework—particularly the LightGBM model—into a production environment and monitoring its behavior over an extended period. Conducting such a longitudinal study would enable a more thorough evaluation of the model’s performance in real-world conditions, particularly considering changing traffic patterns and distributional shifts.

Building on this operational deployment, future research could explore the implementation of **hierarchical reconciliation methods** to ensure coherence between forecasts generated at different aggregation levels. This would be especially beneficial in infrastructure networks where decision-making often depends on aligned predictions across multiple spatial or functional layers.

Another promising direction involves developing **specialized modeling strategies for motorcycle traffic**. Given the high volatility and sensitivity of this vehicle class to external factors, targeted methodological enhancements could contribute substantially to improving overall forecasting accuracy and robustness.

From an analytical perspective, future studies may also incorporate **statistical significance tests**, such as the Wilcoxon signed-rank test, to formally evaluate differences in model performance. Employing such tests would strengthen the empirical support for conclusions drawn in comparative analyses and verify whether observed improvements—such as those attributed to LightGBM—are statistically meaningful rather than attributable to random variation.

Appendix A

Vehicles Classification

A.1 Toll Classes in Portugal

The Portuguese toll system classifies vehicles into five categories, primarily according to their height at the front axle, number of axles, gross vehicle weight, and, in specific cases, emission standards. This classification determines the toll rates applied on national highways and is regulated mainly by Decree-Law No. 39/2005 ([República Portuguesa \(2005\)](#)), and Decree-Law No. 71/2018 ([República Portuguesa \(2018\)](#)).

Descriptive Overview

A.1.0.1 Class 1

Vehicles with a height at the front axle of less than 1.10 meters, typically including most passenger cars and motorcycles. According to the exceptions introduced by Decree-Law No. 39/2005([República Portuguesa \(2005\)](#)) and Decree-Law No. 71/2018([República Portuguesa \(2018\)](#)), certain vehicles such as SUVs, minivans, and crossovers with two axles, gross weight up to 2,300 kg (or up to 3,500 kg if seating five or more), and a height between 1.10 m and 1.30 m, may also qualify as Class 1 if they are not equipped with permanent or switchable four-wheel drive and (for vehicles registered after January 1, 2019) comply with EURO 6 emission standards ([Union \(2007\)](#)).

A.1.0.2 Class 2

Vehicles with two axles and a height at the front axle of 1.10 meters or above, with a gross weight below 3,500 kg. This class generally comprises light commercial vehicles such as delivery vans.

A.1.0.3 Class 3

Vehicles with three axles, or with a gross weight over 3,500 kg and a height at the front axle of at least 1.10 meters. This class typically includes small buses and some heavy-duty vehicles.

A.1.0.4 Class 4

Vehicles with four or more axles (including trailers and semi-trailers) and a height at the front axle of at least 1.10 meters. Examples are large trucks and buses.

A.1.0.5 Class 5

A special category introduced for motorcycles using the electronic toll collection system Via Verde, which provides a 30% discount compared to the Class 1 rate (except on the Vasco da Gama Bridge in Lisbon), [Via Verde Portugal \(2024\)](#).

Summary Table

Table A.1: Toll Classes in Portugal

Class	Criteria	Examples
1	Height at front axle < 1.10 m; includes some exceptions for vehicles with height 1.10–1.30 m, specific weight, seats, no 4WD, and EURO 6 compliance	Passenger cars, motorcycles, certain SUVs/minivans/crossovers
2	2 axles; height at front axle $\geq$ 1.10 m; gross weight < 3,500 kg	Large vans, light commercial vehicles
3	3 axles, or gross weight > 3,500 kg and height at front axle $\geq$ 1.10 m	Small buses, heavy vehicles
4	4 or more axles; height at front axle $\geq$ 1.10 m	Large trucks, articulated lorries, big buses
5	Motorcycles using Via Verde system (electronic toll tag); 30% discount from Class 1 (except Vasco da Gama Bridge)	Motorcycles with electronic toll device

Appendix B

Pattern Analysis by Vehicle Class and Environment

This appendix provides the detailed pattern analysis for each vehicle class across both urban and suburban environments. For each class, the first subfigure (A) refers to the urban environment and the second subfigure (B) refers to the suburban environment. The visualizations include yearly overlays, weekly box plots, and monthly averages, as described in Section 3.2.3.

