FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO

A Survival-Based Health Index Framework for Predictive Maintenance of Bike-Sharing Systems

Gabriel Eggert de Freitas



Mestrado em Engenharia Eletrotécnica e de Computadores

Supervisor: Prof. Dr. Daniel G. Costa

September 23, 2026

© Gabriel Eggert de Freitas, 2026

A Survival-Based Health Index Framework for Predictive Maintenance of Bike-Sharing Systems

Gabriel Eggert de Freitas

Mestrado em Engenharia Eletrotécnica e de Computadores

September 23, 2026

Resumo

Os sistemas de partilha de bicicletas afirmaram-se como componentes essenciais da mobilidade urbana sustentável nas cidades inteligentes. Contudo, a gestão operacional destas frotas enfrenta desafios significativos decorrentes do desgaste mecânico contínuo dos ativos, onde os paradigmas tradicionais de manutenção reativa e preventiva por intervalos fixos resultam em indisponibilidade do serviço, riscos de segurança e custos operacionais elevados. Esta limitação é agravada por uma assimetria estrutural na disponibilidade de dados abertos, em que os padrões públicos cobrem a localização e disponibilidade dos veículos em tempo real, mas omitem o histórico detalhado de reparações efetuadas nas oficinas.

Para responder a esta lacuna, esta dissertação desenvolve um enquadramento de manutenção preditiva baseado em Análise de Sobrevivência, capaz de transformar registos operacionais de viagens e de exposição meteorológica em modelos probabilísticos de degradação ao nível do componente. A metodologia utiliza dados longitudinais do sistema MetroBikeShare de Los Angeles, integrando a harmonização semântica de logs de oficina, o cálculo de rotas e a fusão de dados ambientais. O pipeline constrói unidades de Operação de Manutenção sujeitas a filtragem de censura à esquerda e auditoria de data leakage, treinando arquiteturas semiparamétricas CPH e não-lineares GBSA para seis subsistemas mecânicos primários.

As funções de sobrevivência condicionada resultantes são mapeadas num Índice de Saúde contínuo e multinível, que agrega a integridade física desde o componente individual até à totalidade da frota através de ponderações baseadas no tempo de vida característico de cada montagem. O sistema discretiza as trajetórias de degradação em quatro bandas operacionais definidas por percentis empíricos de falha e culmina num protótipo de painel de monitorização interativo. Esta interface web disponibiliza vistas de linha temporal por veículo e matrizes de estado para a frota, permitindo identificar precocemente ativos em risco e confirmar visualmente as recuperações de saúde após intervenções de oficina.

Os resultados empíricos demonstram que os modelos condicionados em covariáveis atingem elevada capacidade discriminativa, com valores de C-IPCW entre 0.836 e 0.939, e excelente calibração de probabilidades, registando pontuações de IBS entre 0.022 e 0.101, superando decisivamente os baselines populacionais não-condicionados. O estudo quantifica ainda o impacto da tecnologia de propulsão, comprovando que as bicicletas elétricas sofrem um ciclo de trabalho significativamente acelerado. Em conjunto, o enquadramento metodológico, o Índice de Saúde e o painel de monitorização comprovam a viabilidade prática de transitar de rotinas de reparação corretiva para uma gestão proativa e informada pelo risco.

Palavras-chave: Sistemas de partilha de bicicletas • Manutenção preditiva • Análise de sobrevivência • Índice de saúde • Apoio à decisão

Abstract

Bike-sharing systems have become indispensable components of sustainable urban mobility within smart cities. However, the operational management of these fleets faces persistent challenges due to continuous mechanical asset degradation, where conventional reactive and fixed-interval preventive maintenance regimes lead to service downtime, safety risks, and unnecessary operational costs. This inefficiency is compounded by a structural data accessibility gap in the public domain, where open specifications expose real-time vehicle location and availability while keeping granular workshop repair logs strictly proprietary.

To address this limitation, this dissertation presents a predictive maintenance framework based on Survival Analysis, designed to convert routine trip records and meteorological exposure data into probabilistic component-level degradation models. Using longitudinal operational logs from the MetroBikeShare fleet in Los Angeles, the pipeline integrates semantic harmonization of workshop records, route geometry enrichment, and meteorological data fusion. The architecture constructs Maintenance Operation units under strict left-censoring filtration and data leakage auditing, fitting semi-parametric CPH and non-linear GBSA models across six primary mechanical assemblies.

The predicted conditional survival functions are mapped into a continuous, multi-level Health Index that aggregates physical integrity from individual components to composite vehicle and fleet-wide health scores, weighted by subsystem characteristic lifetimes. Continuous health trajectories are discretized into four operational condition bands using empirical failure percentiles and rendered within a web-based interactive monitoring dashboard. Featuring individual asset timelines with step-function recovery resets alongside fleet-wide status matrices, the interface provides operations teams with transparent decision-support capabilities.

Empirical validation demonstrates that covariate-conditioned survival models achieve high predictive discrimination, recording C-IPCW scores between 0.836 and 0.939, and superior probability calibration, yielding IBS values between 0.022 and 0.101, decisively outperforming unconditioned population baselines. Furthermore, the analysis isolates propulsion technology as a primary degradation driver, quantifying the accelerated duty cycle of electric bicycles. Together, the survival modeling pipeline, multi-level Health Index, and interactive dashboard validate the practical feasibility of transitioning shared micro-mobility operations from reactive repairs toward proactive, condition-based fleet management.

Keywords: Bike-sharing systems • Predictive maintenance • Survival analysis • Health Index • Decision support

The United Nations Sustainable Development Goals

The United Nations Sustainable Development Goals (SDGs) provide a global framework to achieve a better and more sustainable future for all. It includes 17 goals to address the world's most pressing challenges, including poverty, inequality, climate change, environmental degradation, peace, and justice.

Implementing predictive maintenance (PdM) in bike-sharing systems directly supports urban sustainability by enhancing the reliability and safety of green transportation. By maximizing the remaining useful life of components and reducing the frequency of corrective repairs, this research contributes to a circular economy and lower carbon footprints for municipal infrastructure.

The specific Sustainable Development Goals addressed by this research have the following names:

SDG 9 Build resilient infrastructure, promote inclusive and sustainable industrialization and foster innovation

SDG 11 Make cities and human settlements inclusive, safe, resilient and sustainable

SDG 12 Ensure sustainable consumption and production patterns

SDG	Target	Contribution	Performance Indicators and Metrics
9	9.4	Retrofitting shared bicycle fleets with data-driven predictive maintenance technologies improves resource-use efficiency and safety.	Reduction in component waste and increase in asset life-cycle.
11	11.2	Predictive maintenance ensures a more reliable and available bike fleet, providing a safe and sustainable transportation alternative.	Mean Time Between Failures (MTBF) and fleet availability percentage.
12	12.5	Transitioning from preventive to predictive maintenance prevents the premature disposal of functional parts, reducing environmental waste.	Percentage reduction in "over-maintenance" replacements.

Acknowledgements

First and foremost, I want to thank my parents, for giving me the necessary opportunities to go after my dream, and always being there for me, in my worst and best moments. I love you.

To my childhood friends, specially every member of Macaco Elástico, everyone is a brother and sister to me. You live in my heart, thank you so much to be part of my life.

To my close friends in Portugal, including Bernardo Espíndola, Enzo Magno, Enzo Sekito, Estevão Brasil, Igor Liberato, Levy Mitsugui, Paulo Silva, Pedro Rezende, Vinicius Caparróz, and Vinicius Pirassoli, thank you for always helping me, either by hearing my struggles or just making me company. Your support was essential to survive the last 7 years.

I also extend my thanks to my colleagues at the IEEE UP Student Branch. As everyone knows, the branch has a special space in my heart, and it's half the reason I'm the man I am today, both personally and professionally.

I thank as well my work colleagues, Rafael Silva, Analcísio Rodino, Carolina Salvador, José Alpuim, José Alves, and Miguel Girão. Working and taking the Master's Degree at the same time wouldn't be possible without your support and understanding.

To my supervisor, Professor Daniel G. Costa, thank you so much for your guidance, availability, feedback, and support throughout this dissertation, it was essential to the development and completion of my work.

Finally, I thank Jordi Grau-Escolano and Thommas K. S. Flores for your valuable support and technical exchanges regarding survival analysis methodologies. Your insights and advice were very helpful in this thesis.

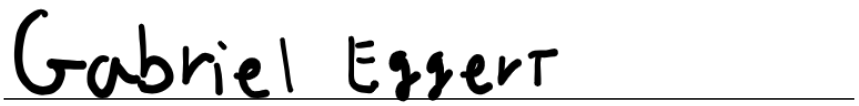
Gabriel Eggert de Freitas

Declaration of honour

I declare that this dissertation is my own work and has not previously been submitted or assessed at this or any other institution. References to other authors (statements, ideas, thoughts), as well as the use of published works, strictly comply with the rules of attribution and are duly identified in the text and in the bibliographic references, in accordance with the applicable referencing standards. I am aware that the practice of plagiarism or self-plagiarism constitutes an academic offence, as established in the U.Porto Code of Ethics.

I also declare that I have used Artificial Intelligence tools to complete part(s) of this dissertation, and that this use and its results are duly noted and have been carefully checked for errors, inaccuracies, unfounded attributions or other forms of bias in content and arguments, as well as determining authorship.

Gabriel Eggert de Freitas



"A lesson without pain is meaningless. That's because no one can gain without sacrificing something. But by enduring that pain and overcoming it, he shall obtain a powerful, unmatched heart. A fullmetal heart."

Edward Elric

Contents

1 Introduction **1**
- 1.1 Problem Definition . . . 2
- 1.2 Objectives and Research Question . . . 3
- 1.3 Contributions . . . 4
- 1.4 Document Structure . . . 4

2 Background and State of the Art **6**
- 2.1 Bike-Sharing Systems and Smart Cycling . . . 6
  - 2.1.1 Soft Mobility and Smart Cycling . . . 6
  - 2.1.2 Anatomy of a Bike-Sharing System . . . 7
  - 2.1.3 The Shared Bicycle as a Maintenance Asset . . . 8
- 2.2 Maintenance Paradigms in Bike-Sharing Systems . . . 9
  - 2.2.1 Fundamentals and Data Requirements . . . 9
  - 2.2.2 Predictive Maintenance in Industrial and Vehicle Fleets . . . 10
  - 2.2.3 Predictive Maintenance in Bike-Sharing Systems . . . 11
- 2.3 Fundamentals of Survival Analysis . . . 12
  - 2.3.1 Time-to-Event Data and Censoring . . . 12
  - 2.3.2 Survival and Hazard Functions . . . 13
  - 2.3.3 Covariate-Conditioned Models . . . 14
  - 2.3.4 Evaluation Metrics for Censored Data . . . 15
- 2.4 Health Indices for Fleet Monitoring . . . 15
- 2.5 Identified Research Gaps . . . 16

3 Methodology **18**
- 3.1 Research Question and Hypotheses . . . 18
- 3.2 Scope of the Study . . . 19
- 3.3 Problem Formulation . . . 20
  - 3.3.1 Maintenance Operation Units and Event Definition . . . 20
  - 3.3.2 Covariates and Feature Aggregation . . . 21
  - 3.3.3 Models and Health Index Mapping . . . 22
- 3.4 Evaluation Metrics and Hypothesis Verification . . . 23
- 3.5 Methodological Pipeline . . . 24
- 3.6 Validation Strategy . . . 25

4 Implementation **27**
- 4.1 Tools and Technologies . . . 27
- 4.2 Data Acquisition and Fleet Selection . . . 29
- 4.3 Data Processing and Harmonization . . . 33

4.3.1 Trip Cleaning . . . 33
4.3.2 Maintenance Taxonomy and Harmonization . . . 34
4.3.3 Route Geometry Enrichment . . . 35
4.3.4 Weather Enrichment . . . 37
4.3.5 Feature Screening . . . 38
4.3.6 Daily Dataset Assembly . . . 42
4.4 MO Unit Construction and Feature Audit . . . 43
4.4.1 Builder and Resulting Structure . . . 43
4.4.2 Feature Configurations and Leakage Audit . . . 44
4.5 Survival Model Training and Evaluation . . . 45
4.6 Health Index Computation . . . 45
4.7 Monitoring Dashboard Generation . . . 46

5 Results and Discussion **48**
5.1 Fleet Profile and Operational Overview . . . 48
5.1.1 Fleet Utilization and Temporal Dynamics . . . 48
5.1.2 Maintenance Log Taxonomy and Workshop Workload . . . 49
5.1.3 Censoring Structure and Usage Profiles . . . 51
5.2 Data Pipeline Quality . . . 54
5.3 Survival Model Performance and Hypothesis Verification . . . 54
5.3.1 Population-Level Behavior: Kaplan-Meier and Weibull . . . 54
5.3.2 Covariate-Conditioned Models Against the Baselines . . . 57
5.4 Health Index Behavior . . . 62
5.5 Monitoring Dashboard and Decision Support . . . 65
5.5.1 Individual Asset Health Evaluation . . . 65
5.5.2 Fleet-Wide Health Snapshot and Categorization . . . 66
5.6 Synthesis and Limitations . . . 68

6 Conclusion **70**
6.1 Summary and Contributions . . . 70
6.2 Hypothesis Verification . . . 71
6.3 Future Work . . . 72

References **74**

A Drop-Column Feature Sensitivity Analysis **81**

B Component-Level Model Interpretability **88**

C Time-Varying Discrimination and Calibration Diagnostics **91**

List of Figures

1.1 Common failures of bicycle parts and components 2

2.1 Dockless bike-sharing infrastructure example 8
2.2 The bathtub curve 10
2.3 Observational trajectories and censoring types 13
2.4 Weibull hazard and survival functions under different shape parameters 14

3.1 Overview of the proposed methodological workflow 24

4.1 Spatial distribution of the seven participating operators 30
4.2 MetroBikeShare station infrastructure and components 32
4.3 Origin-destination route geometries across Los Angeles 37
4.4 Spearman correlation matrix of daily bicycle features 39
4.5 Spearman correlation matrix of weather variables 40
4.6 Meteorological variable screening criteria 41

5.1 Daily trip counts for the MetroBikeShare fleet 49
5.2 Proportion of maintenance operations by primary component 50
5.3 Distribution of maintenance days per bicycle 50
5.4 Daily count of bicycles undergoing workshop servicing 51
5.5 Uncensored and right-censored MO units per subsystem 52
5.6 Cumulative kilometers per MO unit by component 53
5.7 Kaplan-Meier survival curves with fitted Weibull overlays 55
5.8 Kaplan-Meier survival curves stratified by bicycle type 56
5.9 Held-out discrimination performance per component 59
5.10 Predicted versus actual time-to-event per component 60
5.11 Integrated Brier Score per component 61
5.12 Longitudinal Health Index trajectories for individual bicycles 63
5.13 Empirical Health Index distributions at corrective maintenance events 64
5.14 Individual asset health trajectory (CPH view) 65
5.15 Individual asset health trajectory (GBSA view) 66
5.16 Fleet Grid view (Ungrouped spatial matrix) 67
5.17 Fleet Grid view (Grouped by health status) 67

A.1 Drop-column feature importance: Brakes 82
A.2 Drop-column feature importance: Cables/Hydraulics 83
A.3 Drop-column feature importance: Chain 84
A.4 Drop-column feature importance: Saddle 85
A.5 Drop-column feature importance: Tires 86

A.6 Drop-column feature importance: Wheels 87

B.1 CPH hazard ratios per component 89
B.2 GBSA feature importance per component 90

C.1 Brier score over time per component 92
C.2 Time-dependent AUC per component 93

List of Tables

2.1 Comparison of related literature and the proposed framework 17

4.1 Main Python libraries and software tools 28
4.2 Overview of operators providing usable maintenance data 30
4.3 Trip-level outlier removal rules 33
4.4 Component taxonomy for maintenance harmonization 34
4.5 Action taxonomy and MO unit classification 35
4.6 Structural schema of an MO unit record 44

5.1 MO unit population and censoring rates per component 52
5.2 Population-level Weibull parameters per component 55
5.3 Held-out predictive performance per component 57
5.4 Empirical Health Index condition band boundaries per model 64

List of Acronyms

API	Application Programming Interface
AUC	Area Under the ROC Curve
BSS	Bike-Sharing System
CORAL	Correlation Alignment
CPH	Cox Proportional Hazards
CV	Coefficient of Variation
C-IPCW	Concordance Index under Inverse Probability of Censoring Weighting
GBFS	General Bikeshare Feed Specification
GBSA	Gradient Boosting Survival Analysis
GPS	Global Positioning System
HeiGIT	Heidelberg Institute for Geoinformation Technology
HI	Health Index
IBS	Integrated Brier Score
IoT	Internet of Things
IPCW	Inverse Probability of Censoring Weighting
MAPE	Mean Absolute Percentage Error
MLE	Maximum Likelihood Estimation
MO	Maintenance Operation
PBF	Protocolbuffer Binary Format
PdM	Predictive Maintenance
PHM	Prognostics and Health Management
RMSE	Root Mean Square Error
RUL	Remaining Useful Life
SA	Survival Analysis
SoH	State of Health
TPE	Tree-structured Parzen Estimator
VHI	Vehicle Health Index

Chapter 1

Introduction

The integration of digital innovation and sustainability objectives in urban planning has established the smart city as an operational framework for municipal infrastructure [1]. Within this ecosystem, Bike-Sharing System (BSS) platforms constitute an established component of multimodal public transit [2], [3]. By offering a flexible and low-carbon alternative for short-distance travel, shared micro-mobility contributes to reducing vehicular emissions and traffic congestion while supporting active transport [4]. As municipalities incorporate shared bicycles into broader decarbonization strategies, fleet deployments have spread across increasingly diverse geographical, topographic, and climatic conditions.

This operational expansion introduces significant management challenges centered on physical asset preservation. Historically, maintenance in shared mobility was treated as a secondary logistical task, subordinate to rapid fleet deployment [2]. However, the long-term viability of these networks depends directly on the mechanical integrity of the vehicles, as out-of-service or defective bicycles degrade service reliability, reduce user trust, and elevate operational overheads.

Maintenance regimes in BSS still rely predominantly on two traditional strategies. Reactive maintenance follows protocols in which repairs occur only after a component or system has already broken down. This approach is inherently inefficient, as it implies that the asset is already out of service, resulting in high downtime, user dissatisfaction, and potential safety risks before the repair is completed. Alternatively, preventive maintenance involves servicing bicycles at predetermined intervals, based either on time (e.g., every six months) or on usage (e.g., every 1,000 km). While this reduces the probability of catastrophic failures, it suffers from the "mean-time fallacy" [5], [6], by assuming that mechanical degradation is linear and uniform across the entire fleet.

In reality, bicycle wear is highly context-dependent: identical bicycles degrade at different rates depending on their "stress histories" [7], [8], such as exposure to road salt, coastal salinity, or steep urban gradients. Figure 1.1 presents a visual schema of typical failures in regular bicycles. Rigid maintenance intervals applied to such heterogeneous conditions lead either to over-maintenance or to under-maintenance, and both outcomes waste financial and material resources, being inadequate for the dynamic and stochastic usage patterns of modern smart cities [8].



Figure 1.1: Common failures of bicycle parts and components. Figure generated with the support of Gemini and manually reviewed by the author.

To ensure the safety, reliability, and financial viability of urban cycling networks, this dissertation pursues a shift toward Predictive Maintenance (PdM). By combining historical maintenance logs, trip records, and environmental data, PdM estimates the failure risk of each asset from its actual usage and exposure, allowing operators to intervene only when necessary, maximizing the Remaining Useful Life (RUL) of components and preventing safety hazards [9]. Since fleet maintenance records are, in essence, time-to-event data in which many components have not yet failed by the end of the observation period, this dissertation adopts Survival Analysis (SA) as its central modeling framework, using it to estimate component-level failure risk over time and to translate those estimates into practical tools for fleet monitoring.

1.1 Problem Definition

The fundamental obstacle to implementing Predictive Maintenance in the bike-sharing sector is the scarcity of high-quality, granular maintenance data in the public domain. While the growth of urban

mobility platforms has generated terabytes of usage statistics, internal maintenance records remain proprietary and strictly guarded by operators. Publicly available standards, such as the General Bikeshare Feed Specification (GBFS) [10], reinforce this asymmetry by covering asset availability but not mechanical condition.

Despite the vast availability of trip records, a pronounced imbalance persists in publicly accessible fleet data. Although operational usage logs are abundant, detailed repair records linking specific mechanical failure modes, such as chain elongation, brake wear, or frame fatigue, to asset usage histories are virtually absent from open repositories. A systematic audit of public repositories conducted during the preliminary phase of this research and published in [11] identified only two public datasets containing maintenance information [12], [13]. Both datasets, however, are either limited in scale or lack the ground-truth failure labels necessary to train predictive models. Existing studies on bicycle fleet reliability depend almost exclusively on proprietary, single-operator partnerships [8], [14], which restricts the generalizability of their empirical findings.

This scarcity is compounded by a gap in the literature itself, which has historically prioritized the rebalancing problem (the logistical redistribution of assets) [15], [16] and demand forecasting [17] over mechanical reliability. As fleets age and usage intensity escalates, this neglect leads to high operational costs and diminished service quality.

The problem addressed in this dissertation is, therefore, the operational inefficiency of conventional reactive and fixed-interval preventive maintenance regimes in shared bicycle fleets, combined with the inability of current survival modeling approaches to convert raw time-to-event predictions into actionable asset-level health indicators. Although initial research has demonstrated the utility of SA for isolated bicycle components [8], fleet operators still lack an integrated framework that maps conditional risk functions into a unified Health Index across components, individual vehicles, and operational fleets. To resolve this limitation, this dissertation proposes a predictive maintenance framework that derives multi-level Health Indices from censored survival models, systematically benchmarks their predictive gain over unconditioned baselines, and integrates these indicators into an interactive dashboard for proactive fleet monitoring.

1.2 Objectives and Research Question

Based on the problem described in the previous section, this dissertation is guided by the following research question: *"How can Survival Analysis turn the operational and maintenance records of a bike-sharing fleet into a component-level Health Index capable of monitoring the mechanical condition of individual bicycles and of the fleet as a whole?"* The formal hypothesis derived from this question, together with its null counterpart, is stated in Chapter 3.

To answer this question, the following objectives were defined and pursued:

1. The development of a predictive maintenance framework based on Survival Analysis, training Cox Proportional Hazards (CPH) and Gradient Boosting Survival Analysis (GBSA) models on real usage and maintenance data from the MetroBikeShare fleet (Los Angeles, USA);

2. The statistical comparison of these covariate-conditioned models against a non-parametric baseline (Kaplan-Meier) and a parametric baseline (Weibull), using evaluation metrics appropriate for censored data;
3. The derivation of a Health Index (HI) from the survival functions produced by the trained models, aggregable at the component, bicycle, and fleet levels;
4. The development of an interactive visualization tool (dashboard) that translates the Health Index into individual and fleet-wide monitoring views.

In addition to these central objectives, the transferability of the trained models to fleets with different technological and geographical profiles (LongBeach BikeShare and PikeRide) was investigated as a complementary extension and accepted for publication in [18].

1.3 Contributions

In line with the objectives above, the main contributions of this dissertation are the following:

- A predictive maintenance framework for bike-sharing fleets based on Survival Analysis, validated with real operational and maintenance data;
- A systematic comparison between covariate-conditioned survival models (CPH and GBSA) and unconditioned baselines (Kaplan-Meier and Weibull);
- A Health Index derived from survival functions, aggregable per component, per bicycle, and per fleet;
- An interactive prototype dashboard for fleet monitoring, supporting both individual and fleet-wide views.

The research conducted in the scope of this dissertation also resulted in two scientific contributions: an analysis of existing datasets for bike maintenance [11], and a cross-domain evaluation of survival models across bike-sharing fleets [18].

1.4 Document Structure

This dissertation is organized into six chapters. Chapter 2 introduces bike-sharing systems and smart cycling, reviews the maintenance paradigms applied to shared bicycle fleets, the fundamentals of Survival Analysis and Health Index construction, closing with the research gaps that motivate it. Chapter 3 formalizes the research question and hypotheses, defines the scope of the study, the mathematical formulation of the problem, and the evaluation metrics, and describes the methodological pipeline together with its validation strategy. Chapter 4 details the implementation of each stage of that pipeline, from data acquisition and processing to the survival models, the Health Index, and the

dashboard. Chapter 5 presents and discusses the results, following the same order as the validation strategy defined in Chapter 3. Finally, Chapter 6 summarizes the contributions of this work, verifies the research hypothesis, and outlines directions for future research.

Chapter 2

Background and State of the Art

Evaluating component degradation in shared bicycle fleets requires shifting from static time-to failure approximations toward event-history statistical frameworks. The stochastic operational stressors of urban cycling render fixed maintenance schedules inadequate, necessitating a theoretical foundation rooted in survival analysis. By accommodating right-censored observation windows and multi-vector operational covariates, survival modeling provides the mathematical framework required to quantify component-level risk profiles and derive a continuous asset Health Index for fleet monitoring.

2.1 Bike-Sharing Systems and Smart Cycling

Shared bicycles represent a distinct class of municipal infrastructure, operating continuously under dynamic usage patterns and unmonitored environmental conditions. Understanding the physical composition of these fleets, the digital data streams generated by smart cycling platforms, and the specific wear mechanisms affecting shared components provides the foundation for data-driven reliability modeling.

2.1.1 Soft Mobility and Smart Cycling

Soft mobility, frequently designated as active mobility, encompasses urban transport modes propelled by human energy, predominantly walking and cycling. These modes are characterized by the complete absence of direct tailpipe emissions, a spatial footprint far smaller than that of private automobiles during both circulation and parking, low operating costs for users, and substantial public health co-benefits derived from routine physical activity, making soft mobility a central instrument for municipalities seeking to decarbonize short-distance travel [3], [4]. Cycling occupies a strategic position within active transport hierarchies because it efficiently covers travel distances where walking becomes impractical and motorized transport introduces disproportionate vehicular congestion and localized environmental degradation.

The smart city paradigm has introduced a digital management layer to municipal infrastructure through pervasive sensing, ubiquitous connectivity, and data analytics [1], [19]. Applied to urban

cycling, this integration is defined as smart cycling, representing the instrumentation of the cycling ecosystem with connected bicycles, intelligent docking stations, real-time spatial data feeds, and data-driven operational services [20], [21]. Rather than installing invasive hardware retrofits or mounting dedicated sensor arrays on individual mechanical components, this dissertation operates within the scope of event-history predictive maintenance. This paradigm leverages the digital traces that a modern mobility service naturally produces, such as trip timestamps, geographical coordinates, and workshop records, transforming raw operational records into quantitative indicators of physical asset health.

2.1.2 Anatomy of a Bike-Sharing System

A BSS is an urban public transport service that grants short-term access to a distributed fleet of bicycles, catering primarily to first-mile and last-mile connections within multimodal transit networks [2]. The technological evolution of BSS is categorized across distinct generations, progressing from unregulated free-to-use bicycles to magnetic-card docked fleets, and ultimately to Internet of Things (IoT)-enabled, Global Positioning System (GPS)-tracked smart fleets [2], [20]. Modern BSS deployments are fundamentally differentiated by their docking architecture and propulsion technology, both of which exert a direct influence on asset tracking and fleet maintenance requirements.

In terms of spatial configuration, station-based systems constrain trip origins and destinations to fixed physical docks, offering predictable asset consolidation at the cost of rigid infrastructure constraints. These deployments rely on dedicated physical docking terminals that secure vehicles when idle while facilitating automated locking and asset consolidation. Conversely, as illustrated in Figure 2.1, free-floating or dockless systems permit users to leave bicycles anywhere within a geofenced service area, increasing user flexibility while complicating asset tracking, retrieval logistics, and maintenance monitoring [15], [22]. Regarding propulsion, modern fleets increasingly deploy pedal-assist electric bicycles alongside conventional mechanical bicycles [8], [23]. Electric bicycles expand service catchment areas and demographic reach by reducing physical exertion over hilly terrain, but their increased weight and motorized torque introduce distinct degradation dynamics that alter mechanical maintenance requirements [8].

From an information perspective, an operational BSS continuously generates three distinct data streams. The first stream consists of trip records, which document every journey with anonymized asset identifiers, departure and arrival timestamps, and spatial origin-destination pairs. The second stream comprises real-time availability feeds, standardized by the GBFS, which publish the instantaneous location and dock occupancy of the fleet. The third stream consists of internal workshop maintenance logs, which record every inspection, repair, and part replacement performed by technicians. A structural asymmetry exists across these data streams, as trip records and availability feeds are frequently published as open data by municipal authorities, whereas granular component-level maintenance logs remain strictly proprietary and guarded by system operators.



Figure 2.1: The Yuhuan Bicycle-sharing System as an example of dockless infrastructure [24].

2.1.3 The Shared Bicycle as a Maintenance Asset

Mechanically, a bicycle is an assembly of kinematically coupled wear components. The transmission, comprising the chain, chainrings, cassette sprockets, and bottom bracket, converts rider pedaling force into forward motion while suffering progressive material elongation. The braking system dissipates kinetic energy through frictional contact, progressively consuming brake pads and calipers. Structural loads are sustained by the frame, saddle, and wheel assemblies, with pneumatic tires absorbing pavement irregularities, while gear and brake actuation inputs are transmitted through mechanical cables or hydraulic lines. In electric bicycles, this mechanical core is complemented by an electric motor, controller, and battery pack.

What distinguishes a shared bicycle from a privately owned vehicle is not its underlying mechanical architecture, but its operational regime. A shared bicycle operates under an intensive, unmonitored, and highly stochastic duty cycle, ridden multiple times daily by diverse users who exhibit different body masses, pedaling styles, gear-shifting habits, and levels of mechanical care [25]. Shared assets lack an individual owner to detect early wear symptoms, such as minor brake misalignment or chain slack, and intervene before functional failure occurs. These vehicles remain parked outdoors continuously, subjected to environmental exposure, atmospheric corrosion, rain-induced lubricant degradation, UV radiation, and seasonal thermal cycling [26], [27].

This operational environment results in a wear regime that is simultaneously accelerated and dispersed. Bicycles deployed on the same date accumulate different stress histories based on localized urban topographies, weather conditions, and user behaviors [7], [8]. Consequently, rigid

preventive maintenance intervals based on fleet-average lifetimes inevitably lead to maintenance inefficiencies, resulting either in over-maintenance, where functional parts are replaced prematurely, or under-maintenance, where components break down unexpectedly during active rentals [5], [6]. The integration of pedal-assist electric bicycles intensifies this stress dispersion, as motor torque, higher sustained operational speeds, and increased overall vehicle mass accelerate the wear rate of consumable friction components compared to conventional mechanical bicycles operating within the same city [8], [18]. Estimating failure risk at the individual component level, conditional on asset-specific stress histories, is therefore the central technical requirement for efficient fleet management.

2.2 Maintenance Paradigms in Bike-Sharing Systems

Maintenance strategies are commonly organized into three paradigms that differ in what triggers an intervention. Reactive maintenance intervenes only after a functional failure has occurred; it extracts the maximum physical life from each component, but at the cost of unplanned downtime, emergency logistics, and safety exposure for the user [6]. Preventive maintenance replaces failure detection with fixed schedules, servicing assets at predetermined time or usage intervals; it reduces the probability of catastrophic failures, but its efficiency depends on the assumption that all assets degrade at a similar and predictable rate [5]. PdM, in turn, can be defined as an intervention triggered by an algorithmic estimation of failure probability, derived from the specific usage history and operating context of each asset [9], [11]. This dissertation adopts the predictive paradigm, following the direction established for bike-sharing systems by Grau-Escolano et al. [8], whose study of Barcelona's Bicing system is used as a recurring methodological reference throughout this document.

2.2.1 Fundamentals and Data Requirements

The decision logic of PdM requires estimating, for each asset, how failure risk evolves as a function of accumulated stress. A classical starting point in reliability engineering is the bathtub curve, which describes the failure rate of a component through three phases: an infant mortality phase with elevated failures due to manufacturing or installation defects, a useful life phase with a low and approximately constant failure rate, and a wear-out phase in which the failure rate grows as the component approaches the end of its physical life [5], [6]. Figure 2.2 illustrates the three phases and their superposition.

Fixed preventive intervals implicitly assume that every unit of a fleet traverses these phases at the same pace. PdM removes this assumption by conditioning the risk estimate on the observed history of each individual asset, the same shift toward individually conditioned, data-driven maintenance reported across recent vehicle maintenance research [29], [30].

This conditioning imposes concrete data requirements. First, PdM needs a failure history: a log of maintenance interventions with sufficient granularity to identify which component of which asset was serviced, and when. Second, it needs exposure covariates that describe how each asset

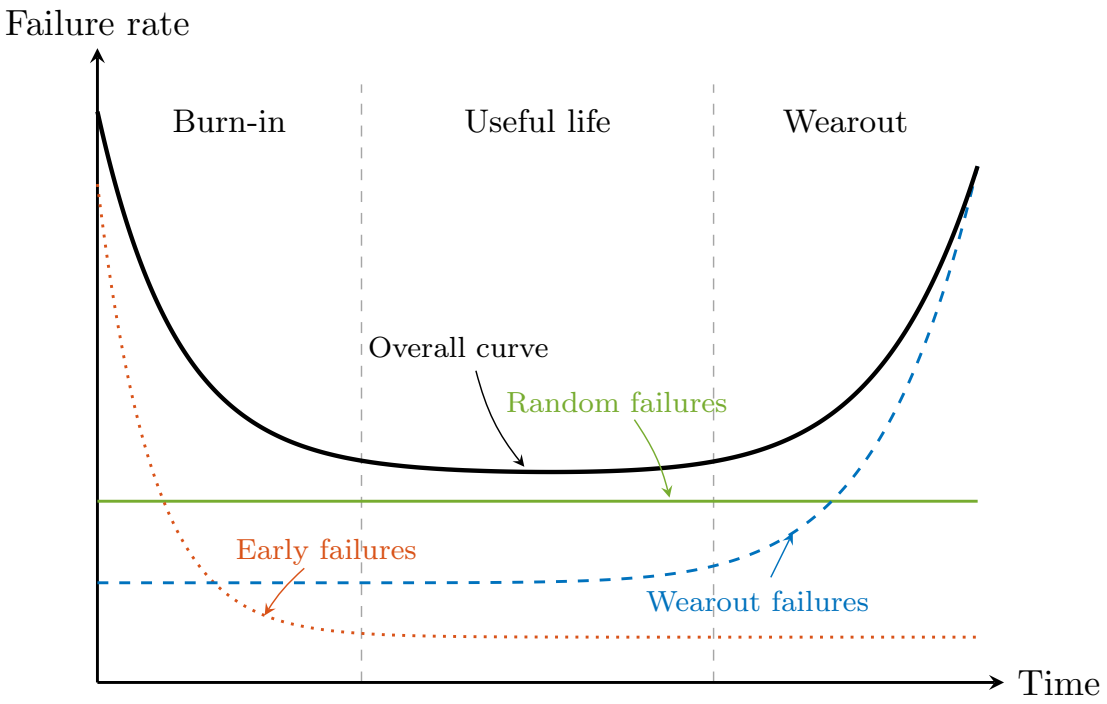


Figure 2.2: The bathtub curve as the superposition of decreasing early-life failures, a constant random-failure rate, and increasing wear-out failures, following the classical formulation in [28].

was used between interventions, such as accumulated distance, usage intensity, and environmental conditions. Third, it needs a stable asset identity, so that usage and failures can be linked to the same physical unit over time. In heavy industry, these requirements are typically met by dedicated condition-monitoring sensors installed on the equipment. Shared bicycles, however, are low-cost assets deployed in large numbers, and instrumenting every component of every bicycle is not economically feasible. The same argument applies to any numerous, low-value asset without a strong safety case for dedicated sensors, from station doors to elevators [31]. The viable alternative, and the one followed in this dissertation, is event-history PdM: reconstructing exposure from operational records (trips) and learning failure behavior from the maintenance log itself [8], [11].

2.2.2 Predictive Maintenance in Industrial and Vehicle Fleets

Outside the bike-sharing domain, the combination of PdM with time-to-event modeling is an established and growing research line. Si et al. [7] review the statistical data-driven approaches for RUL estimation, framing degradation as a stochastic process whose parameters can be learned from historical data. Survival models in particular have been applied to a wide range of fleet and equipment contexts: mobile work machinery [32], battery electric trucks [33], heavy-duty truck components [34], [35], [36], smart energy infrastructures [37], and general-purpose maintenance optimization systems [38]. Comparative studies within this literature are directly informative for model selection: Abdelhafid et al. [39] directly compare Cox and Weibull models for industrial

component RUL prediction, Rahat et al. [40] compare regression-based RUL prediction with SA, and Abad [41] couples survival-based PdM with model explainability methods.

A parallel line of research applies survival analysis and degradation models to battery prognostics, focusing on estimating the State of Health (SoH) and RUL of lithium-ion cells [42], [43], [44], [45]. This body of literature addresses electrochemical degradation mechanisms and battery management telemetry, providing valuable insights into energy storage reliability across electric transport systems.

2.2.3 Predictive Maintenance in Bike-Sharing Systems

The academic literature that accompanied the growth of bike-sharing described in Subsection 2.1.2 concentrated, however, on the demand side of operations. The logistical redistribution of bicycles across stations or service areas is a mature research topic with exact and heuristic optimization formulations [15], [16], and demand forecasting has progressed from classical spatio-temporal models [17] to graph neural networks for station-level prediction and network expansion [46], [47]. Maintenance, in contrast, received far less attention, and when addressed at all it was approached from the operational side, as a task-assignment and dispatching problem [22], rather than as a failure prediction problem.

The mechanical justification for asset-level prediction is nonetheless well documented. Component wear in bicycles is strongly modulated by the operating environment: chain lubrication degrades under rain exposure [26], drivetrain efficiency and wear vary with load and usage regime [48], coastal salinity accelerates atmospheric corrosion of steel parts [27], and winter de-icing salt is aggressive enough that municipalities weigh its effect on cycling infrastructure and equipment [49], [50]. These mechanisms make degradation asset-specific and context-dependent, the same variability that fixed preventive schedules assume away.

To the best of the knowledge gathered in this work, only a small number of studies address failure prediction for shared bicycles directly. In the Lisbon GIRA system, Monteiro [14] applied artificial neural networks to maintenance data for binary health classification, identifying cumulative distance and topographical stress as primary predictors of failure. Building upon the same network, Mendonça [51] applied linear mixed models and generalized additive models to predict remaining useful life from riding duration and user-centric covariates such as end-of-trip user ratings and socio-economic status. However, that framework aggregated maintenance events at the whole-bicycle level through sequential maintenance groups, meaning that a repair on a single component affected the tracking index for the entire vehicle. The most complete precedent, and the primary methodological comparator of this dissertation, is the study of Grau-Escolano et al. [8] on Barcelona's Bicing system: using operator-provided trip and maintenance records, the authors model component-level time-to-failure with SA, adopting a one-model-per-component strategy in which usage covariates condition individual risk estimates. Their work establishes both the feasibility of survival-based PdM in bike-sharing and the reference experimental design that this dissertation adapts. Expanding beyond localized fleet models, cross-domain evaluations across distinct systems demonstrate that survival models can transfer risk estimation capabilities across operators, where unsupervised

feature alignment techniques such as Correlation Alignment (CORAL) mitigate domain shift caused by operational discrepancies [18], [52]; transfer learning for survival models remains an emerging topic in the broader reliability literature [53].

All of these studies share one structural limitation: they depend on proprietary operator data. As established in Section 1.1 and quantified in the dataset audit conducted in the preliminary phase of this work [11], the public data landscape offers abundant trip records but almost no maintenance logs, the only exceptions being [12] and [13], while the de facto data standard for the sector exposes only availability and location information [10]. The present dissertation is no exception to this dependence: its trip records are public, but its maintenance log was obtained directly from the operator, through the acquisition campaign described in Chapter 4.