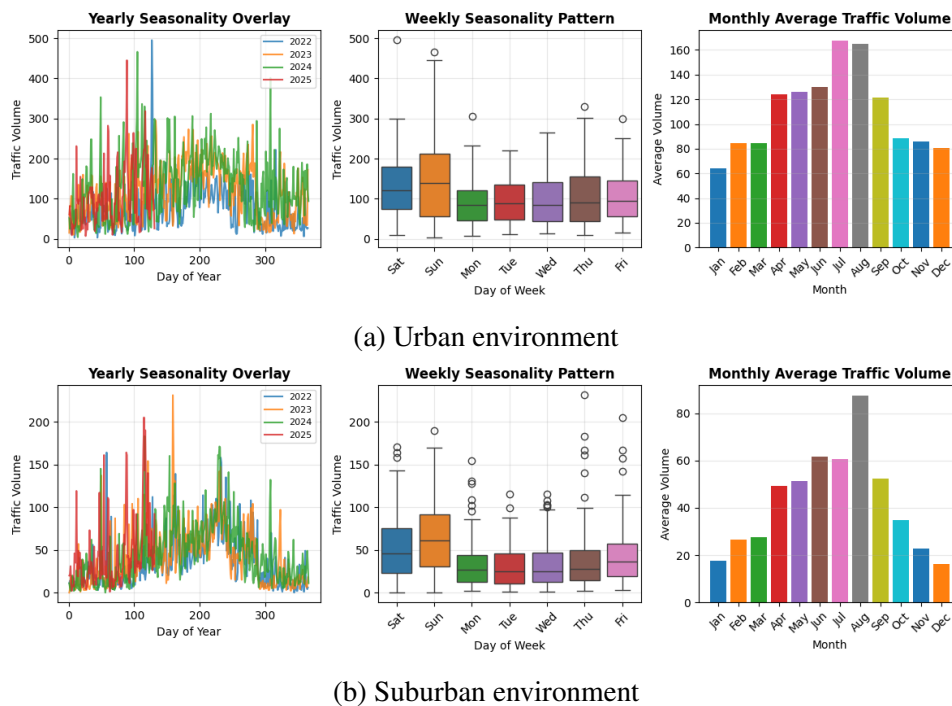


Figure B.1: Pattern analysis for class 1 vehicles in the (A) urban and (B) suburban environments.