2.3 Fundamentals of Survival Analysis

Survival Analysis is the statistical framework dedicated to modeling time until an event of interest occurs. Originally established within biostatistics, its mathematical formalism adapts directly to engineering reliability by mapping health endpoints to mechanical component failure [54], [55]. This section establishes the theoretical foundations utilized throughout the dissertation, providing the mathematical definitions for censoring, risk estimation, and model evaluation that underpin the methodology formalized in Chapter 3.

2.3.1 Time-to-Event Data and Censoring

The distinguishing feature of time-to-event datasets is censoring, which occurs when the event of interest is not observed for every unit within the study window. In a fleet context, a bicycle component that accumulates usage without failing prior to the end of data collection represents a right-censored observation. This record carries partial information establishing that component lifetime exceeds the observed duration [54]. Discarding right-censored observations or treating censoring dates as functional failures introduces systematic bias into conventional regression algorithms, whereas SA estimators incorporate this incomplete temporal information [55]. By contrast, left-censoring[1] arises when the origin of an observation interval is unobserved, such as when maintenance tracking commences while an asset is already in service. To prevent parameter contamination from unknown initial wear states, the standard mitigation consists of discarding the initial, incompletely tracked interval of each asset-component pair [8].

The handling of observational boundaries is central to time-to-event formulations, as component lifecycles are rarely fully observed within a fixed data collection window. Observational trajectories are subject to distinct censoring mechanisms, as illustrated in Figure 2.3, where intervals are classified according to the visibility of their commissioning and failure events.

[1] Strictly, the mechanism described is left truncation (delayed entry) rather than left censoring, since the component is already at risk when observation begins and its exact age at that instant is unknown. The term is retained throughout for consistency with the terminology adopted in [8].

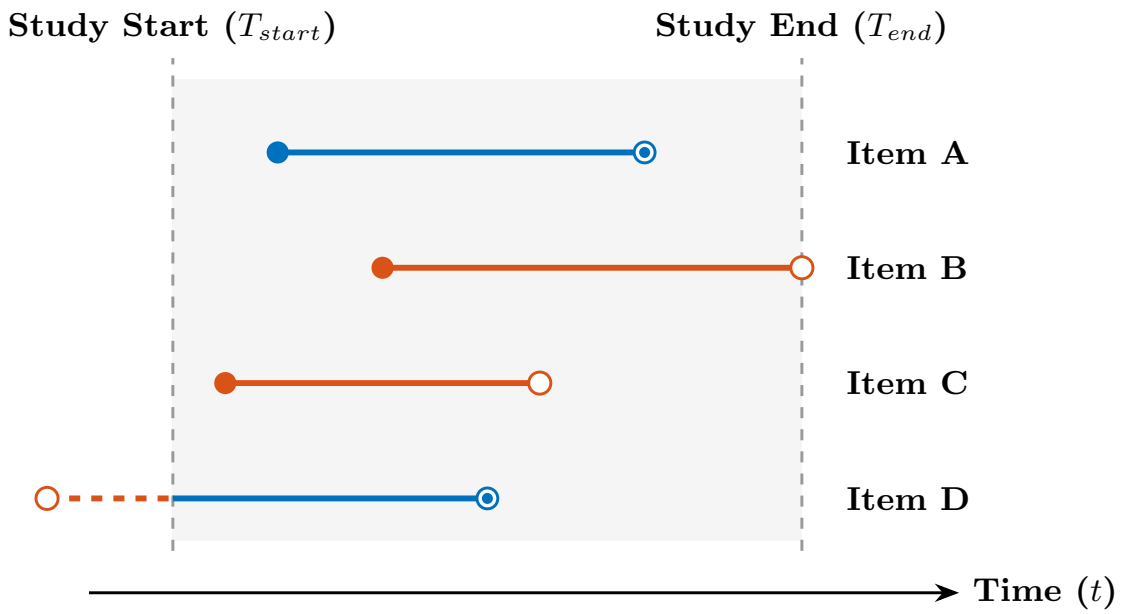


Figure 2.3: Schematic representation of observational trajectories and censoring mechanisms within a study window: uncensored failure (Item A), right-censoring from window termination or loss to follow-up (Items B and C), and left-censoring due to pre-study commissioning (Item D).

2.3.2 Survival and Hazard Functions

Let T denote a non-negative random variable representing time to the next mechanical failure. The survival function defines the probability that a component remains operational beyond time t:

$$S(t) = P(T > t), \tag{2.1}$$

while the hazard function represents the instantaneous failure rate at time t, conditional upon survival up to that instant:

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{P(t \leq T < t + \Delta t \mid T \geq t)}{\Delta t}. \tag{2.2}$$

The Kaplan-Meier estimator provides a non-parametric calculation of $S(t)$ from censored samples [56]:

$$\hat{S}(t) = \prod_{i:t_i \leq t} \left(1 - \frac{d_i}{n_i}\right), \tag{2.3}$$

where d_i represents observed failure events at time t_i and n_i denotes the count of operational units at risk immediately preceding t_i. Operating without distributional assumptions or covariate inputs, the Kaplan-Meier estimator describes population-level baseline reliability across a fleet [57].

The Weibull distribution provides the classical parametric formulation. Parametrized by scale

η and shape β, its survival and hazard functions take the form:

$$S(t) = \exp\left[-\left(\frac{t}{\eta}\right)^{\beta}\right], \qquad \lambda(t) = \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1}. \tag{2.4}$$

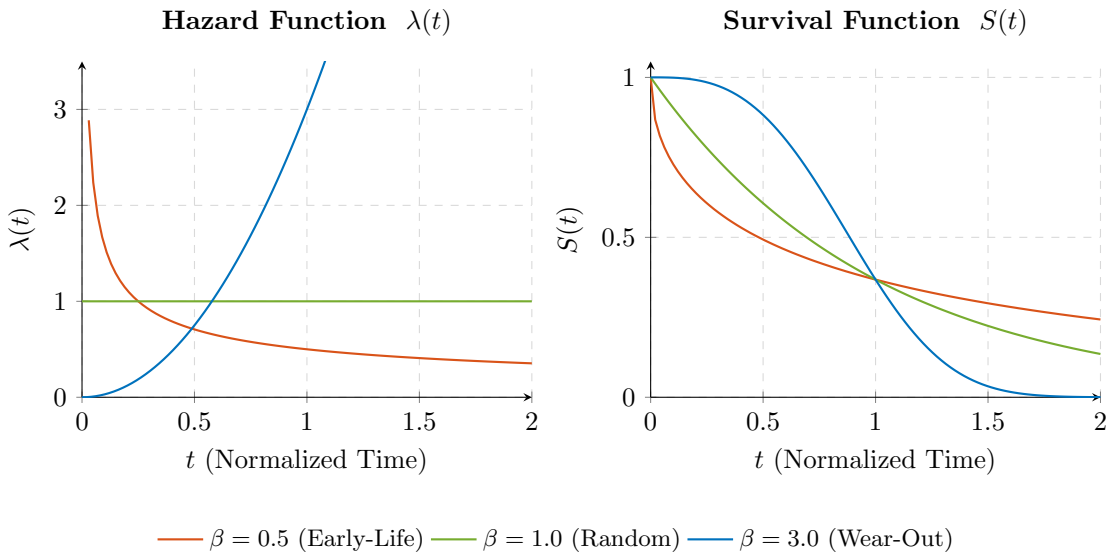


Figure 2.4: Weibull hazard and survival functions for shape parameters $\beta = 0.5$, $\beta = 1.0$, and $\beta = 3.0$ under unit scale ($\eta = 1.0$).

The shape parameter β offers direct physical insight into degradation regimes. Values of $\beta < 1$ yield a decreasing hazard indicative of early-life infant mortality, $\beta \approx 1$ produces a constant failure rate characteristic of random external shocks, and $\beta > 1$ yields an increasing hazard reflecting progressive wear-out degradation [7], [55]. Figure 2.4 illustrates these parametric regimes, aligning with the theoretical stages of the bathtub curve established in Subsection 2.2.1. Due to its mathematical tractability and clear physical interpretation, the Weibull distribution serves both as a parametric benchmark and as an anchoring function for health indexing.

2.3.3 Covariate-Conditioned Models

While population estimators capture global fleet behavior, predictive maintenance requires individualized risk profiles. The CPH model connects population baselines to individual asset profiles by expressing the conditional hazard for a unit with covariate vector $\mathbf{x}$ as:

$$\lambda(t \mid \mathbf{x}) = \lambda_0(t) \exp\left(\beta^{\top}\mathbf{x}\right), \tag{2.5}$$

where $\lambda_0(t)$ represents an unspecified baseline hazard and the exponential term scales risk multiplicatively. The regression coefficients β are estimated by maximizing the partial likelihood without specifying $\lambda_0(t)$, yielding interpretable log hazard ratios for individual stress factors. This semi-parametric efficiency assumes proportional hazards, implying that relative covariate

effects remain constant over time. Advanced formulations extend this baseline by incorporating time-varying covariates, time-dependent coefficients, and recurrent event structures suitable for repairable assets [58], [59], [60].

To overcome linear predictor constraints, GBSA replaces the linear combination $\beta^\top \mathbf{x}$ with an ensemble of regression trees $F(\mathbf{x})$, trained via gradient boosting on the Cox partial log-likelihood [61], [62]:

$$\ell(F) = \sum_{i:\,\delta_i=1} \left[F(\mathbf{x}_i) - \log \sum_{j:\,t_j \geq t_i} e^{F(\mathbf{x}_j)} \right], \tag{2.6}$$

where δ_i denotes event occurrence. This formulation captures complex non-linear feature interactions while retaining the survival structures required to construct individualized conditional survival functions $S(t \mid \mathbf{x})$, providing the foundational outputs required to derive component health indices [38], [41].

2.3.4 Evaluation Metrics for Censored Data

Evaluating models on censored lifetime data requires specialized statistical metrics. The concordance index measures discrimination by calculating the proportion of comparable pairs whose predicted risk ordering agrees with observed failure sequences [63]. Under heavy censoring, standard concordance metrics exhibit positive bias because pair comparability depends on the censoring distribution. The Inverse Probability of Censoring Weighting correction resolves this limitation, yielding the Concordance Index under Inverse Probability of Censoring Weighting (C-IPCW) statistic as the primary discrimination metric for fleet failure ranking [64].

Discrimination metrics alone do not reflect probability calibration, as models can order risk correctly while outputting miscalibrated survival probabilities [65]. Calibration over operational horizons is evaluated using the Integrated Brier Score (IBS), which computes the time-integrated mean squared error between predicted survival curves and observed binary outcomes, weighted by censoring probabilities [66]. Model stability depends fundamentally on the absolute count of observed failure events rather than total sample volume, adhering to events-per-variable statistical guidelines during feature selection and component scope definition [67], [68].

2.4 Health Indices for Fleet Monitoring

Probabilistic survival models generate continuous hazard trajectories and cumulative risk functions, yet fleet operators require interpretable metrics to evaluate physical asset condition. The operational instrument that bridges survival estimates and maintenance scheduling is the HI: a normalized indicator bounded within the $[0, 1]$ domain, where 1 denotes nominal operational integrity and values near 0 signal imminent functional failure. Within Prognostics and Health Management (PHM), the Health Index serves as an intermediate state representation in two-stage failure forecasting pipelines, where multidimensional usage and telemetry variables are compressed into a single health trajectory before estimating RUL [7], [32].

This conceptual structure builds upon long-standing reliability paradigms, including electrochemical SoH metrics established in battery health monitoring [42] and systemic Vehicle Health Index (VHI) frameworks adopted in automotive diagnostics [69]. Functional Data Analysis offers additional mathematical tools for modeling these continuous wear paths, with functional principal components connecting longitudinal telemetry streams directly to survival outcomes [70], [71].

Two primary paradigms govern index construction in reliability engineering. Deterministic formulations compute health scores through fixed algebraic expressions, such as weighted sums of normalized mileage ratios. Although computationally lightweight, deterministic metrics assume a uniform degradation rate across all operational conditions, inheriting the main theoretical limitation of mileage-based preventive maintenance. Model-based formulations, however, derive health values dynamically from fitted degradation or probabilistic risk models.

The framework formalized in this dissertation belongs to the model-based paradigm and exploits a structural identity: the conditional survival probability $S(t \mid x)$ introduced in Subsection 2.3.3, evaluated at the accumulated usage t of a specific component given its stress profile x, represents the probability that the component reaches its present state without requiring corrective intervention.

Anchoring this probability on component-specific parametric baselines and aggregating component-level scores into bicycle-level and fleet-level indices using characteristic failure weights produces a monitoring representation derived strictly from operational logs. The mathematical formalization of this multi-level hierarchy is established in Chapter 3.

Within shared micro-mobility, structured Health Index architectures remain largely unaddressed in the literature. As established in Subsection 2.1.2, the sector's data standard exposes no mechanical condition data at all. Consequently, operators lack a standardized quantitative metric to evaluate component-level degradation from operational telemetry, creating an unaddressed monitoring gap that directly motivates the research objectives formalized in Section 2.5.

2.5 Identified Research Gaps

The critical review of the state of the art highlights four primary structural limitations in existing literature, defining the scientific background and operational boundaries of this dissertation.

The primary constraint stems from data scarcity. Granular maintenance records for shared micro-mobility fleets are almost entirely absent from the public domain. The sector's data standard covers availability but not mechanical condition [10]. The audit reported in Section 4.2 confirms that public repositories contain virtually no timestamped repair logs suitable for component-level degradation modeling, forcing existing studies to depend on proprietary single-operator logs.

Second, operational research in bike-sharing has overwhelmingly prioritized fleet rebalancing logistics and demand forecasting [15], [16], [17], [46]. Maintenance is conventionally treated as a reactive or scheduled operational cost rather than a probabilistic estimation problem, despite mechanical evidence showing that component wear is asset-specific and driven by cumulative operational stress.

Third, a generalizability gap exists across heterogeneous operating environments. Existing studies on bicycle component failure prediction rely exclusively on single-operator datasets within isolated urban contexts [8], [14]. The extent to which predictive degradation models transfer across fleets operating under differing technological configurations, climatic conditions, and topographies remains understudied in the literature.

Finally, a diagnostic monitoring gap exists within operational software. No structured Health Index framework specifically tailored for bike-sharing fleets was identified in the reviewed literature. Existing platforms lack computational mechanisms to translate conditional survival probabilities into standardized vehicle-level and fleet-level condition metrics required for decision support.

The methodology formalized in Chapter 3 directly addresses the core challenges of data accessibility, probabilistic failure modeling, and operational health visualization in response to the research question and objectives formulated in Section 1.2.

To clearly delineate the contributions of this dissertation from prior research, Table 2.1 provides a structured comparison between representative literature and the proposed methodology across key technical dimensions.

Table 2.1: Comparison of related literature and the proposed predictive maintenance framework.

Reference	**BSS Domain**	**Survival Analysis**	**Real Maint. Logs**	**Health Index Formulation**	**Decision Dashboard**
Lu et al. [22]	✓	×	✓	×	×
Monteiro [14]	✓	×	✓	×	×
Mendonça [51]	✓	×	✓	×	×
Grau-Escolano et al. [8]	✓	✓	✓	×	×
Coutinho et al. [38]	×	✓	✓	×	×
Proposed Framework	✓	✓	✓	✓	✓

As highlighted in Table 2.1, while existing bike-sharing studies have increasingly explored data-driven maintenance, they either rely on classification heuristics or restrict survival modeling to raw time-to-event predictions without translating risk curves into monitoring indicators. On the other hand, broader predictive maintenance literature establishes mathematical formulations for Health Indices but rarely validates them on complex, highly censored urban mobility fleets. This dissertation fills this gap by integrating censored survival modeling, component-to-fleet Health Index derivation, and interactive visual monitoring within a unified architecture.

Chapter 3

Methodology

Predictive maintenance in shared micro-mobility relies on processing routine operational records through probabilistic time-to-event architectures. Mechanical degradation is modeled at the component level using survival analysis, incorporating censoring-aware evaluation metrics such as C-IPCW and IBS to assess model discrimination and calibration. These probabilistic estimates directly inform the mathematical formulation of the Health Index, establishing a unified scale for individual asset assessment and fleet-wide operational monitoring.

3.1 Research Question and Hypotheses

The research question stated in Section 1.2 investigates how survival analysis can transform operational and maintenance logs into a component-level HI for monitoring individual bicycles and overall fleet condition. This question involves two key dimensions. First, it requires a practical method to map predicted survival probabilities into an aggregated HI, which is developed in Subsection 3.3.3 and evaluated in Chapter 5. Second, it rests on a fundamental statistical assumption: that routine usage and exposure records carry enough predictive signal to forecast component failures better than population-level averages.

Building on the maintenance paradigms and statistical foundations established in Chapter 2, verifying this premise requires comparing covariate-conditioned models against population baselines across two complementary performance dimensions: discrimination, which measures the ability to rank asset failure risk under right-censoring, and calibration, which assesses the numerical accuracy of predicted survival probabilities over time.

This reasoning leads to the following testable hypotheses:

H_1. Survival models conditioned on individual usage intensity and environmental exposure history (CPH and GBSA) achieve superior discrimination and calibration in predicting component failures compared to covariate-blind baselines (Kaplan-Meier and Weibull), when evaluated on operational bike-sharing data using censoring-aware metrics.

H_0**.** Conditioning on individual usage and exposure covariates provides no predictive advantage over covariate-blind baselines, indicating that operational trip logs and environmental data contain insufficient signal to differentiate individual component risk beyond population averages.

Rejecting H_0 confirms that routinely collected operator data carries sufficient predictive information to drive asset-specific, condition-based monitoring. The verification protocol for these hypotheses is established in Section 3.4, and the empirical findings are presented in Chapter 5.

3.2 Scope of the Study

The scope of this dissertation is delimited across four core dimensions: the target fleet, the evaluated mechanical components, the prediction target, and the selected modeling families.

Fleet. The primary empirical analysis is conducted on the MetroBikeShare system of Los Angeles. Its trip records are publicly available, while its component-level maintenance log was acquired through the targeted operator engagement campaign described in Chapter 4. Following a quality triage of candidate repositories, this fleet was selected for its fine maintenance record granularity, multi-year operational coverage, and stable bicycle identifiers that allow usage histories to be linked to individual physical assets over time. Two additional fleets with distinct technological, climatic, and topographic profiles, LongBeach BikeShare and PikeRide in Colorado Springs, were harmonized under the same data processing pipeline and evaluated in a dedicated cross-domain study [18].

Components. Six mechanical components are modeled: brakes, saddle, wheels, tires, chain, and cables/hydraulics. These subsystems were selected due to their high representation frequency in the maintenance log, their consistent presence across harmonized fleets, and their sufficient failure event counts to support model fitting in accordance with events-per-variable guidelines [67]. The traction battery is deliberately excluded from this mechanical framework: its degradation is driven by electrochemistry and charge-discharge cycles rather than accumulated mechanical stress, requiring fundamentally different diagnostic signals that do not reflect structural bicycle health. This exclusion is maintained consistently across exploratory analysis, predictive modeling, and Health Index derivation.

Within the component taxonomy detailed in Table 4.4, modeling is applied to the category that carries the recorded failure events: the *Brakes* model is fitted on the `Braking (General)` records and the *Wheels* model on `Wheels (General)`, while `Tires` and `Cables/Hydraulics` are modeled as components in their own right. Named sub-parts such as brake pads, calipers, spokes, rims, and hubs are catalogued in the harmonized log but are not modeled individually. This restricts the scope of the braking and wheel models to the general-category records and is one factor behind the low brake event count reported in Table 5.1.

Prediction target. The prediction target is the operational duration until the next corrective intervention for a specific component on a given bicycle, formulated as a right-censored time-to-event variable. This study does not attempt to isolate specific physical failure modes, pinpoint

localized subassembly defects, or directly prescribe automated maintenance schedules. Instead, it estimates component failure risk and translates these probabilistic risk functions into a structured monitoring representation.