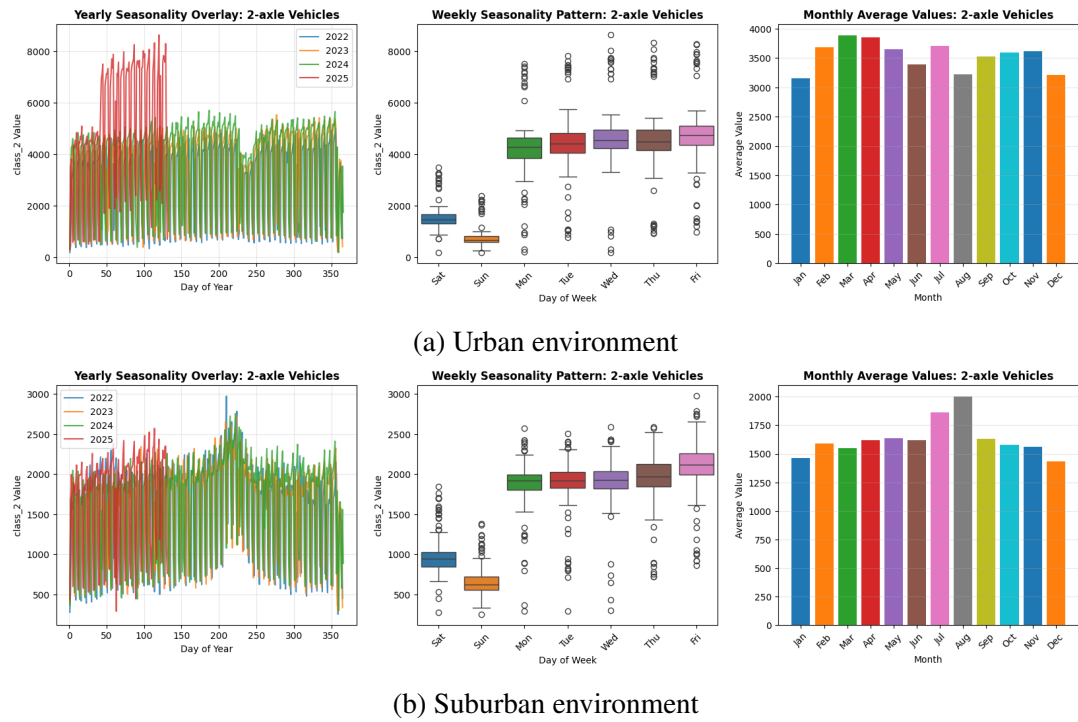


Figure B.2: Pattern analysis for class 2 vehicles in the (A) urban and (B) suburban environments.