Modeling families. Two primary covariate-conditioned survival model families are trained per component: the semi-parametric CPH model, retained for its interpretability and direct comparability with reference bike-sharing literature [8], and GBSA, adopted as a non-linear counterpart capable of capturing complex feature interactions while maintaining methodological continuity with related cross-domain work [18]. These models are evaluated against the non-parametric Kaplan-Meier and parametric Weibull baselines defined in Section 3.1.

3.3 Problem Formulation

Translating mechanical fleet degradation into a survival analysis framework requires mapping discrete operational records into formal statistical representations. The central analytical unit adopted in this dissertation is the Maintenance Operation (MO) unit, following the interval-based formulation and terminology established by Grau-Escolano et al. [8]. This construct converts continuous temporal logs and discrete repair events into structured time-to-event observations suitable for survival modeling.

3.3.1 Maintenance Operation Units and Event Definition

Because shared bicycles are repairable physical assets, each bicycle-component pair generates a temporal sequence of recurring lifetimes rather than a single terminal trajectory, a structure known in reliability literature as recurrent event data [60]. The formulation segments this sequence into discrete intervals. A MO unit represents the interval of a specific component c on bicycle b between two consecutive maintenance interventions, defined by the tuple

$$\mathrm{MO}_{b,c,k} = (\mathbf{x}_{b,c,k},\, t_{b,c,k},\, \delta_{b,c,k})\,, \tag{3.1}$$

where k indexes the sequence of interventions, $t_{b,c,k}$ denotes the operational duration in calendar days, $\delta_{b,c,k}$ is the binary event indicator, and $\mathbf{x}_{b,c,k}$ is the feature vector summarizing usage intensity and environmental strain accumulated over that specific interval. Measuring duration in calendar days establishes a uniform temporal scale across bicycles operating under varying usage regimes, while cumulative mileage is integrated as an explanatory covariate in accordance with established fleet predictive maintenance practices [18], [33]. The assignment of event status $\delta_{b,c,k}$ depends on the severity classification of the intervention closing the interval, determined by the standardized maintenance taxonomy detailed in Chapter 4. Interventions classified at or above severity level 3, which encompass component repairs and complete part replacements, mark functional life termination and are assigned $\delta_{b,c,k} = 1$. Lighter interventions, such as inspections, cleaning, lubrication, or mechanical adjustments (severity levels 1 and 2), do not reset the degradation clock; instead, they are tracked as a minor-maintenance count covariate ($\text{maint_count}_{c,\text{minor}}$) within

the active interval. Intervals terminated by purely preventive servicing or remaining open at the conclusion of the data collection window are classified as right-censored observations ($\delta_{b,c,k} = 0$), aligning with industry modeling standards [72]. Left-censoring is addressed by discarding the first MO unit of each bicycle-component pair, as its true installation date cannot be established from historical logs [8], while retaining all associated trip data to maintain accurate fleet-wide usage accumulation.

3.3.2 Covariates and Feature Aggregation

The feature vector $\mathbf{x}_{b,c,k}$ is derived from a daily per-bicycle operational dataset that merges three primary data streams: individual trip logs, route geometry extracted from self-hosted routing services, and historical meteorological records [18]. Because survival models ingest a static covariate vector per MO unit while operational exposures fluctuate on a daily basis, time-varying daily metrics are transformed into interval-level summary statistics.

For each MO interval, daily time series capturing usage intensity, mechanical speed, and environmental exposure are aggregated into statistical measures that summarize exposure level, dispersion, and temporal trajectory. Specifically, the mean, standard deviation, and linear gradient are calculated across the interval duration for candidate metrics. To ensure feature independence and mitigate multicollinearity without making linear distribution assumptions, two distinct Spearman rank correlation analyses are performed during feature screening. First, candidate operational and usage predictors are evaluated, applying an absolute threshold of $|\rho| > 0.85$ to eliminate redundant daily metrics while retaining the variable demonstrating higher physical interpretability or direct mechanical relevance to component wear. Second, an independent Spearman screening combined with hierarchical clustering, Coefficient of Variation, and sparsity checks is applied to the meteorological stream, consolidating internally collinear weather blocks, such as pressure and temperature variants, into a parsimonious set of environmental covariates.

Consistent with the foundational formulation established by Grau-Escolano et al. [8], the feature vector incorporates the component-specific cumulative distance accumulated within the observation interval (duration_km), providing a direct quantitative metric of mechanical work. This usage metric is complemented by daily operational rate and dispersion statistics, as well as cross-component degradation indicators such as the interval-averaged accumulated mileage ($\text{km_accum}_{c',\text{mean}}$) and the initial maintenance age ($\text{days_since}_{c',\text{start}}$) of other mechanical subsystems $c' \neq c$ on the same asset. To prevent data leakage, other non-distance cumulative interval totals, such as total active hours or trip counts, are explicitly excluded when they serve as deterministic proxies for the observation duration $t_{b,c,k}$ [18]. Combined with the minor-maintenance count covariate ($\text{maint_count}_{c,\text{minor}}$) defined in Subsection 3.3.1, this feature aggregation and dual-screening strategy aligns with data-driven predictive maintenance frameworks for transport fleets [33], balancing statistical parsimony with predictive expressiveness.

3.3.3 Models and Health Index Mapping

Modeling is conducted independently for each of the six selected mechanical components, adhering to the one-model-per-component structure established by Grau-Escolano et al. [8]. Dedicated component-level models prevent the dilution of distinct wear mechanisms, such as frictional degradation in brake pads, structural material fatigue in wheels, or environmental oxidation in drive chains.

For a given component c, the trained survival model estimates the conditional individual survival function $S_c(t \mid \mathbf{x}_{b,c})$, representing the probability that component c on bicycle b survives beyond operational age t given its specific interval exposure vector $\mathbf{x}_{b,c}$ without requiring corrective repair. The component-level Health Index ($\mathrm{HI}_{b,c}$) is defined directly through this conditional survival probability:

$$\mathrm{HI}_{b,c}(t) = S_c(t \mid \mathbf{x}_{b,c})\,. \tag{3.2}$$

To provide fleet operators with an interpretable asset-level monitoring signal, individual component indices are aggregated into a composite bicycle HI (HI_b). The aggregation employs a weighted linear combination where component weights w_c are defined inversely proportional to each component's empirical characteristic lifetime η_c:

$$\mathrm{HI}_b(t) = \frac{\sum_c w_c \,\mathrm{HI}_{b,c}(t)}{\sum_c w_c}, \quad \text{where } w_c = \frac{1}{\eta_c} \tag{3.3}$$

The characteristic lifetime η_c represents the scale parameter of a two-parameter Weibull distribution fitted via Maximum Likelihood Estimation (MLE) on the unconditioned observation intervals of component c, explicitly accounting for right-censored records in the training partition. Physically, η_c denotes the operational duration by which approximately 63.2% of the component population is expected to undergo corrective intervention. Weighting components by $1/\eta_c$ ensures that high-wear, shorter-lived mechanical subsystems exert a proportionally greater influence on the composite vehicle health state. System-wide monitoring views are subsequently constructed by evaluating Equation 3.3 across all active fleet assets at any given calendar date.

To quantify the predictive benefit of conditioning risk estimates on operational and environmental covariates, a deterministic baseline index is established using unconditioned parametric survival. The baseline index ($\mathrm{HI}^{\mathrm{W}}_{b,c}$) evaluates survival strictly as a function of operational age t via a two-parameter Weibull model:

$$\mathrm{HI}^{\mathrm{W}}_{b,c}(t) = \exp\left[-\left(\frac{t}{\eta_c}\right)^{\beta_c}\right], \tag{3.4}$$

where shape parameter β_c and scale parameter η_c are fitted via MLE on the training partition. Contrasting the covariate-conditioned index of Equation 3.2 against the deterministic baseline of Equation 3.4 directly isolates the performance gains attributable to operational intensity and environmental stressors.

3.4 Evaluation Metrics and Hypothesis Verification

Model quality is assessed along the two complementary analytical axes established in Subsection 2.3.4: discrimination and calibration.

Discrimination is measured with the C-IPCW [64], adopted as the primary evaluation metric of this dissertation. This choice follows directly from the operational data regime: fleet maintenance logs exhibit high censoring rates, a condition under which Harrell's classical C-index becomes optimistically biased because the set of comparable pairs it evaluates depends directly on the censoring distribution itself. C-IPCW re-weights comparable pairs by the inverse probability of censoring, restoring an unbiased estimate of the probability that the model ranks failure times correctly, where a value of 0.5 corresponds to random ranking and 1.0 represents perfect concordance.

Calibration is measured with the IBS [66], defined as the time-integrated mean squared error between predicted survival probabilities and observed outcomes, weighted for censoring. The IBS offers an absolute reference point: an uninformative model constantly predicting a probability of 0.5 yields a score of 0.25, meaning any informative predictive model must remain clearly below this baseline threshold. IBS is reported alongside C-IPCW because correct ranking does not guarantee well-calibrated probabilities: a survival model can rank asset failure risks accurately while producing poorly calibrated absolute probabilities, and recent literature explicitly demonstrates that discrimination metrics should not be optimized in isolation [65]. During exploratory analysis, population-level survival contrasts between asset groups, such as mechanical versus electric bicycles, are examined by comparing their respective Kaplan-Meier curves and fitted Weibull scale parameters [55]. To establish a direct numerical bridge with the reference bike-sharing study, which relies on regression-style metrics, Root Mean Square Error (RMSE) is also calculated strictly on uncensored observations in the test partition [8].

Hypothesis verification is built directly on this dual-axis evaluation framework. For each of the six mechanical components, the covariate-conditioned models (CPH and GBSA) are evaluated on the held-out test partition using C-IPCW and IBS. Values are read against two reference points established analytically rather than measured from a fitted baseline: the random-ordering level (C-IPCW $= 0.5$), which any covariate-blind model, including Kaplan-Meier and Weibull, produces by construction, since identical predictions for every unit tie every comparable pair; and the null-model threshold (IBS $= 0.25$), the calibration error of a model that constantly predicts a 0.5 survival probability. The null hypothesis H_0 is rejected for a given component if the covariate-conditioned models clear both reference points, achieving a C-IPCW clearly above 0.5 and an IBS clearly below 0.25. The main research hypothesis H_1 is supported at the study level if this performance superiority holds across the subsystem set rather than in isolated components. Because Kaplan-Meier and Weibull are covariate-blind and yield population-average estimates, this contrast serves as a direct statistical verification of whether individual operational usage and environmental exposure features hold genuine predictive value for fleet degradation.

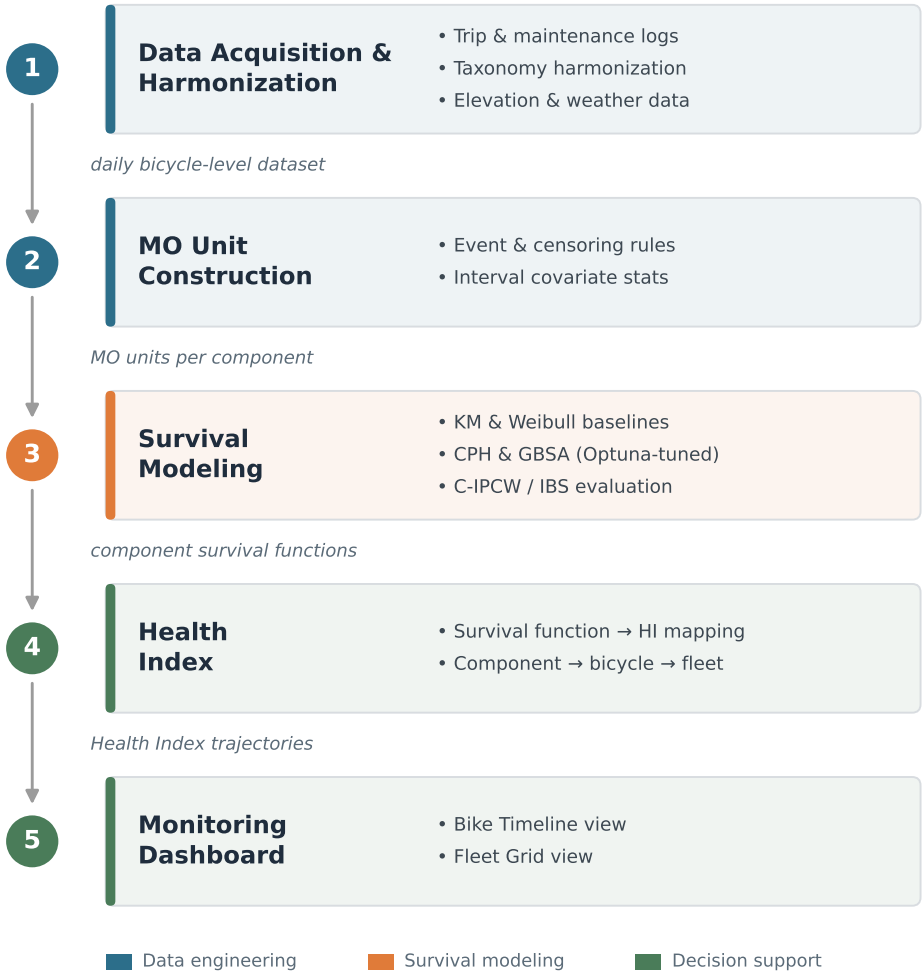


Figure 3.1: Overview of the proposed methodological workflow. Figure generated with the support of Claude and manually reviewed by the author.

3.5 Methodological Pipeline

The methodology is organized as a layered pipeline, in which each layer consumes the validated output of the previous one.

The pipeline comprises five layers, illustrated in Figure 3.1:

1. **Data Acquisition and Harmonization.** Ingests the operator's trip and maintenance records, applies quality filters, harmonizes the maintenance free-text into the component taxonomy, and produces the daily bicycle-level dataset enriched with route distance and historical weather.

2. **MO Unit Construction.** Constructs the MO units of Subsection 3.3.1, applying the event definition and censoring rules, and aggregates covariates at the interval level as detailed in Subsection 3.3.2.

3. **Survival Modeling.** Fits the two covariate-blind baselines and trains the two covariate-conditioned models per component, with hyperparameters for the latter selected through systematic optimization [73], and evaluates the covariate-conditioned models against the reference points established in Section 3.4.

4. **Health Index.** Derives the Health Index of Subsection 3.3.3 from the trained survival functions, aggregable from the component to the bicycle and to the fleet.

5. **Monitoring Dashboard.** Renders the resulting Health Index in the interactive monitoring dashboard.

The implementation of each layer is the subject of Chapter 4.

3.6 Validation Strategy

Each layer of the pipeline is validated with criteria appropriate to its nature, and Chapter 5 reports the outcomes in this same order.

Data pipeline. The harmonization layer is validated through data quality auditing: verification of the cleaning filters against the distributions that motivated them, coverage checks of the component taxonomy over the maintenance log, and consistency checks of the resulting daily dataset (asset identity, temporal continuity, and physical plausibility of derived features). A dedicated leakage audit inspects the correlation structure between candidate covariates and the survival target, removing features that are deterministic functions of the outcome; its findings materially constrained the feature space and are reported with the implementation.

Survival models. The modeling layer is validated on a held-out test partition obtained by an 80/20 split at the level of the bicycle, not of the observation unit: every interval of a given bicycle falls on the same side of the split, so that no model is evaluated on a bicycle it has already seen. A purely temporal split was discarded after producing severely imbalanced event rates between partitions. For each component, CPH and GBSA are evaluated on C-IPCW and IBS against the random-ordering and null-model reference points defined in Section 3.4, and are additionally inspected for internal validity (coefficient plausibility for CPH, feature relevance for GBSA). This layer carries the hypothesis test of Section 3.4, and includes the drop-column sensitivity screening reported in Appendix A, which informs the feature configuration used in all subsequent experiments.

Health Index. The monitoring layer is validated through behavioral criteria: component and bicycle trajectories must be bounded in $[0, 1]$, decrease with accumulated usage between interventions, and recover after corrective maintenance; degenerate behaviors (oscillation, discontinuities not aligned with maintenance events) are treated as defects to be diagnosed and corrected. The bicycle-level aggregation is inspected for coherence with its component inputs; the model-based index is contrasted with the deterministic Weibull baseline of Equation 3.4, making explicit how much the covariates change the monitoring picture; and the resulting fleet distribution is examined for discriminative spread and for the plausibility of the threshold-based categorization derived from it, since an index that assigns similar values to all assets carries no planning value.

Dashboard. The visualization layer is validated functionally: the Bike Timeline and Fleet Grid views must render the computed indices without transformation errors, support the interaction flows required for planning (asset selection, temporal navigation, fleet snapshot), and remain a faithful projection of the underlying model outputs.

This multi-tiered evaluation methodology establishes a cohesive validation framework across the entire system. The data layer confirms the technical integrity and physical plausibility of the underlying features, while the predictive modeling layer formally addresses the primary research question. The monitoring and visualization layers demonstrate that statistical risk estimations can be effectively translated into actionable operational instruments, fulfilling the main applied contributions detailed in Section 1.3.

Chapter 4

Implementation

The methodology formalized in Chapter 3 defines the conceptual architecture of the predictive maintenance framework. The software pipeline implements the methodological requirements by converting unstandardized operational feeds and workshop logs into an audited feature space. This chapter details the technical implementation of that pipeline, documenting the engineering decisions, cleaning thresholds, semantic harmonization taxonomies, model training procedures, Health Index computations, and user interface deployment.

4.1 Tools and Technologies

The implementation is developed in Python, combining dedicated survival analysis libraries with the standard scientific computing and geospatial ecosystem; the monitoring dashboard, in contrast, is implemented as a self-contained web artifact with no Python dependency at runtime. Table 4.1 summarizes the tools used across the pipeline, from survival modeling to route enrichment and fleet visualization.

The survival modeling stack is split across two libraries because no single package covers both the classical and the tree-based side of the framework established in Chapter 3. Lifelines [74] implements the non-parametric Kaplan-Meier estimator, the parametric Weibull fit, and the semi-parametric CPH model with its Breslow baseline hazard, the three formulations that admit a closed, interpretable form. Scikit-survival [75], in turn, implements GBSA as a gradient-boosted tree ensemble, and is also the source of the two evaluation metrics that the fleet's censoring rates specifically require: the Concordance Index under Inverse Probability of Censoring Weighting and the Integrated Brier Score, discussed in Section 3.4.

Table 4.1: Main Python libraries and software tools used in the implementation.

Library	**Main purpose**	**Use in this dissertation**
lifelines	Survival analysis	Fitting non-parametric Kaplan-Meier baselines, parametric Weibull models, and semi-parametric CPH regression.
scikit-survival	Survival analysis	Fitting GBSA models and evaluating censoring-aware metrics, specifically C-IPCW and IBS.
Optuna	Hyperparameter tuning	Automating hyperparameter optimization for CPH and GBSA using Tree-structured Parzen Estimator sampling with median pruning.
pandas	Tabular data processing	Managing trip records, maintenance logs, daily bicycle summaries, and Maintenance Operation units across processing pipelines.
NumPy	Numerical computing	Supporting vectorized mathematical operations for daily usage feature aggregation and cumulative matrix computations.
SciPy	Scientific computing	Executing Spearman rank correlation, hierarchical clustering, and Maximum Likelihood Estimation routines for statistical screening.
Matplotlib / Seaborn	Data visualization	Generating diagnostic figures, correlation matrices, Kaplan-Meier curves, and feature importance visual representations.
Requests	HTTP client operations	Issuing automated API queries to the local OpenRouteService routing engine and historical weather repositories.
OpenRouteService	Geographic routing engine	Extracting route geometries, road network distances, and point-by-point elevation profiles for trip origin-destination pairs.
Docker	System containerization	Self-hosting local OpenRouteService instances to bypass public API request limits and unify route calculation parameters.
Open-Meteo	Meteorological data source	Providing historical weather covariates, including temperature, precipitation, and wind speeds matched to local fleet coordinates.
HTML / CSS / JS	Front-end interfaces	Implementing the standalone interactive fleet monitoring dashboard without reliance on external server infrastructure.
Chart.js	Dynamic web visualization	Rendering time-series charts and status grids within the Bike Timeline and Fleet Grid interactive dashboard modules.

Route and weather enrichment relies on two external services rather than a bundled Python library. OpenRouteService, self-hosted in a Docker container for the reasons detailed in Subsection 4.3.3, computes route geometry and elevation for every trip; Open-Meteo supplies the matching

historical weather record, with the data gathering process detailed in Subsection 4.3.4. Both are reached through Requests, keeping the enrichment layer a thin, auditable HTTP client rather than a dependency on a proprietary mapping Application Programming Interface (API). Development itself was carried out on Google Colab, with every stage of the pipeline reading from and writing to a shared Google Drive folder structure, so that a script's inputs and outputs remain traceable to the fleet and model configuration that produced them.

With the exception of the operator-provided maintenance log, every tool listed in Table 4.1 is open-source or freely accessible, and every dataset that feeds the geospatial and weather covariates can be regenerated by any operator willing to publish an equivalent trip and maintenance record.

4.2 Data Acquisition and Fleet Selection

Public trip feeds for bike-sharing systems are widely accessible, whereas mechanical maintenance data remains largely unpublished. A systematic audit across public repositories, including Kaggle, Zenodo, Mendeley Data, GitHub, and IEEE DataPort, was executed using the search query `("bike" OR "bicycle") AND ("maintenance" OR "repair" OR "fix" OR "failure")`. Although this search yielded numerous mobility datasets, only two datasets containing maintenance-related logs were identified in total: the Santander Cycle Scheme servicing records released by Transport for London [13] and a user-contributed bicycle predictive maintenance dataset hosted on Kaggle [12]. However, neither dataset satisfies the structural requirements for component-level survival modeling. The first consists of high-level workshop logs that omit unique bicycle identifiers and trip histories, making it impossible to reconstruct accumulated mileage or usage stress for individual assets. Conversely, the Kaggle repository lacks real-world operational dynamics, verified failure modes, and localized environmental context. Most remaining mobility datasets are built exclusively for spatio-temporal demand forecasting and carry no maintenance information whatsoever. Without linking cumulative mechanical load and environmental stressors to specific repair events on identifiable bicycles, estimating RUL or training component-level survival models is impossible. Primary maintenance logs had to be requested directly from operators worldwide.

The Meddin Bike-sharing World Map [76] served as the primary directory. A custom Python scraper extracted operator names, geographic locations, and operational statuses, producing an initial database that was manually curated to isolate active systems. This process yielded 358 candidate operators across six continents, each contacted directly with formal requests for anonymized maintenance logs. The complete directory of all 358 contacted operators, including their contact details, outreach attempt histories, and operational response statuses, is made publicly available at [77]. This dataset is shared to facilitate future empirical research and outreach within the bike-sharing domain.

The response rate from the outreach campaign highlights the proprietary nature of fleet health data: only 1.96% of the contacted operators (seven in total) provided their maintenance records. Among the non-responding majority, 8.94% explicitly reported holding no structured maintenance

logs, while the remaining 89.11% either declined participation or remained unresponsive. Table 4.2 characterizes the seven participating operators, and Figure 4.1 illustrates their geographic distribution.

Table 4.2: Overview of operators providing usable maintenance data, including their geographical, climatic, and topographic characteristics.

Operator	**Location**	**Climate / Topography**	**Notable Stressor**
Encicla	Medellín, Colombia	Tropical, 1,500 m valley	Humidity-driven oxidation [26]; high drivetrain torque on hillside routes
Tartu Smart Bike	Tartu, Estonia	Subarctic, severe winters	Road salt corrosion [50]; e-bikes stored indoors in winter
WeGoShare	Portugal & Scotland	Mediterranean / Oceanic	Identical hardware under two climates; saline wind versus UV exposure
Flamingo	New Zealand	Southern Hemisphere, dockless	Inverted seasonality; structural damage from tipping and vandalism
MetroBikeShare	Los Angeles, USA	Mediterranean, flat terrain	Start-stop traffic acceleration and drivetrain wear [48]
LongBeach BikeShare	Long Beach, USA	Mediterranean, coastal	Atmospheric salinity and coastal frame corrosion [27]
PikeRide	Colorado Springs, USA	Semi-arid, high altitude	Electric motor cooling under thermal and altitude stress [78]



Figure 4.1: Spatial distribution of the seven participating operators characterized in Table 4.2.

Following data ingestion, each candidate repository underwent an exploratory quality triage to evaluate completeness, spatio-temporal granularity, and input consistency. Four candidate fleets were excluded during this assessment due to severe structural limitations:

- **Flamingo:** The dataset was dominated by uninformative manual entries, with over half of the recorded interventions indicating no maintenance action taken and a significant share concerning helmet management rather than bicycle mechanics.
- **Tartu Smart Bike:** The repository consisted of unnormalized manual logs recording only total fleet mileage without component-level attribution. Tthe complete absence of trip records made it impossible to correlate environmental exposure or usage intensity with component wear.
- **Encicla:** Maintenance tags were recorded at the general system level rather than attributed to specific bicycle components, and individual trip logs were absent (providing only daily aggregated distance).
- **WeGoShare:** Despite having a comprehensive and detailed structure, it presented an extremely low volume of maintenance events and trip records relative to its timeframe, leading to severe data sparsity that hinders machine learning and survival model convergence.

The triage retained three fleets: MetroBikeShare, LongBeach BikeShare, and PikeRide. Among these, MetroBikeShare was selected as the primary dataset for this dissertation due to its structural completeness, superior data volume, and longitudinal coverage. Spanning over four years of continuous logging, MetroBikeShare contains over 1.71 million raw trip records and 247,948 raw workshop records across a fleet of 2,637 bicycles. The system is technologically heterogeneous, operating both standard mechanical bicycles and pedal-assist electric bicycles within identical geographical and climatic boundaries. To formally distinguish between these technologies, standard mechanical bicycles are designated as *m-bikes* while pedal-assist electric bicycles are referred to as *e-bikes*, adopting the terminology established by Grau-Escolano et al. [8]. This dual-propulsion structure makes it possible to isolate drivetrain and propulsion mechanisms as explanatory wear factors without introducing environmental confounding variables.

The remaining two fleets, LongBeach BikeShare (a fully mechanical fleet with shaft-drive units) and PikeRide (a fully electric fleet operating in a high-altitude, semi-arid environment), were harmonized under the same processing pipeline. These datasets were reserved for the cross-domain transfer and domain adaptation evaluations presented in [18], testing the generalizability of the trained survival architectures across operationally distinct urban environments.

To illustrate the physical deployment of the MetroBikeShare fleet within its operating environment, Figures 4.2a and 4.2b show the station infrastructure located at the intersection of Bagley and Venice in Los Angeles, highlighting the solar-powered smart kiosk and docked e-bikes and m-bikes.



(a) Physical docking array and fleet assets of the MetroBikeShare system at the Bagley and Venice station in Los Angeles, featuring docked e-bikes and m-bikes [79].



(b) Solar-powered user payment station of the MetroBikeShare fleet at the Bagley and Venice station in Los Angeles [80].

Figure 4.2: MetroBikeShare station infrastructure at Bagley and Venice: (a) docking terminals and fleet assets; (b) solar-powered payment kiosk.

4.3 Data Processing and Harmonization

The processing layer ingests three primary data streams, each presenting distinct structural and logging challenges. Public trip feeds generated automatically by station docking infrastructure inherently contain sensor logging noise and measurement artifacts, including non-physical travel times, unrealistic operating speeds, and unclosed rental sessions resulting from improperly docked bicycles. In contrast, maintenance logs recorded manually by workshop technicians present semantic heterogeneity, featuring unstandardized free-text component descriptions and action labels that vary across operators and individual work orders. Finally, route geometry and local weather exposure are absent from operator databases and require external enrichment. Because MetroBikeShare, LongBeach BikeShare, and PikeRide are processed through a unified framework, all cleaning thresholds, taxonomy mappings, and feature engineering procedures are enforced consistently across fleets to ensure methodological comparability and reproducibility.

4.3.1 Trip Cleaning

To preserve data integrity and prevent systematic distortion in downstream predictive models, a sequential filtering pipeline was implemented based on the empirical distributions of each fleet. Table 4.3 summarizes these sequential rules and their quantitative impact on the primary MetroBikeShare feed.

Table 4.3: Trip-level outlier removal rules, applied sequentially to the MetroBikeShare feed.

Rule	**Criterion**	**Trips removed**
Implausibly short	Duration below 2 minutes	44,033 (2.63%)
Implausibly long	Duration above 150 minutes	42,399 (2.53%)
Stuck / abandoned bicycle	One-way trip, duration above 60 minutes, displacement below 1 km	10,513 (0.63%)
Implausible speed	Average speed above 25 km/h	5,809 (0.35%)
Total removed	Sequential application of criteria	**102,754 (6.13%)**

The raw MetroBikeShare dataset opens with 1,710,605 trip entries. Initial spatial filtering removes 35,479 entries associated with three non-station identifiers used internally by the operator. Almost all of these belong to a single virtual station code recording bicycle withdrawals for servicing or rebalancing; a second code flags out-of-service-area returns, and a third corresponds to a free-floating mode that the system no longer operates and that matches no trip in the observation window. The three are filtered together for completeness. For stations with sporadic coordinate omissions that appear with valid coordinates elsewhere in the dataset, missing values are imputed using historical station lookups, ensuring zero loss of valid trip records due to coordinate formatting errors. Following this initial station-level filtering, 1,675,126 trips remain for outlier evaluation.

Bicycles that record trip activity but never appear in the servicing logs are explicitly retained in the dataset. The absence of a maintenance entry does not signify defective recording; rather, it

represents a valid right-censored observation wherein the component survived the entire monitoring window without experiencing a corrective intervention, adhering to the censoring framework defined in Subsection 3.3.1.

The outlier filtering rules are applied sequentially to reflect logical dependencies in trip characteristics. Implausibly short trips (under two minutes) are filtered first to remove accidental unlocks and immediate docking re-trials, followed by the removal of implausibly long trips (exceeding 150 minutes). The average speed distribution is then recalculated on the remaining subset, confirming that a 25 km/h upper limit falls between the 99.5th and 99.9th percentiles of the observed speed distribution. The speed threshold and the stuck-bicycle criterion are evaluated jointly in the final filtering step. Circular trips originating and ending at the same station are explicitly exempt from the stuck-bicycle rule, preserving legitimate recreational usage patterns that involve low net displacement over longer durations. After removing a cumulative total of 102,754 outliers (6.13% of the candidate feed), 1,572,372 cleaned trips are retained.

4.3.2 Maintenance Taxonomy and Harmonization

Raw maintenance logs present semantic heterogeneity because work orders are entered manually by workshop technicians across different operational teams. Identical mechanical failure modes and servicing actions are recorded using disparate part descriptions, shorthand abbreviations, and regional technical terms. Standardizing these raw entries into a structured, auditable taxonomy is a prerequisite for defining MO units and fitting component-level survival models.

The harmonization framework organizes servicing interventions into a two-level taxonomy encompassing six mechanical and electrical subsystems, with each subsystem divided into named components and a residual general category. Table 4.4 outlines this global classification structure. Cataloguing every component maintains a complete operational audit trail across the entire fleet regardless of modeling boundaries.

Table 4.4: Component taxonomy used to harmonize free-text maintenance entries into standardized mechanical and electrical subsystems.

Subsystem	**Standardized Components**
Braking	Brake Pads, Cables/Hydraulics, Calipers, Braking (General)
Wheels	Tires, Spokes, Rims, Hubs, Wheels (General)
Drivetrain	Chain, Cassette, Derailleur, Bottom Bracket, Pedals, Shifting, Drivetrain (General)
Structure	Frame, Kickstand, Seat, Handlebar, Fork, Docking Pin, Structure (General)
Electronics	Battery, Motor, GPS, RFID, Electronics (General)
Accessories & Aesthetics	Basket, Bell, Helmet, Stickers, Fender, Lock, Accessories (General)

In addition to component identification, each log entry is tagged with one of eight standardized maintenance actions, which dictate how the entry is processed within the survival modeling layer.

Table 4.5 details these standardized actions alongside their event classification. Interventions that resolve a functional failure or replace an exhausted component, specifically *Replace* and *Repair*, terminate the active observation interval as an event ($\delta = 1$). Non-corrective interventions such as *Adjust*, *Tighten*, *Lubricate*, *Clean*, *Inspect*, and *Preventive* do not reset component age or accumulated mileage; instead, they are tracked as minor maintenance counts within the active interval to serve as indicators of degradation intensity.

Table 4.5: Action taxonomy and the resulting MO unit classification.

Action	**Description**	**Classification**
Replace	Full substitution of the component at the end of its functional life	Event ($\delta = 1$)
Repair	Recovery of the component without full substitution	Event ($\delta = 1$)
Adjust	Calibration or tension adjustment of a mechanical system	Minor intervention
Tighten	Fastener retightening to restore structural integrity	Minor intervention
Lubricate	Application of lubricant to reduce friction and inhibit oxidation	Minor intervention
Clean	Removal of dirt, debris, or road salt to prevent corrosion	Minor intervention
Inspect	Diagnostic check or safety verification	Minor intervention
Preventive	Scheduled inspection triggered by elapsed time or distance	Minor intervention

To address the semantic heterogeneity of free-text workshop logs, the harmonization layer employs explicit, operator-specific mapping dictionaries tailored to the individual vocabulary and logging structure of each fleet. For the primary MetroBikeShare feed, a dedicated lookup dictionary comprising 430 unique part and action descriptors was constructed. This dictionary maps distinct raw description strings directly to standardized component/action pairs, expands multi-component work order rows into individual single-component records, inherits maintenance action tags from ticket-level metadata when item descriptions are non-specific, and applies dual raw and normalized lookup passes to absorb whitespace and typographical variants.

Through this operator-specific mapping architecture, unclassifiable records in the MetroBikeShare log were reduced from roughly 35,900 raw entries to 270, all of which represent legitimately non-component items such as shop tool servicing or deleted placeholder rows. A final manual audit resolved residual boundary cases.

4.3.3 Route Geometry Enrichment

Trip feeds generated by bike-sharing systems natively record origin and destination docking stations but omit intermediate route geometries and altitude profiles. Because straight-line distance between stations understates the mechanical work actually performed, reconstructing the travelled path is a prerequisite for any distance-based covariate. To reconstruct actual travel paths, the enrichment layer incorporates OpenRouteService (v9.9.0) [81], an open-source routing engine developed by

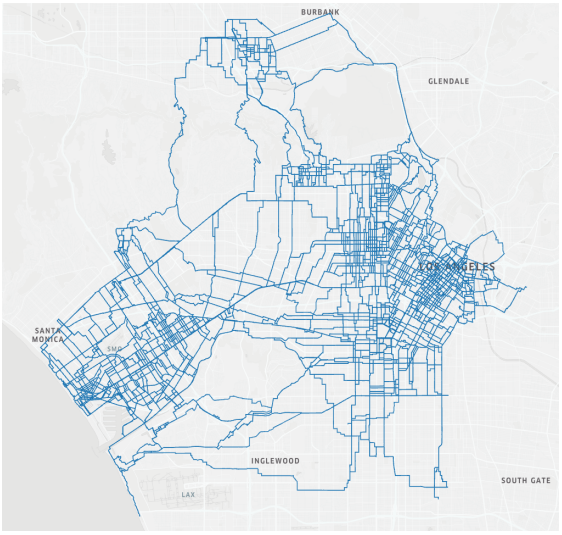
the Heidelberg Institute for Geoinformation Technology (HeiGIT) based on OpenStreetMap [82] road network topologies.

Evaluating the primary MetroBikeShare feed required generating route profiles for 27,509 unique non-circular origin-destination pairs. Relying on the public OpenRouteService API was unfeasible due to strict daily request caps. The routing engine was containerized using Docker and deployed on local infrastructure, bypassing external rate limits and unifying routing parameters across target fleets. Map extracts in Protocolbuffer Binary Format (PBF) were acquired from Geofabrik [83], covering Southern California.

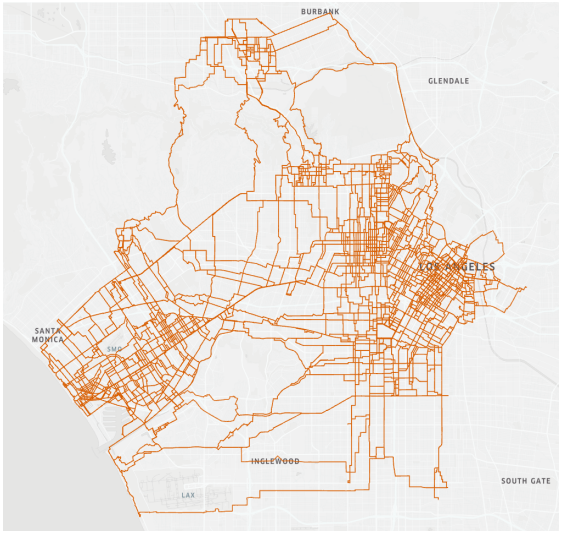
To account for potential routing divergences between propulsion technologies, two dedicated cycling profiles were configured within OpenRouteService: *cycling-regular* for m-bikes and *cycling-electric* for e-bikes. The electric routing profile accounts for motor assistance, prioritizing paths that maintain momentum and leverage dedicated cycling infrastructure. Evaluating a random sample of 2,000 origin-destination pairs under both profiles revealed that routing paths diverged in 17.0% of cases (339 pairs), producing a minor mean distance difference of -3.88 m ($SD = 73.04$ m, ranging from $-1,588.3$ m to $+643.4$ m). While this small global distance delta confirms that overall trip length remains largely stable across profiles, the adoption of dual-profile routing is driven by network coverage across the system. Across the complete MetroBikeShare network, 3,404 origin-destination pairs were traveled exclusively by m-bikes and 6,222 exclusively by e-bikes, while 17,883 pairs were shared by both propulsion types. Because nearly 35% of all active origin-destination pairs are specific to one vehicle class, enforcing propulsion-matched routing profiles ensures that spatial geometries reflect the distinct sub-networks utilized by each fleet component.

Route geometries returned by OpenRouteService for each origin-destination pair were exported as `.geojson` files, which were subsequently loaded into Kepler.gl for spatial rendering. The calculated origin-destination route network across the Los Angeles metropolitan area, visualized in Figure 4.3, illustrates this spatial distribution across vehicle types. By evaluating mechanical and electric trips under dedicated network graphs, the enrichment pipeline captures spatial coverage variations along transit corridors and localized road segments without imposing unobserved geometries onto vehicle-specific journeys. This dual-profile configuration enables the extraction of route distances and elevation profiles, which are subsequently aggregated into the daily feature space described in Subsection 4.3.6. However, these elevation-derived covariates were ultimately excluded following the data leakage audit detailed in Subsection 4.4.2.