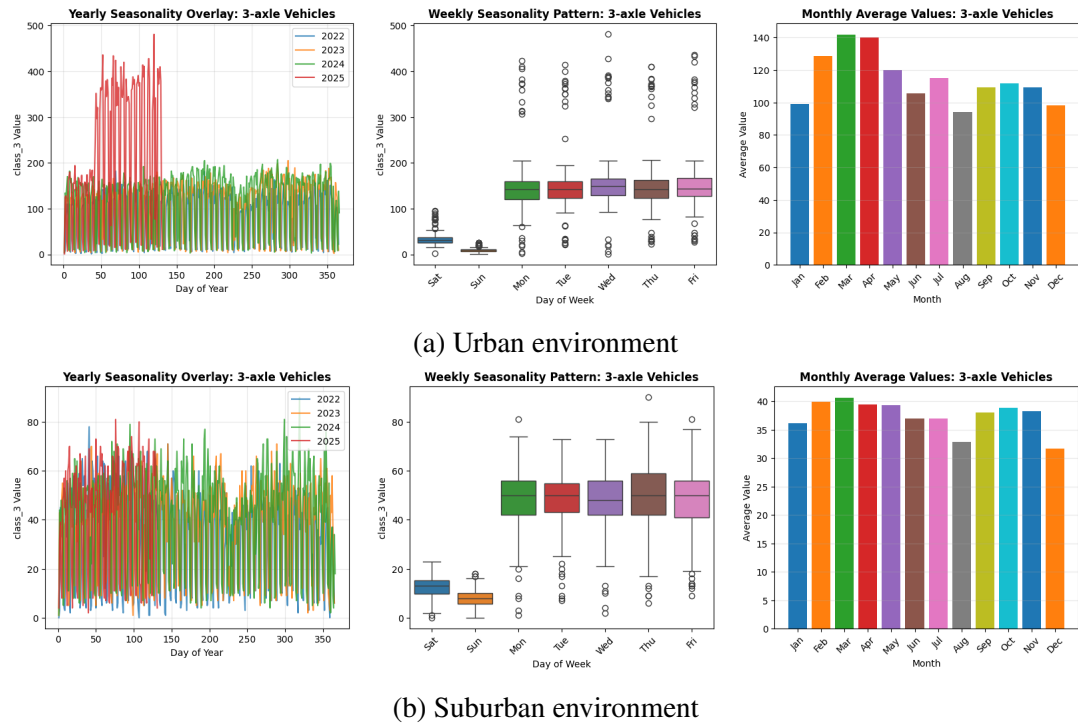
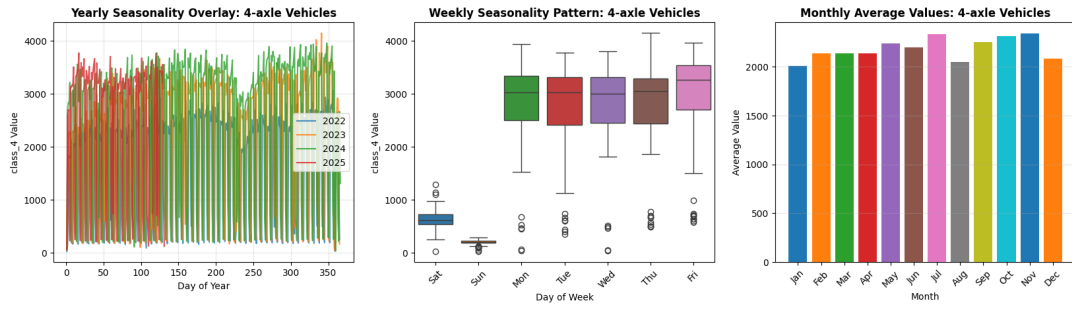
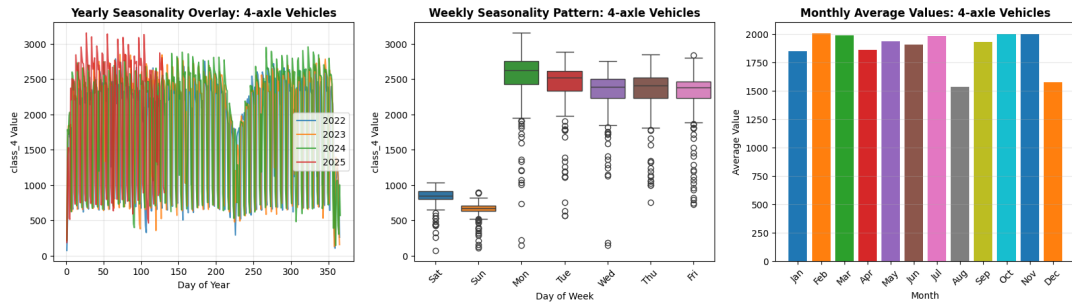


Figure B.3: Pattern analysis for class 3 vehicles in the (A) urban and (B) suburban environments.

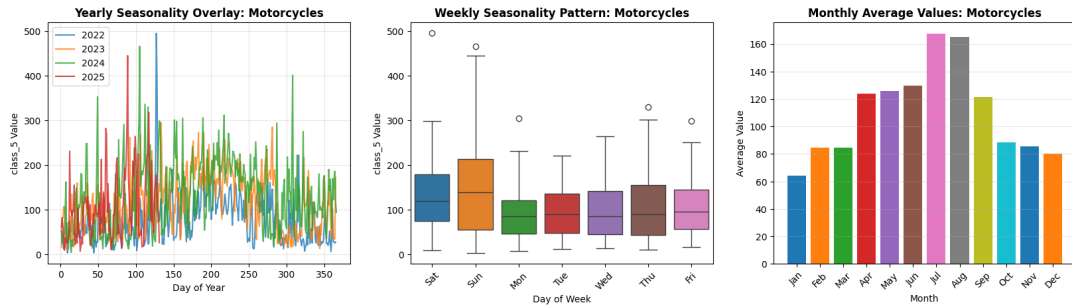


(a) Urban environment

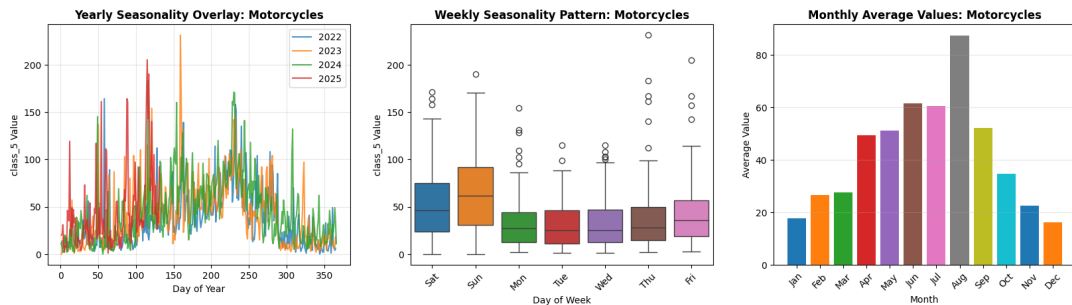


(b) Suburban environment

Figure B.4: Pattern analysis for class 4 vehicles in the (A) urban and (B) suburban environments.



(a) Urban environment



(b) Suburban environment

Figure B.5: Pattern analysis for class 5 vehicles in the (A) urban and (B) suburban environments.

References

- Ascendi. Business areas. Available at <https://www.ascendi.pt/areas-de-negocio/>, 2025a. Retrieved February 11, 2025.
- Ascendi. Presentation. Available at <https://www.ascendi.pt/apresentacao/>, 2025b. Retrieved February 02, 2025.
- Ascendi. Our roads. Available at <https://www.ascendi.pt/as-nossas-vias/>, 2025c. Retrieved February 11, 2025.
- Marie-Aude Aufaure and Esteban Zimányi, editors. *Business Intelligence: Second European Summer School, eBISS 2012, Brussels, Belgium, July 15–21, 2012, Tutorial Lectures*, volume 138 of *Lecture Notes in Business Information Processing*. Springer, Berlin, Heidelberg, 2013. ISBN 978-3-642-36317-7. doi: 10.1007/978-3-642-36318-4.
- Annarita Carianni and Andrea Gemma. Overview of traffic flow forecasting techniques. *IEEE Open Journal of Intelligent Transportation Systems*, 2025. doi: 10.1109/OJITS.2025.3580802. Received 7 Dec. 2024; revised 14 Apr. 2025 and 5 Jun. 2025; accepted 13 Jun. 2025; published 18 Jun. 2025; current version 11 Jul. 2025.
- Peter Chapman. Crisp-dm. In *CRISP-DM 1.0: Step-by-step data mining guide*, 2000. URL <https://api.semanticscholar.org/CorpusID:59777418>.
- Erlangung. *Quantifying and Interpreting Uncertainty in Time Series Forecasting*. PhD thesis, Karlsruhe Institute of Technology, 2023.
- R. Fu, Z. Zhang, and L. Li. Using lstm and gru neural network methods for traffic flow prediction. In *Proceedings of the 31st Youth Academic Annual Conference of Chinese Association of Automation (YAC)*, pages 324–328, November 2016.
- Tilman Gneiting. Editorial: Probabilistic forecasting. *Journal of the Royal Statistical Society Series A: Statistics in Society*, 171(2):319–321, 03 2008. ISSN 0964-1998. doi: 10.1111/j.1467-985X.2007.00522.x. URL <https://doi.org/10.1111/j.1467-985X.2007.00522.x>.
- J. Guo, W. Huang, and Billy M. Williams. Adaptive kalman filter approach for stochastic short-term traffic flow rate prediction and uncertainty quantification. *Transportation Research Part C: Emerging Technologies*, 43:50–64, June 2014.