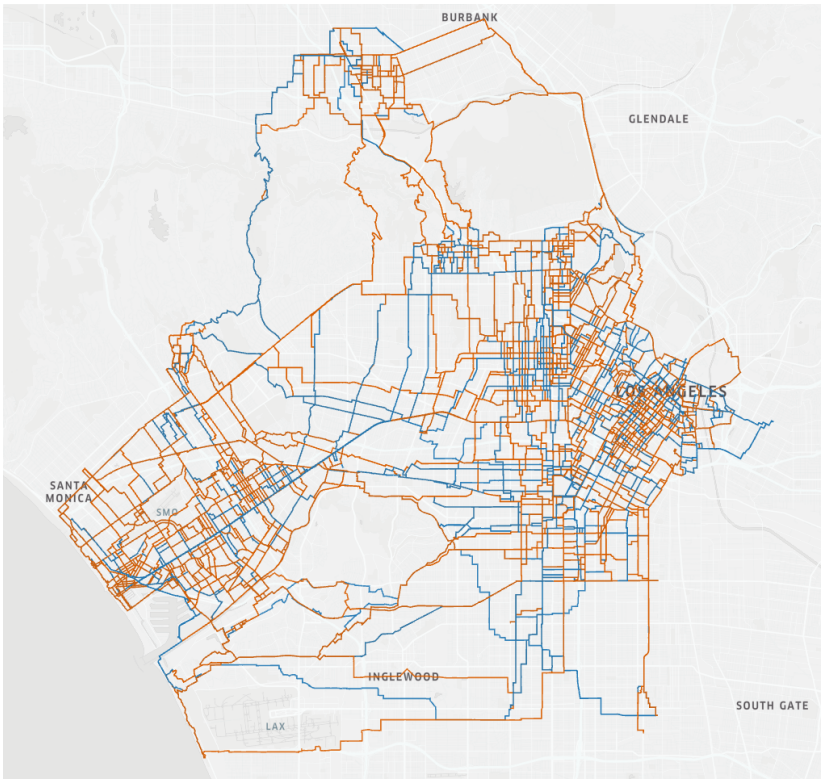
Circular trips, where origin and destination stations coincide (comprising 318,573 trips, or 19.0% of the MetroBikeShare feed), cannot yield a deterministic spatial route. Rather than discarding these valid rentals, circular trips are retained across all usage statistics. Their travel distances are imputed using the product of trip duration and the average operating speed of non-circular trips within the fleet, while elevation gain and loss are set to zero and tracked through a dedicated `n_round_trips` feature during the daily dataset assembly described in Subsection 4.3.6. This approach preserves legitimate rental volume without introducing artificial topographic stress into the feature space.



(a) cycling-regular profile



(b) cycling-electric profile



(c) Combined profile overlay showing spatial divergence

Figure 4.3: Spatial distribution of calculated origin-destination route geometries across the Los Angeles metropolitan area, visualized via Kepler.gl.

4.3.4 Weather Enrichment

Environmental conditions such as ambient temperature fluctuations, relative humidity, and rainfall act as secondary degradation drivers, accelerating metal oxidation, lubricant breakdown, and electronic moisture ingress. To capture these external stressors, historical daily weather records

were retrieved using the Open-Meteo [84] archive service. Open-Meteo consolidates observational data and global meteorological reanalysis models, providing spatial retrieval grounded on exact geographic coordinates without daily request constraints. For the primary MetroBikeShare fleet, query parameters were mapped directly to the spatial centroid of central Los Angeles, matching the longitudinal temporal window of the dataset from July 2021 to November 2025. Because Los Angeles accounts for over 80% of all recorded fleet trips and neighboring operational areas such as Santa Monica share immediate geographic proximity, microclimatic differences across the service area are negligible on a daily scale, validating centroid-based meteorological sampling as representative of the entire system.

The raw meteorological stream ingested daily metrics covering minimum, maximum, and mean ambient temperatures, total precipitation, mean wind speed and direction, relative humidity, cloud cover, solar radiation, and surface atmospheric pressure. Data validation procedures addressed structural sparsity, such as filling missing precipitation entries with zero to reflect dry operational days, while ensuring complete alignment along the daily calendar index. The resulting weather matrix was merged directly into the daily master dataset by date. Because several raw environmental attributes share underlying seasonal trends and exhibit substantial mutual collinearity, this initial set is subsequently filtered using the statistical screening procedures described in Subsection 4.3.5, retaining a parsimonious subset of environmental covariates for survival model training.

4.3.5 Feature Screening

Combining aggregated daily bicycle usage indicators with local meteorological variables yields a high-dimensional candidate feature set. To prevent multicollinearity from compromising parameter estimation and hazard ratio interpretation in downstream survival models, candidate variables from both data sources were screened independently using Spearman rank correlation (ρ). This non-parametric metric was chosen due to its robustness against non-linear relationships and the heavy-tailed, skewed distributions inherent to mobility usage records.

The correlation structure for candidate daily bicycle usage attributes is illustrated in Figure 4.4. Feature pairs exhibiting absolute correlation coefficients exceeding $|\rho| > 0.85$ were categorized as redundant, leading to the retention of a single representative attribute selected on grounds of physical interpretability and mechanical relevance. The identified redundancies directly reflected the mathematical definitions of the aggregated metrics. Specifically, upper speed percentiles tracked daily mean speed ($\rho = 0.95$ between `speed_p75` and `speed_mean`), maximum trip duration correlated strongly with mean active duration ($\rho = 0.96$), maximum single-trip distance mirrored mean trip distance ($\rho = 0.94$), and daily elevation loss strictly tracked daily elevation gain ($\rho = 0.94$) due to closed urban commuting paths. Attributes lacking direct mechanical links to component fatigue, such as the timestamp of the first trip of the day, were excluded on domain grounds.

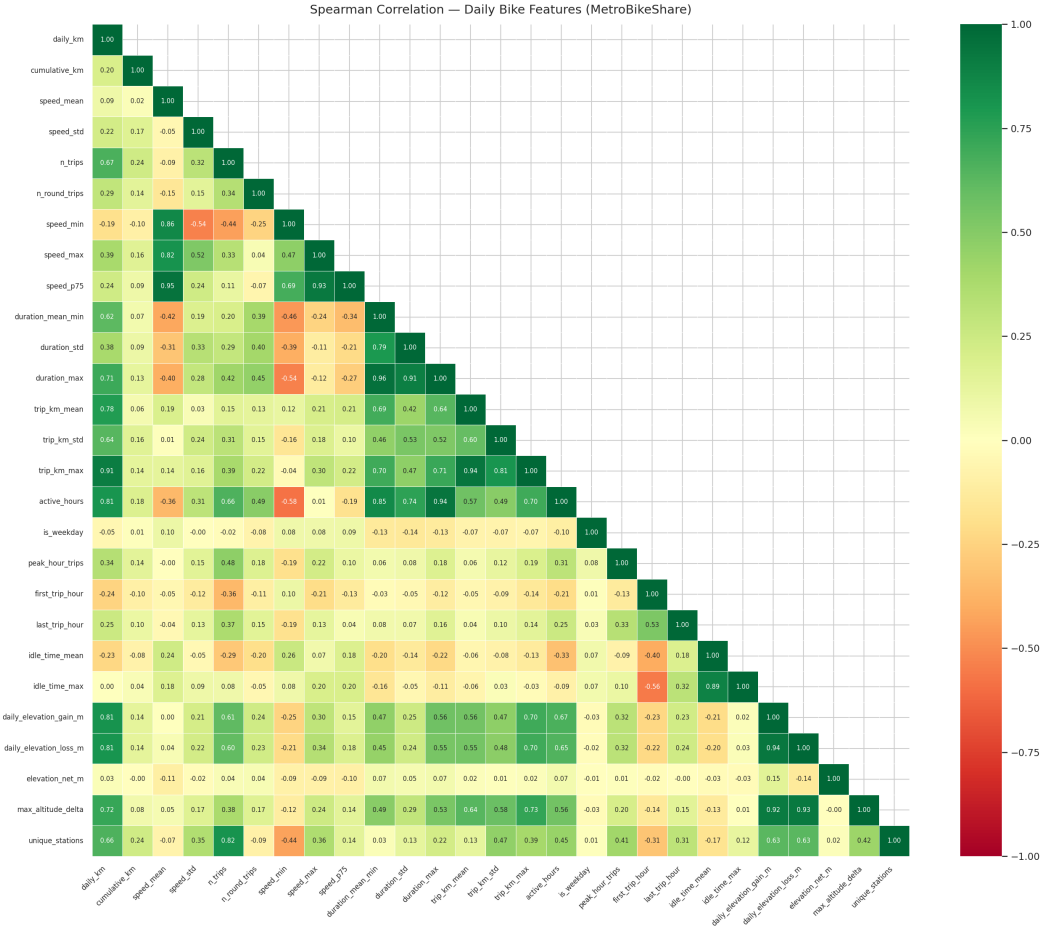


Figure 4.4: Spearman rank correlation matrix of candidate daily bicycle features.

Meteorological attributes required a dual screening framework incorporating both redundancy and dispersion metrics, as shown in Figure 4.5 and Figure 4.6. Correlation analysis collapsed the weather archive into collinear variable clusters: atmospheric pressure variants were mutually redundant ($\rho = 0.999$ between surface and sea-level pressure), ambient temperature metrics formed an internally collinear block ($\rho \approx 0.98$), and sunshine duration strongly correlated with solar radiation ($\rho = 0.88$). One representative attribute was selected from each collinear group.

To evaluate whether uncorrelated environmental features provided sufficient variance for statistical learning, the Coefficient of Variation (CV) and zero-value proportions were evaluated, as depicted in Figure 4.6. In Los Angeles, atmospheric pressure demonstrated near-zero dispersion (CV $\approx 0.1\%$), establishing that barometric pressure remains virtually static in this climate and offers negligible predictive contrast. Daily precipitation exhibited a high coefficient of variation (486%) alongside zero values on 82% of recorded days, capturing a sparse but mechanically significant degradation stressor.

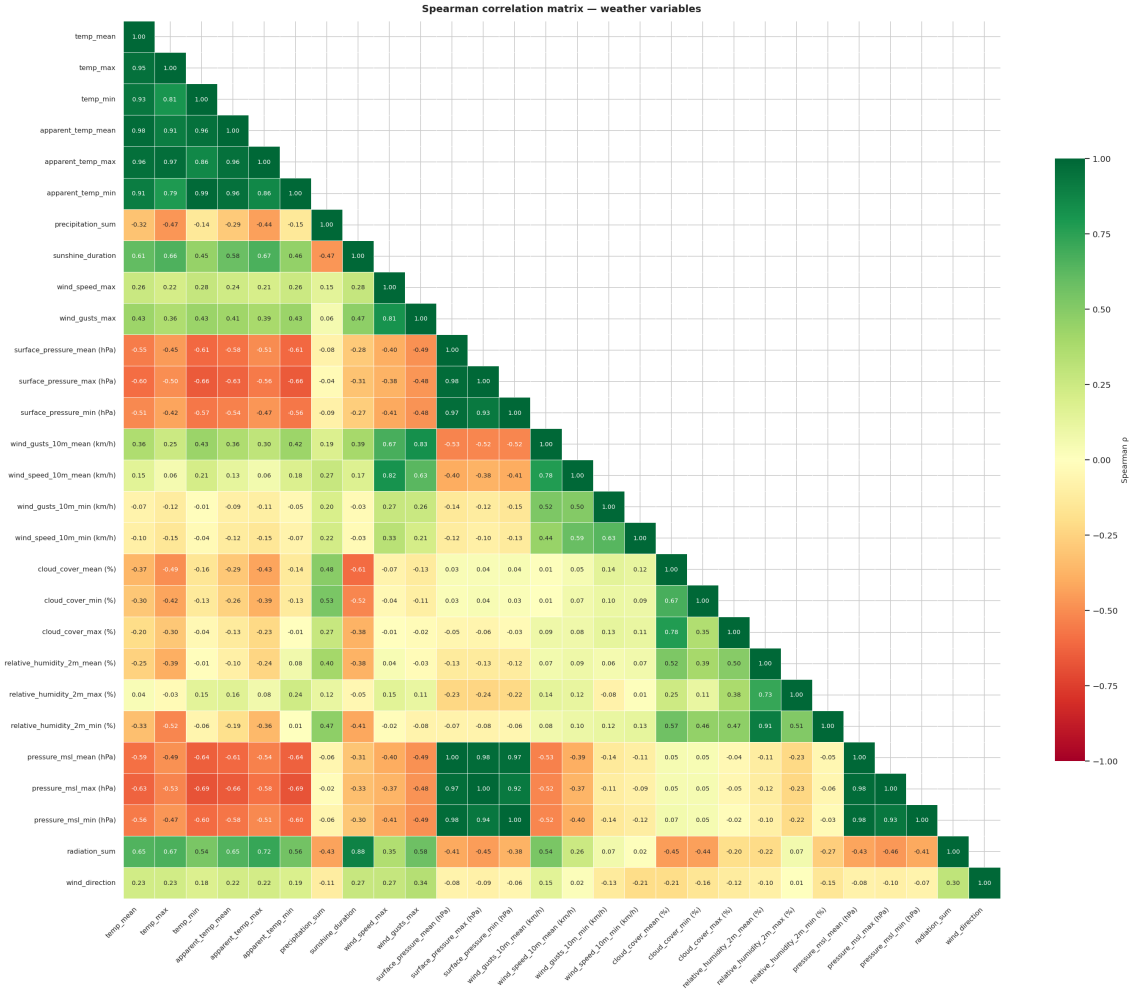


Figure 4.5: Spearman rank correlation matrix of historical weather variables extracted from Open-Meteo.

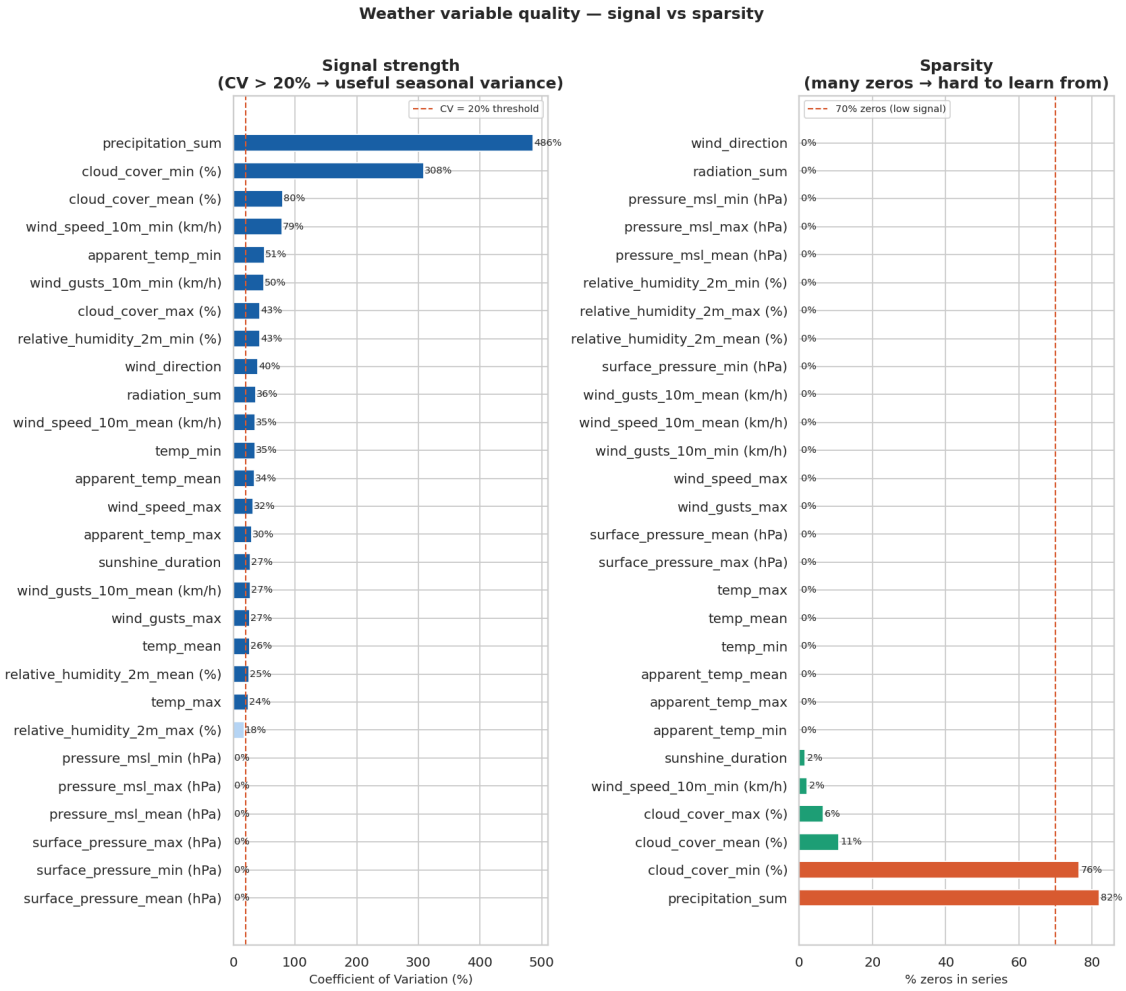


Figure 4.6: Meteorological variable screening criteria. Left: coefficient of variation assessing seasonal dispersion. Right: proportion of zero values highlighting variable sparsity.

The statistical screening process reduced the high-dimensional candidate pool to a parsimonious feature space optimized for time-to-event modeling, organized into three primary functional groups. This audited variable subset directly feeds the daily master dataset and subsequent MO unit assembly detailed in Subsection 4.3.6.

Usage and work indicators. Retains total daily trip count (`n_trips`), count of circular trips where origin equals destination (`n_round_trips`), total distance traveled in the day (`daily_km`), and lifetime accumulated distance (`cumulative_km`). These features capture the primary mechanical work and load accumulation on structural components.

Speed and operational intensity. Incorporates mean operating speed (`speed_mean`), standard deviation of operating speed (`speed_std`), average distance per trip (`trip_km_mean`), total active riding duration (`active_hours`), and mean idle duration between consecutive trips (`idle_time_mean`). These metrics represent dynamics that drive friction, thermal stress, and peak mechanical loads.

Environmental covariates. Encompasses mean ambient temperature (`temp_mean`) and total daily precipitation (`precip_sum`), alongside mean surface atmospheric pressure (`pressure_mean`) deliberately retained as a negative control variable to verify statistical baseline behavior.

4.3.6 Daily Dataset Assembly

The data processing pipeline converges into a daily master dataset structured at a temporal granularity of one record per unique bicycle per calendar day. For each asset, the timeline commences on the date of its initial recorded activity in either trip or workshop logs and continues sequentially through the final observation date. Non-operational days are explicitly retained with usage metrics set to zero (daily_km $= 0$, n_trips $= 0$), preserving temporal continuity since material degradation and environmental exposure persist during inactive periods.

Along this continuous daily timeline, three component-specific tracking variables are updated for each mechanical subsystem c in the standardized taxonomy:

1. `maint_c`: An ordinal severity indicator reflecting the highest intervention level recorded on that date (0 for no intervention, 1 for inspection, cleaning, or lubrication, 2 for mechanical adjustments, 3 for component repair, and 4 for complete replacement).

2. `km_accum_c`: Cumulative distance in kilometers accumulated since the most recent severe corrective intervention (`maint_c` ≥ 3). This counter accumulates daily distance and resets to zero upon a repair or replacement event, whereas minor adjustments (`maint_c` ≤ 2) do not trigger a reset.

3. `riding_time_accum_c`: Cumulative active operation time in hours accumulated since the last severe corrective event, following an identical reset structure.

Extended non-operational intervals are managed by enforcing a maximum idle threshold of 30 days. Inactivity exceeding this threshold triggers a general servicing reset, restoring accumulated stress counters to zero and re-establishing the asset state prior to returning to active service.

To prepare this master timeline for failure modeling, daily records are transformed into MO units. To eliminate parameter bias arising from unobserved initial wear states, all days preceding the first recorded workshop intervention for a given component are flagged as left-censored (`is_left_censored = True`) and excluded from survival model training. Active intervals concluding with a corrective repair or replacement (`maint_c` ≥ 3) are classified as observed failure events ($\delta = 1$), whereas intervals surviving at the conclusion of data collection are categorized as right-censored ($\delta = 0$).

4.4 MO Unit Construction and Feature Audit

The conversion of continuous daily fleet records into discrete time-to-event observations constitutes the core link between data harmonization and statistical survival modeling. While the daily dataset assembled in Subsection 4.3.6 tracks asset state sequentially, survival architectures require structured intervals bounded by operational resets or censoring thresholds. This section presents the automated construction of MO units, detailing the interval builder logic, the schema of the derived records, and the systematic feature leakage audit performed to ensure temporal validity before model training.

4.4.1 Builder and Resulting Structure

A fleet-agnostic builder module transforms the continuous daily master dataset into component-specific observation records. Following the formulation of Subsection 3.3.1, the builder identifies terminal events based on the corrective action classifications formalized in Subsection 4.3.2, setting the failure indicator $\delta = 1$ for intervals concluding with a repair or replacement, while assigning $\delta = 0$ to right-censored intervals active at the conclusion of data collection.