- M. M. Hamed, H. R. Al-Masaeid, and Z. M. B. Said. Short-term prediction of traffic volume in urban arterials. *Journal of Transportation Engineering*, 121(3):249–254, May 1995.
- Jiamu He. Research on traffic flow prediction based on arima model. In *Proceedings of the 1st International Conference on Engineering Management, Information Technology and Intelligence (EMITI 2024)*, pages 16–21, 2024. doi: 10.5220/0012887800004508.
- S. L. Ho, M. Xie, L. C. Tang, K. Xu, and T. N. Goh. Neural network modeling with confidence bounds: A case study on the solder paste deposition process. *IEEE Transactions on Electronics Packaging Manufacturing*, 24(4):323–332, October 2001.
- Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural Computation*, 9(8):1735–1780, 1997. doi: 10.1162/neco.1997.9.8.1735. URL <https://pubmed.ncbi.nlm.nih.gov/9377276/>.
- R. J. Hyndman and G. Athanasopoulos. *Forecasting: Principles and Practice*. OTexts, 3rd edition, 2021. Available at <https://otexts.com/fpp3/single-level.html>.
- Rob J. Hyndman and Anne B. Koehler. Another look at measures of forecast accuracy. *International Journal of Forecasting*, 22(4):679–688, 2006. doi: 10.1016/j.ijforecast.2006.03.001.
- Guolin Ke, Qi Meng, Thomas Finley, Taifeng Wang, Wei Chen, Weidong Ma, Qiwei Ye, and Tie-Yan Liu. Lightgbm: A highly efficient gradient boosting decision tree. arXiv:1712.01043, 2017a. URL <https://arxiv.org/abs/1712.01043>.
- Guolin Ke, Qi Meng, Thomas Finley, Taifeng Wang, Wei Chen, Weidong Ma, Qiwei Ye, and Tie-Yan Liu. Lightgbm: A highly efficient gradient boosting decision tree. In *Advances in Neural Information Processing Systems (NeurIPS)*, pages 3146–3154, 2017b. URL <https://proceedings.neurips.cc/paper/6907-lightgbm-a-highly-efficient-gradient-boosting-decision-tree.pdf>.
- Abbas Khosravi, Saeid Nahavandi, Douglas Creighton, and Amir F. Atiya. Comprehensive review of neural network-based prediction intervals and new advances. *IEEE Transactions on Neural Networks*, 22(9):1341–1356, September 2011. doi: 10.1109/TNN.2011.2162110.
- Ibai Laña, Javier Del Ser, Manuel Vélez, and Eleni I. Vlahogianni. Road traffic forecasting: Recent advances and new challenges. *IEEE Intelligent Transportation Systems Magazine*, 10(2):93–109, 2018. doi: 10.1109/MITS.2018.2806634.
- Ibai Laña, Ignacio Olabarrieta, and Javier Del Ser. Measuring the confidence of traffic forecasting models: Techniques, experimental comparison and guidelines towards their actionability. arXiv preprint arXiv:2210.16049, 2022. URL <https://arxiv.org/abs/2210.16049>. Version 1, 46 pages.

- Xu Li, Yutong Xia, Yuxuan Liang, Junfeng Hu, Yiwei Wang, Lei Bai, Chao Huang, Zhenguang Liu, Bryan Hooi, and Roger Zimmermann. Largest: A large-scale, long-term traffic forecasting dataset with rich metadata. In *37th Conference on Neural Information Processing Systems (NeurIPS 2023) Track on Datasets and Benchmarks*, 2023.
- H. Lin, R. Zito, and M. Taylor. A review of travel-time prediction in transport and logistics. In *Proceedings of the Eastern Asia Society for Transportation Studies*, volume 5, pages 1433–144, Bangkok, Thailand, 2005.
- Pawan Lingras, Satish C. Sharma, Phil Osborne, and Iftekhar Kalyar. Traffic volume time-series analysis according to the type of road use. *Computer-Aided Civil and Infrastructure Engineering*, 15(4):365–373, 2000.
- Longshun Liu. A short-term traffic flow prediction method based on svr. In *2021 2nd International Conference on Urban Engineering and Management Science (ICUEMS)*, 2021.
- Y. Liu, H. Zheng, X. Feng, and Z. Chen. Short-term traffic flow prediction with conv-lstm. In *Proceedings of the 9th International Conference on Wireless Communications and Signal Processing (WCSP)*, pages 1–6, October 2017.
- Marcin Majka. Arimax: Time series forecasting with external variables. 09 2024.
- Spyros Makridakis, Evangelos Spiliotis, and Vassilios Assimakopoulos. M5 accuracy competition: Results, findings, and conclusions. *International Journal of Forecasting*, 38(4):1346–1364, 2022. doi: 10.1016/j.ijforecast.2021.10.009.
- L. Manibardo, I. Laña, and J. D. Ser. Deep learning for road traffic forecasting: Does it make a difference? *IEEE Transactions on Intelligent Transportation Systems*, 23(7): 6164–6188, 2021. doi: 10.1109/TITS.2021.3059097.
- Kentaro Mori and Kara M. Kockelman. Day-of-week, month, and seasonal demand variations: Comparing flow estimates across new travel data sources. *Findings*, July 2024. doi: 10.32866/001c.118815. URL <https://doi.org/10.32866/001c.118815>.
- Petter Næss and Arvid Strand. Traffic forecasting at ‘strategic’, ‘tactical’ and ‘operational’ level. *Journal of Critical Realism*, 51(2):41–48, 2015.
- Evgeny Nikulchev and Alexander Chervyakov. Prediction intervals: A geometric view. *Symmetry*, 15:781, 03 2023. doi: 10.3390/sym15040781.
- Nixtla. Autoarima model — nixtla statsforecast. <https://nixtlaverse.nixtla.io/statsforecast/docs/models/autoarima.html>, 2025. Accessed: YYYY-MM-DD.
- Isao Okutani and Yorgos J. Stephanedes. Dynamic prediction of traffic volume through kalman filtering theory. *Transportation Research Part B: Methodological*, 18(1):1–11, February 1984.

Alfredo M. Pereira and Jorge M. Andraz. On the economic and fiscal effects of investments in road infrastructures in Portugal. *International Economic Journal*, 2011.

Renascença. Porto. trucks exempt from tolls at crep from 2026. Available at <https://rr.pt/noticia/pais/2025/04/03/porto-camioes-isentos-de-portagens-na-crep-a-partir-de-2026/420207/>, 2025.

República Portuguesa. Decreto-Lei n.º 39/2005, de 17 de fevereiro de 2005. Diário da República n.º 34/2005, Série I-A. Available at: [urlhttps://diariodarepublica.pt/dr/detalhe/decreto-lei/39-2005-605324](https://diariodarepublica.pt/dr/detalhe/decreto-lei/39-2005-605324), 2005. Introduces exceptions allowing certain vehicles with specific features to be considered Class 1.

República Portuguesa. Decreto-Lei n.º 71/2018, de 5 de setembro de 2018. Diário da República n.º 172/2018, Série I. Available at: [urlhttps://dre.pt/dre/detalhe/decreto-lei/71-2018-116160172](https://dre.pt/dre/detalhe/decreto-lei/71-2018-116160172), 2018. Further adjusts Class 1 and Class 2 criteria, including EURO 6 compliance.

J. Salotti, S. Fenet, R. Billot, N.-E. El Faouzi, and C. Solnon. Comparison of traffic forecasting methods in urban and suburban context. In *2018 IEEE 30th International Conference on Tools with Artificial Intelligence (ICTAI)*, pages 846–853, 2018. doi: 10.1109/ICTAI.2018.00132.

M. Shaygan, C. Meese, W. Li, X. G. Zhao, and M. Nejad. Traffic prediction using artificial intelligence: Review of recent advances and emerging opportunities. *Transportation Research Part C: Emerging Technologies*, 145:103921, December 2022a.