To eliminate parameter estimation bias resulting from unknown baseline wear, left-censoring is programmatically enforced during interval construction. The builder flags the initial interval of each bicycle-component pair as left-censored (`is_left_censored = True`) whenever asset trip logs predate the opening of workshop records, as detailed in Subsection 4.3.6. These incomplete initial units are filtered out prior to survival model training, whereas assets introduced after maintenance logging commenced retain fully observed initial intervals.

For each valid MO record, daily usage and meteorological variables are aggregated across the interval bounds using descriptive summary statistics, specifically the mean, standard deviation, and ordinary least squares linear trend. In accordance with the taxonomy established in Subsection 4.3.2, minor interventions are accumulated into dedicated count features, providing secondary degradation signals without triggering an interval reset. Cross-component maintenance counts are appended to allow single-component models to observe the broader servicing history of the asset. Table 4.6 outlines the structural schema of the resulting MO unit records.

To guarantee adequate statistical power during model training, the builder enforces an empirical sample size requirement of at least ten observed failure events ($\delta = 1$) per component. Across the MetroBikeShare fleet, seven mechanical and electrical subsystems satisfy this condition: brakes,

saddle, wheels, tires, chain, cables, and battery. The battery is excluded from the mechanical framework for the reasons established in Section 3.2.

Table 4.6: Structural schema of an MO unit record generated by the automated builder pipeline.

Field Group	**Attributes and Content**
Identity	Bicycle identifier, vehicle propulsion type, and target component
Interval Boundaries	Start date, end date, and total active duration in days
Survival Target	Survival duration (T), event indicator (δ), and closing action type
Censoring Flags	Left-censoring indicator (`is_left_censored`)
Usage Covariates	Interval mean, standard deviation, and linear trend of daily distance, speed, trip counts, active duration, and idle time
Exposure Covariates	Interval mean, standard deviation, and trend of ambient temperature, precipitation, and surface pressure
Maintenance Covariates	Minor intervention counts (`maint_count_minor`) and cross-component corrective maintenance counts

4.4.2 Feature Configurations and Leakage Audit

Prior to survival model training, a data leakage audit was conducted across all constructed MO units by evaluating the bivariate correlation structure between candidate covariates and the target survival duration (T). This audit identified a critical vulnerability: interval-total usage indicators, such as total accumulated distance, total active hours, total trip count, and total cumulative elevation gain over the interval, exhibit deterministic linear correlation ($\rho = 1.0$) with interval duration. When provided with total interval distance, a baseline CPH model achieved a concordance index of 1.0, signaling complete data leakage rather than genuine predictive learning. All cumulative interval totals were systematically purged from the feature space, replacing them with daily rate metrics, dispersion parameters, and temporal trend coefficients.

A secondary source of data leakage was identified in cross-component cumulative usage counters. Cumulative distance accumulated on auxiliary components over the target interval similarly leaked interval duration. These variables were removed and replaced strictly with discrete minor and major maintenance event counts recorded on other components prior to or during the interval.

Following the leakage audit, feature selection was refined through drop-column sensitivity analysis across each mechanical subsystem, evaluating model performance variations when candidate covariates were systematically removed from the feature set. The set of drop-column feature importance plots for all six modeled components is documented in Appendix A. Model discrimination consistently improved when pruning redundant environmental and unstable usage metrics. The final feature configuration adopted for all primary modeling experiments combines the audited lean feature set with the distance accumulated within the interval, a retrospective covariate discussed in Section 5.6, ensuring discrimination while maintaining strict temporal integrity across the predictive pipeline.

4.5 Survival Model Training and Evaluation

For each of the six mechanical components in the harmonized taxonomy, four distinct survival formulations are fitted using the audited MetroBikeShare MO units. Population-level baselines are established through non-parametric Kaplan-Meier estimation using the *lifelines* package and parametric Weibull distributions fitted via censoring-aware maximum likelihood, where the fitted shape parameter (β_c) provides physical insight into the underlying wear regime of each mechanical assembly as detailed in Subsection 2.3.2. Covariate-conditioned modeling incorporates both semi-parametric CPH implemented via *lifelines* and non-linear GBSA implemented in *scikit-survival*. GBSA was selected due to its superior empirical performance in predictive maintenance benchmarks and to maintain methodological alignment with the boosting architectures employed in previous research [38]. Hyperparameters for both conditioned model families are optimized independently for each component using Optuna [73] with Tree-structured Parzen Estimator (TPE) sampling across five-fold cross-validation on the training set, with folds grouped by `bike_id` so that no bicycle is split across folds, incorporating early stopping on held-out fold fractions for gradient boosting fits.

Model training and evaluation strictly adhere to the fleet-level partitioning scheme established in Section 3.6. An 80/20 dataset split is enforced at the bicycle level (`bike_id`), ensuring that every observation interval generated by a given asset is assigned exclusively to either the training or test partition to prevent cross-interval data leakage. For example, in the braking subsystem dataset, 1,216 bicycles contribute 1,360 training units, while 305 held-out bicycles account for the 336 observation units in the test partition. Predictive discrimination is evaluated using the C-IPCW from *scikit-survival*, which corrects for censoring bias inherent to longitudinal fleet data. Predictive calibration is assessed using the IBS, integrated up to τ, fixed per component at the 95th percentile of the observed event times in the training partition. Predicted survival curves are evaluated on a grid bounded by that horizon. The resulting evaluation pipeline outputs component-specific CPH hazard ratios, GBSA permutation feature importance ranks, discrimination indices, and calibration scores on the held-out test partition.

4.6 Health Index Computation

The fleet monitoring layer is structured as a sequential processing chain comprising four modular scripts that compute component-level indices, aggregate component health into asset-level trajectories, derive empirical failure thresholds, and serialize data payloads for web visualization. Both Health Index formulations formalized in Subsection 3.3.3 are computed concurrently. The deterministic baseline evaluates Weibull survival probabilities at seven-day steps along each active MO unit, based on the distance accumulated since the most recent corrective intervention, utilizing shape and scale parameters fitted exclusively on the training partition. The model-based HI evaluates the conditional survival function ($S(t \mid \mathbf{x})$) on the same seven-day grid, using the trained CPH and GBSA models under the lean feature configuration fixed in Subsection 4.4.2.

Component-level index aggregation into a unified bicycle-level Health Index exposed a critical alignment vulnerability during initial development. Because different mechanical components undergo maintenance interventions on distinct calendar dates, raw component health series exist on asynchronous date grids. Naive averaging across misaligned series created artificial, high-frequency trajectory oscillations that distorted asset health interpretation. To resolve this issue, the aggregation module resamples every component series onto a common seven-day calendar grid per bicycle, forward-fills component health values across inactive operational intervals, and computes the weighted aggregate asset index. To preserve the physical meaning of servicing interventions, segment-aware smoothing is enforced, preventing moving averages from blending health values across a corrective maintenance reset. This ensures that post-repair health recoveries appear as sharp, unambiguous step-function discontinuities.

Continuous Health Index trajectories are discretized into four operational condition bands: *Healthy, Attention, Alert, and Critical*. Rather than establishing arbitrary qualitative boundaries, classification thresholds are derived empirically from fleet behavior. The three boundary values correspond to the 25th, 50th, and 75th percentiles of the Health Index distribution observed at the exact moment of 2,720 verified corrective interventions in the test set. These boundaries encode the specific risk values at which components historically failed within the operational environment. A bicycle transitions into the Critical band when its predicted health drops below the threshold at which one-quarter of all real-world failures occurred, while the Alert boundary marks the median asset health on failure days. The resulting boundary values for the three index formulations are reported in Table 5.4.

4.7 Monitoring Dashboard Generation

To demonstrate the practical applicability of the proposed framework and illustrate the temporal evolution of asset health, a proof-of-concept interactive monitoring dashboard was developed as a lightweight web prototype. An automated generation script serializes the derived Health Index time series into optimized JSON payloads, embedding them directly into a single, standalone HTML file constructed with vanilla JavaScript, CSS3, and Chart.js. By operating strictly client-side without external web server dependencies, runtime database queries, or active Python environments, this prototype provides an accessible, self-contained visualization artifact that allows fleet operators and researchers to inspect health trajectories and evaluate the operational feasibility of survival-based monitoring interfaces.

The prototype user interface implements two complementary analytical views designed to support decision-making in fleet health monitoring. The Bike Timeline module renders historical component indices and aggregate bicycle health trajectories for any selected asset, explicitly displaying past maintenance interventions as instant recovery step-functions preserved by the segment-aware smoothing rules formalized in Section 4.6. Complementing this individual asset view, the Fleet Grid module presents a macroscopic snapshot of the entire bicycle fleet as a matrix of color-coded cells corresponding to operational health categories on a selected date. Interactive

temporal sliders and animation playback controls enable users to observe fleet degradation dynamics across continuous operational windows and navigate directly to specific bicycle timelines. All interaction logic is executed directly over the embedded JSON data structures, ensuring that the visualization layer remains an exact, transformation-free projection of upstream survival model outputs as required by the validation protocol established in Section 3.6.

With the complete data processing, survival modeling, Health Index derivation, and prototype visualization pipeline implemented end to end, Chapter 5 presents the formal experimental validation and comparative performance evaluation of the overall framework.

Chapter 5

Results and Discussion

Empirical validation of the predictive maintenance framework leverages longitudinal operational and workshop records from the MetroBikeShare fleet. Evaluating the proposed framework begins with an audit of data pipeline integrity and baseline component failure regimes before comparing conditioned survival architectures against unconditioned statistical benchmarks. Evaluating the resulting Health Index dynamics across individual asset lifecycles and fleet-wide aggregations demonstrates the practical efficacy of the proposed predictive monitoring system.

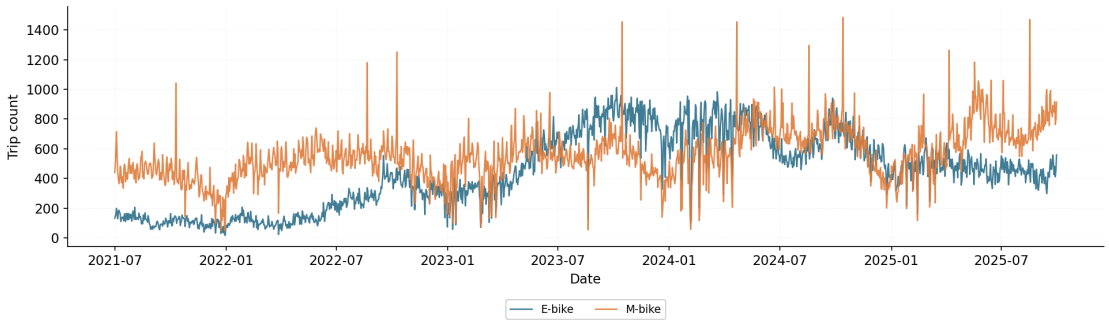
5.1 Fleet Profile and Operational Overview

Understanding the operational environment and failure characteristics of the MetroBikeShare fleet is necessary for contextualizing model performance. The primary dataset combines public trip records spanning from July 2021 to September 2025 with internal workshop maintenance logs recorded between January 2022 and November 2025, encompassing 2,637 unique bicycles. The fleet comprises 516 pedal-assist electric bicycles (19.6%) and 2,121 standard mechanical bicycles (80.4%).

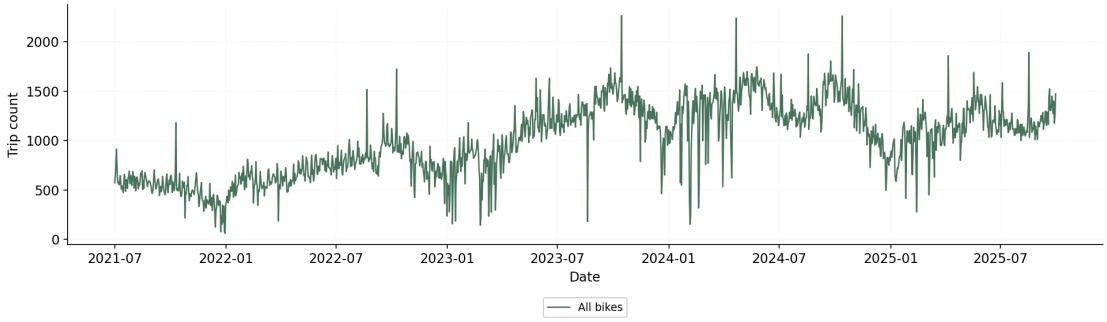
5.1.1 Fleet Utilization and Temporal Dynamics

Trip generation across the MetroBikeShare fleet exhibits pronounced seasonal, weekly, and intra-daily patterns over the multi-year observation window. As shown in Figure 5.1b, total daily trip counts follow a clear annual cycle, fluctuating between winter troughs of approximately 300 trips per day and summer demand peaks exceeding 2,200 daily trips. This macro-level seasonality reflects weather-driven mobility demand, while regular weekly fluctuations highlight distinct peak usage on weekdays driven by urban commuting patterns.

Comparing propulsion technologies across the fleet reveals a shift in composition (Figure 5.1a). E-bikes account for roughly one fifth of daily trips during the second half of 2021, overtake m-bikes in absolute daily volume between mid-2023 and early 2024, and settle near one third of total demand through 2025. Since e-bikes represent only 19.6% of the fleet, this share reflects a substantially higher utilization intensity per asset. This disproportionate demand translates to an accelerated duty



(a) Comparison between E-bike and M-bike daily trips



(b) Total daily trips (All bikes)

Figure 5.1: Daily trip counts for the MetroBikeShare fleet: (a) breakdown by propulsion technology (E-bike vs. M-bike), and (b) combined total fleet trip volume.

cycle for electric assets, subjecting drivetrain, braking, and structural components to greater daily operational strain than their mechanical counterparts.

5.1.2 Maintenance Log Taxonomy and Workshop Workload

Harmonizing raw free-text workshop logs into a standardized taxonomy provides a quantitative overview of servicing demand across the fleet. Expanding multi-component work orders into individual single-component records, as detailed in Subsection 4.3.2, raises the 247,948 raw workshop logs to 296,002 harmonized maintenance entries. To highlight the primary operational drivers and eliminate low-frequency logging noise, Figure 5.2 presents the hardware categories accounting for more than 3.0% of these harmonized records. Together, these fourteen primary categories comprise 85.0% of the total recorded servicing volume, offering a clear view of workshop workload.

As illustrated in Figure 5.2, maintenance interventions span both electronic tracking peripherals and primary mechanical wear assemblies. While electronic modules such as GPS units and RFID receivers account for a notable portion of individual workshop entries, physical maintenance concentrates heavily on critical wear components, particularly braking systems, saddles, drivetrain chains, and tires. Secondary structural elements and auxiliary wiring represent a moderate servicing volume, whereas minor repairs below the 3.0% threshold form a long tail of peripheral interventions.

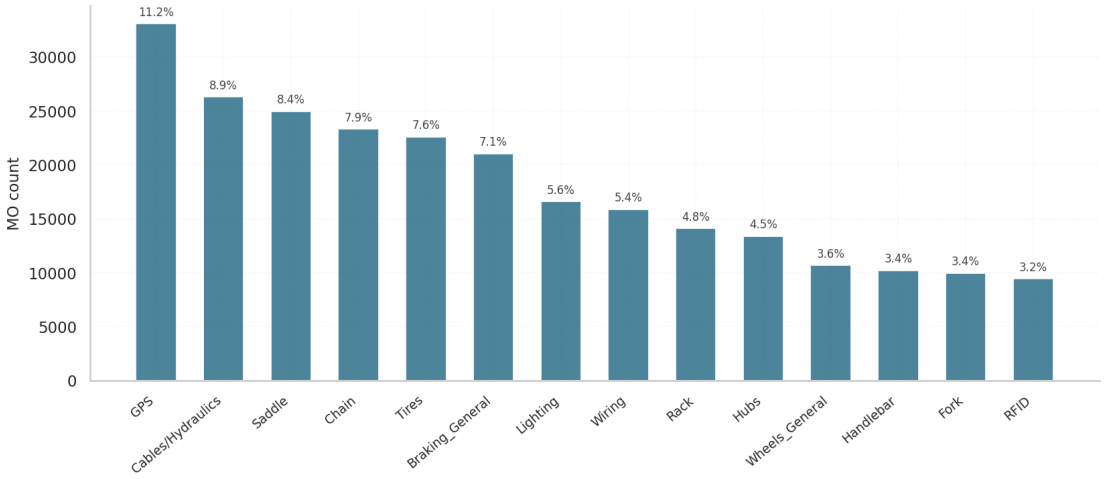


Figure 5.2: Proportion of maintenance operations by primary component categories (exceeding 3.0% representation) across the harmonized MetroBikeShare servicing log.

Although electronic peripherals generate frequent service logs, the probabilistic survival modeling evaluated in subsequent Subsections focuses specifically on primary mechanical wear components, whose physical degradation directly impacts asset availability and user safety.

In addition to component-level variations, the servicing burden is distributed unevenly across individual assets. Figure 5.3 displays the distribution of distinct maintenance days accumulated per bicycle over the four-year observation window.

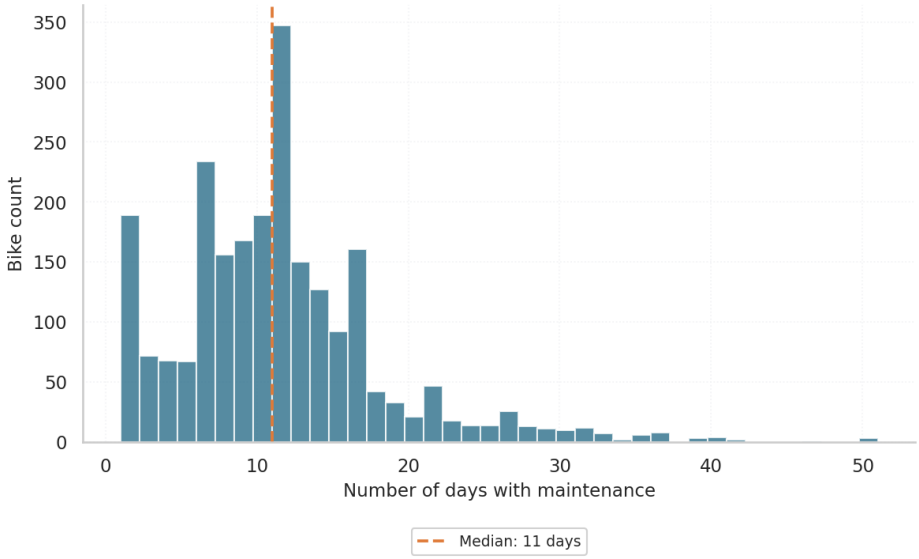
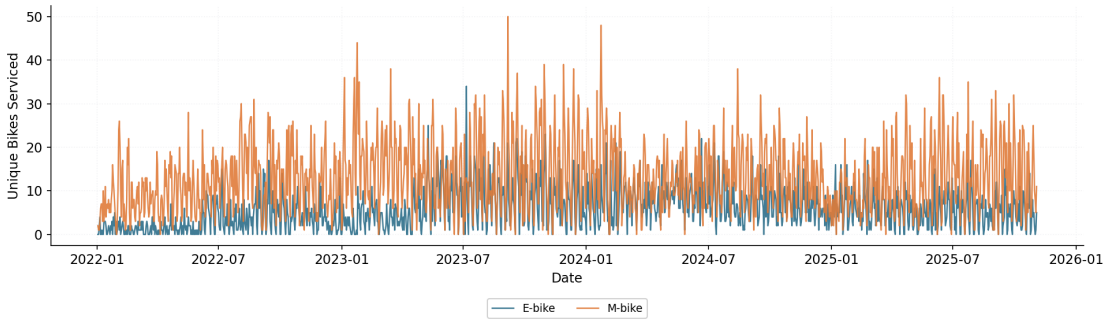


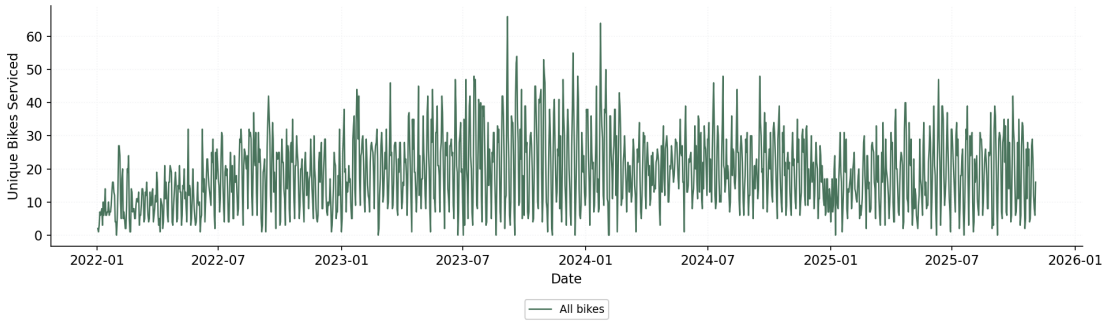
Figure 5.3: Distribution of maintenance days per bicycle across the MetroBikeShare fleet.