M. Shaygan et al. Transportation research part c: Emerging technologies. *Transportation Research Part C*, 145:103921, 2022b.

European Union. Regulation (EC) No 459/2012 of the European Parliament and of the Council of 20 June 2007 on type approval of motor vehicles with respect to emissions from light passenger and commercial vehicles (Euro 5 and Euro 6). Official Journal of the European Union, L 171, 29.6.2007, pp. 1–16. Available at: [urlhttps://eur-lex.europa.eu/eli/reg/2007/715/oj](https://eur-lex.europa.eu/eli/reg/2007/715/oj), 2007.

C. P. I. J. van Hinsbergen, J. W. C. van Lint, and F. M. Sanders. Short term traffic prediction models. In *Proceedings of the 14th World Congress on Intelligent Transport Systems: ITS for a Better Life*, pages 1–18, Beijing, China, 2007. Research Institute of Highway, Chinese Ministry of Communications.

Via Verde Portugal. Via Verde for Motorcycles: Class 5. Available at: [urlhttps://www.viaverde.pt/particulares/apoio-ao-cliente/perguntas-frequentes/perguntas-frequentes?themeId=224&faqId=111](https://www.viaverde.pt/particulares/apoio-ao-cliente/perguntas-frequentes/perguntas-frequentes?themeId=224&faqId=111), 2024. Describes the toll discount and classification for motorcycles using Via Verde.

Eleni I. Vlahogianni, John C. Golias, and Matthew G. Karlaftis. Short-term traffic forecasting: Overview of objectives and methods. *Transport Reviews*, 24(5):533–557, 2004. doi: 10.1080/0144164042000195072.

- J. Wang, J. Wu, Z. Wang, F. Gao, and Z. Xiong. Understanding urban dynamics via context-aware tensor factorization with neighboring regularization. *IEEE Transactions on Knowledge and Data Engineering*, 32(11):2269–2283, 2019.
- S. L. Wickramasuriya, G. Athanasopoulos, and R. J. Hyndman. Optimal forecast reconciliation for hierarchical and grouped time series through trace minimization. *Journal of the American Statistical Association*, 114(526):804–819, 2018. doi: 10.1080/01621459.2018.1448825.
- Billy M. Williams. Multivariate vehicular traffic flow prediction: Evaluation of arimax modeling. *Transportation Research Record*, 1776(1):194–200, January 2001.
- Billy M. Williams and Lester A. Hoel. Modeling and forecasting vehicular traffic flow as a seasonal arima process: Theoretical basis and empirical results. *Journal of Transportation Engineering*, 129(6):664–672, November 2003.
- Billy M. Williams, Praveen K. Durvasula, and David E. Brown. Urban freeway traffic flow prediction: Application of seasonal autoregressive integrated moving average and exponential smoothing models. *Transportation Research Record*, 1644(1):132–141, January 1998.
- Xiaotong Yao, Xiaoli Fu, and Chaofei Zong. Short-term load forecasting method based on feature preference strategy and lightgbm-xgboost. *IEEE Access*, 10:1–1, 01 2022. doi: 10.1109/ACCESS.2022.3192011.
- X. Yin, G. Wu, J. Wei, and H. Qi. Deep learning on traffic prediction: Methods, analysis, and future directions. *IEEE Transactions on Intelligent Transportation Systems*, 22(12): 4928–4943, 2021. doi: 10.1109/TITS.2021.3055863.
- Jun Hua Zhao, Zhao Yang Dong, Zhao Xu, and Kit Po Wong. A statistical approach for interval forecasting of the electricity price. *IEEE Transactions on Power Systems*, 23(2):267–276, May 2008. doi: 10.1109/TPWRS.2008.919309.
- H. Zheng, F. Lin, X. Feng, and Y. Chen. A hybrid deep learning model with attention-based conv-lstm networks for short-term traffic flow prediction. *IEEE Transactions on Intelligent Transportation Systems*, 22(11):6910–6920, November 2021.