The maintenance day distribution exhibits strong right-skewness, with a fleet-wide median of 11 servicing days per bicycle. While the majority of assets require between 5 and 20 servicing days throughout their operational lifecycle, a subset of heavily utilized or failure-prone bicycles accumulates more than 30 servicing days. Despite these asset-level discrepancies, daily workshop

demand remains relatively consistent over time, as depicted in Figure 5.4b. Between 10 and 40 unique bicycles enter workshops on a typical day, with peaks above 60, demonstrating a continuous operational flow that can benefit directly from predictive scheduling to optimize technician allocation and inventory management. As shown in Figure 5.4a, m-bikes account for more daily servicing events than e-bikes across most of the observation window. This ordering follows from fleet composition rather than from reliability, since m-bikes outnumber e-bikes by roughly four to one. Per asset, the electric subfleet generates the heavier servicing load.



(a) Comparison between E-bike and M-bike daily bicycles serviced



(b) Total daily bicycles serviced (All bikes)

Figure 5.4: Daily count of unique bicycles undergoing workshop servicing: (a) breakdown by propulsion technology (E-bike vs. M-bike), and (b) combined total fleet servicing volume.

5.1.3 Censoring Structure and Usage Profiles

To support time-to-event modeling, raw maintenance dates are transformed into MO units representing active operational windows between consecutive corrective interventions on a given asset and component. Table 5.1 summarizes the resulting MO unit counts, observed corrective failure events ($\delta = 1$), right-censoring rates, and vehicle split across candidate components.

As visualized in Figure 5.5, the proportion of uncensored failure events versus right-censored observation intervals varies markedly across mechanical subsystems.

Consumable components subject to frequent replacements, such as saddles (71.4% uncensored) and tires (65.1%), accumulate high event proportions, yielding dense failure labels for statistical learning. Braking assemblies display a low uncensored event rate of 10.7%, with 89.3% of

Table 5.1: MO unit population, observed corrective failure events ($\delta = 1$), right-censoring proportions, and vehicle-type split per component on the MetroBikeShare fleet. Vehicle-type columns report units / events.

Component	MO Units	Events	Censoring (%)	Electric	Standard
Brakes	1,696	181	89.3	578 / 112	1,118 / 69
Saddle	7,341	5,239	28.6	1,785 / 1,318	5,556 / 3,921
Wheels	3,153	1,249	60.4	1,039 / 572	2,114 / 677
Tires	5,826	3,793	34.9	1,471 / 1,007	4,355 / 2,786
Chain	3,797	2,113	44.4	1,361 / 909	2,436 / 1,204
Cables/Hydraulics	3,296	1,313	60.2	857 / 393	2,439 / 920

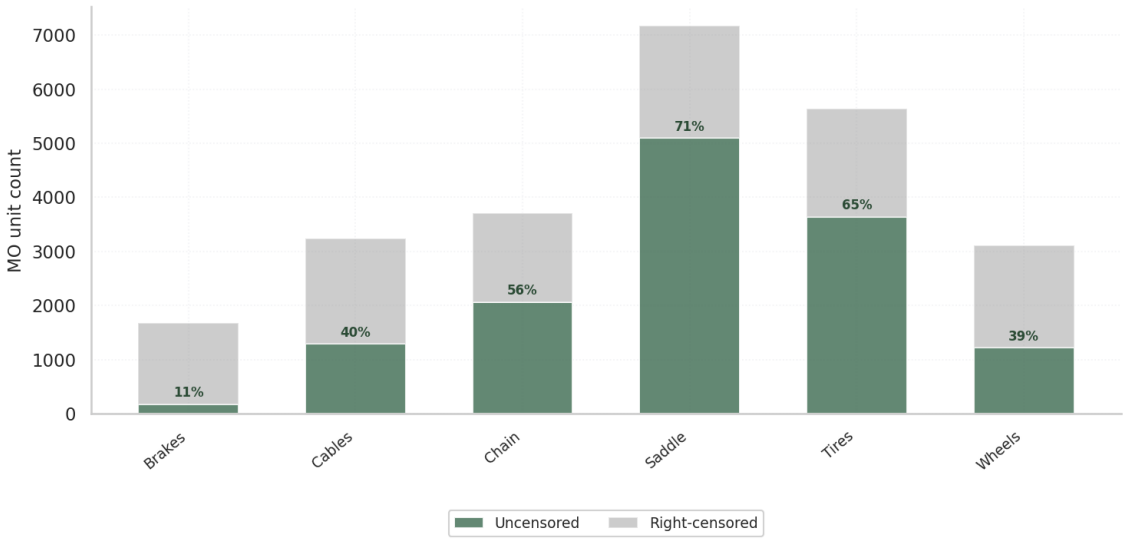


Figure 5.5: Distribution of uncensored ($\delta = 1$) and right-censored ($\delta = 0$) MO units across candidate subsystems.

observation intervals remaining right-censored. This pronounced label sparsity stems directly from the left-censoring filtration rule detailed in Subsection 4.4.1, where discarding the initial unobserved interval of each bicycle-component pair reduces the braking dataset from 3,420 to 1,696 intervals and removes 625 of the 806 recorded brake events, raising the censoring rate from 76.4% to 89.3%.

Across the six modeled subsystems, the corrective interventions that close an MO unit are almost entirely replacements: `Repair` (severity 3) accounts for no events in brakes, chain, tires, cables and wheels, and the saddle is the only component with a mixed profile (4,356 repairs against 2,973 replacements). The severity-3 event threshold defined in Subsection 3.3.1 therefore operates, in practice, as a replacement threshold for five of the six components.

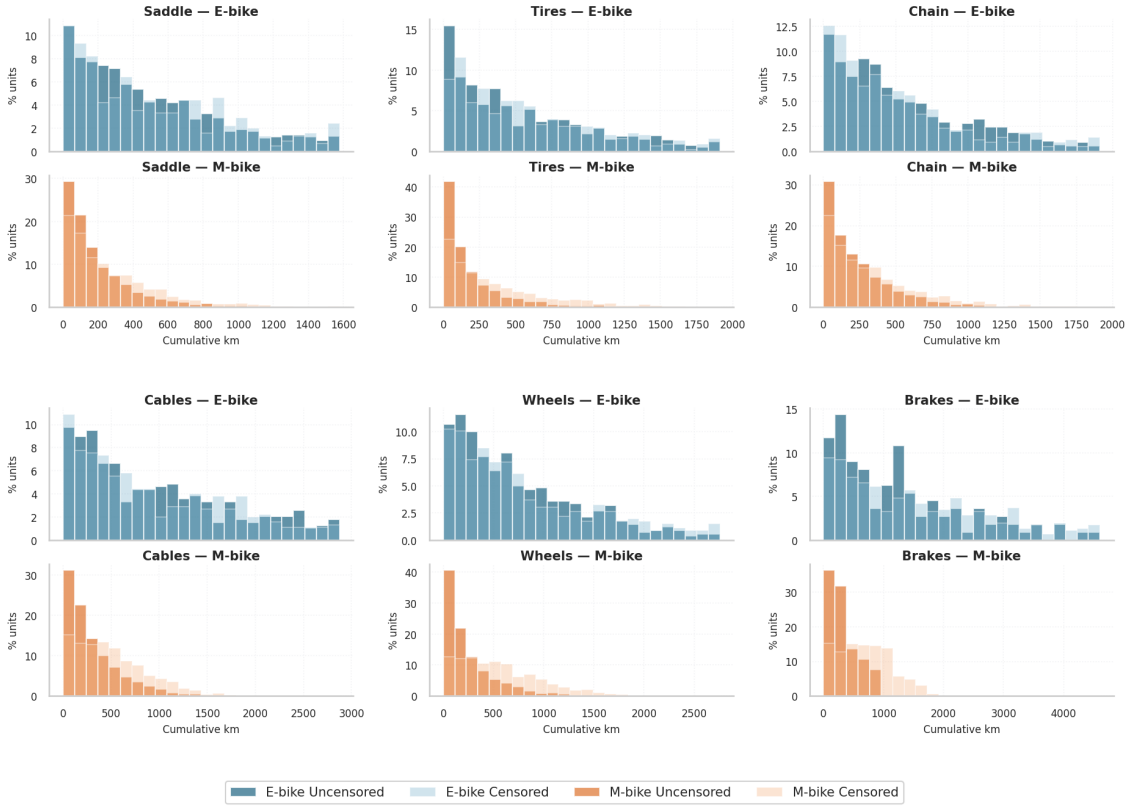


Figure 5.6: Distribution of cumulative kilometers per MO unit by component, stratified by uncensored and right-censored status across e-bikes and m-bikes.

Vehicle propulsion technology further introduces a clear differentiation in accumulated usage per MO unit. Figure 5.6 illustrates the empirical distributions of accumulated distance for uncensored and right-censored MO units, stratified by component type and propulsion class.

While m-bikes concentrate heavily in lower mileage ranges, e-bikes display broader, flatter distributions extending into higher distance bins. This contrast with the shorter calendar lifespans observed in Subsection 5.3.1 is consistent with higher daily usage rather than an accelerated wear rate per kilometer. Nevertheless, retaining vehicle propulsion class as an explanatory covariate remains justified because it captures these disparate duty cycles across the fleet.

5.2 Data Pipeline Quality

The first validation layer establishes the integrity and temporal validity of the data structures feeding the survival models. The cleaning filters defined in Subsection 4.3.1 were verified against empirical distributions, removing sensor logging noise and non-physical trip records. Similarly, the maintenance harmonization framework described in Subsection 4.3.2 achieved high coverage across workshop logs, reducing residual unclassifiable entries to a negligible set. As established in Table 5.1 (Subsection 5.1.3), the resulting MO population for the six audited mechanical subsystems totals 25,109 observation intervals.

Left-censoring filtration accounts for the lower observed event count in several components. Beyond the 625 brake events already noted, the same rule discards 2,090 saddle events and 1,407 chain events. This filtration step is essential to eliminate parameter estimation bias arising from unknown baseline wear, even though it reduces event density for long-lived components.

The censoring structure spans a wide range across subsystems, from 28.6% on saddles to 89.3% on braking assemblies. This structural variance empirically validates two methodological positions established in Chapter 3: model evaluation must be censoring-aware, which motivated adopting C-IPCW and IBS as primary evaluation metrics, and component viability for statistical learning is governed by observed failure events rather than total sample volume, given that saddles contribute nearly twenty-nine times as many failure events as brakes from only four times the sample size.

Data pipeline validation concluded with the verification of the data leakage audit executed in Subsection 4.4.2. Purging deterministic interval-total usage proxies and cross-component cumulative distance features restored mathematical independence between the feature space and target survival duration. Following these structural corrections, model concordance metrics transitioned from implausible near-perfect values to the realistic performance ranges reported in the following Subsections.

5.3 Survival Model Performance and Hypothesis Verification

5.3.1 Population-Level Behavior: Kaplan-Meier and Weibull

Non-parametric Kaplan-Meier survival curves and parametric Weibull distributions were fitted across the entire population of each mechanical component to establish baseline failure regimes. Figure 5.7 illustrates these population-level survival functions, while Table 5.2 summarizes the corresponding Weibull scale parameters (η), shape parameters (β), and observed median lifespans.

These population baselines delineate two distinct longevity profiles within the fleet. High-turnover consumables, including saddles, tires, and chains, exhibit short operational lifespans with median survival times between 232 and 280 days. In contrast, structural assemblies such as cables and wheels survive substantially longer, recording median lifespans of 764 and 790 days, respectively. Braking assemblies remain heavily right-censored, preventing a reliable population median estimation within the observation window. The corresponding scale parameter ($\eta = 10{,}070$ days, and 21,575 days when restricted to the m-bike stratum) lies more than an order of magnitude

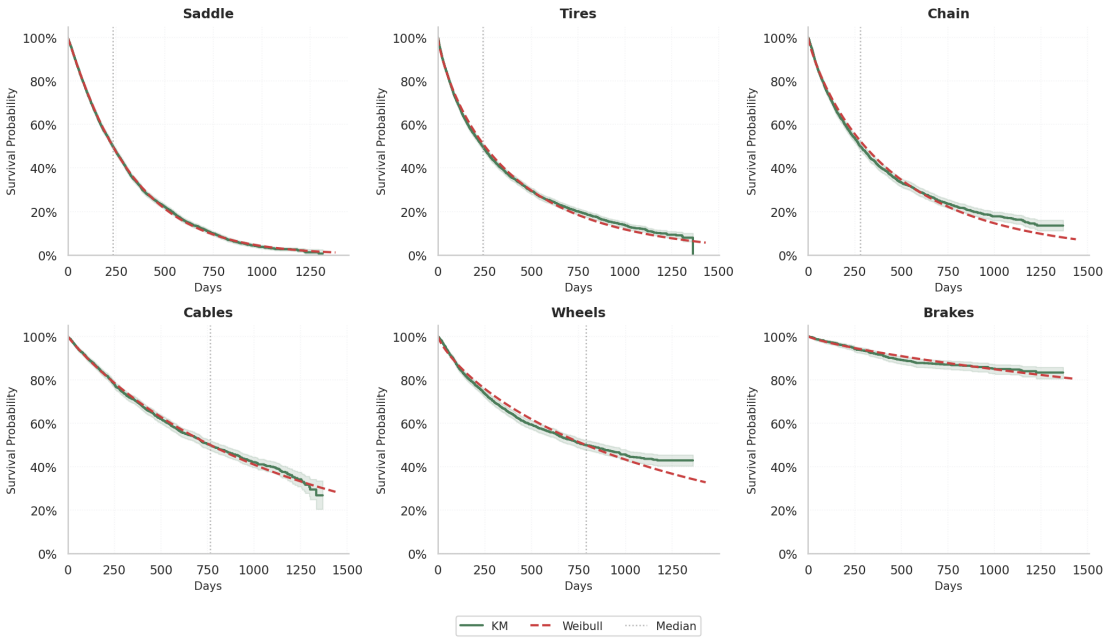


Figure 5.7: Kaplan-Meier survival curves with fitted Weibull overlays per component on the MetroBikeShare fleet. Vertical dotted lines mark median observed lifetimes where reached.

Table 5.2: Population-level Weibull distribution parameters (β shape, η scale in days) and observed median lifespan per component. The brake scale parameter is unidentified under 89.3% right-censoring and exceeds the observation window by more than an order of magnitude.

Component	**Shape (β)**	**Scale (η, days)**	**Median (days)**	**Aggregate Regime**
Brakes	0.781	10,070	—	Population frailty mix
Saddle	1.047	329	232	Approximately constant hazard
Wheels	0.799	1,248	790	Population frailty mix
Tires	0.817	393	239	Population frailty mix
Chain	0.859	466	280	Population frailty mix
Cables/Hydraulics	0.944	1,128	764	Population frailty mix

beyond the 1,587-day observation window and is not identified by the data. It is reported for completeness and should be read as an artifact of fitting a two-parameter distribution to a sample in which 89.3% of intervals are right-censored, not as a lifetime estimate.

For five of the six components, the fitted shape parameter falls below unity ($\beta < 1$), with the saddle subsystem standing as the sole exception ($\beta \approx 1.05$). A shape parameter below one classically implies an infant mortality regime characterized by a hazard rate that decreases over time [5], [6]. However, physical component degradation in shared bicycles is strictly wear-out driven. This apparent contradiction is consistent with unobserved frailty heterogeneity within the unconditioned fleet population [32], [37]. Because baseline models evaluate a mixed asset pool subject to different usage intensities, vulnerable or heavily stressed components fail early and are serviced [8]. The surviving subpopulation becomes increasingly composed of low-usage or inherently robust assets, creating an artificial decline in the aggregate population hazard rate. Unconditioned baselines fail to disentangle individual wear-out dynamics from this survivor selection bias.

Propulsion technology serves as a primary driver of this population heterogeneity. Stratifying Kaplan-Meier estimators by vehicle type confirms that e-bikes undergo systematically accelerated degradation compared to m-bikes. Fitted Weibull scale parameters for consumables are two to more than three times shorter on electric bikes: for drivetrain chains, $\eta \approx 201$ days on e-bikes compared to 661 days on m-bikes; for saddles, 179 days against 380 days; and for tires, 216 days against 460 days. Figure 5.8 shows the aggregated survival curves stratified by bicycle propulsion type.

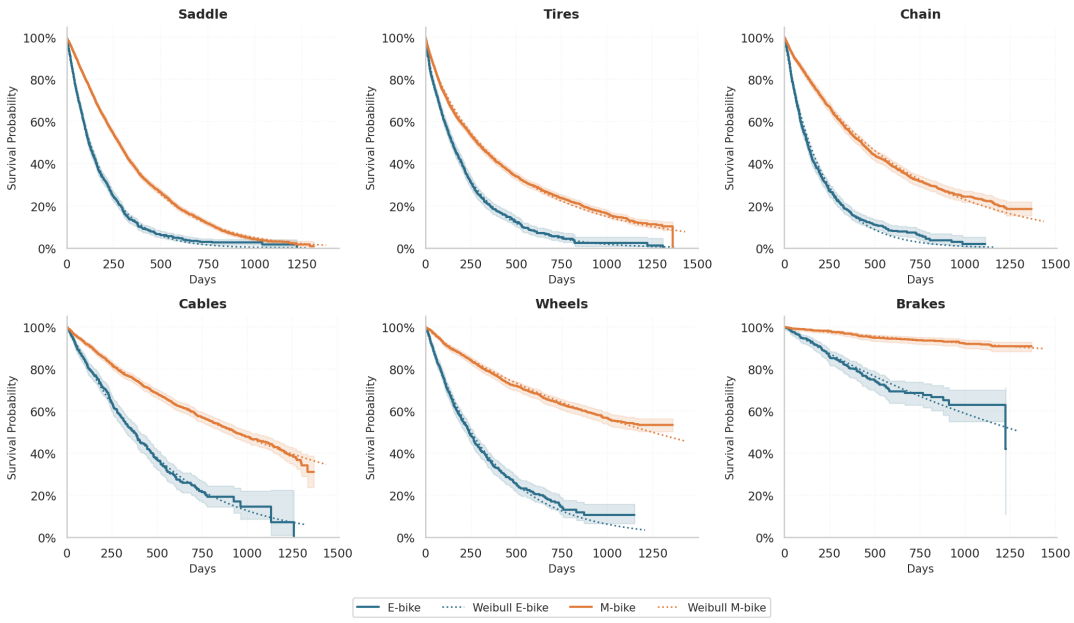


Figure 5.8: Kaplan-Meier survival curves across all evaluated components, stratified by e-bikes and m-bikes, with fitted Weibull overlays.

5.3.2 Covariate-Conditioned Models Against the Baselines

Evaluating covariate-conditioned survival architectures against unconditioned statistical baselines demonstrates the predictive value of integrating operational, temporal, and environmental stress vectors. Table 5.3 summarizes the held-out performance of the semi-parametric CPH model and the non-linear GBSA ensemble across the six mechanical subsystems under the lean feature configuration. Discrimination capacity relative to the unconditioned random-ordering benchmark is illustrated in Figure 5.9, while Figures 5.10 and 5.11 evaluate time-to-event prediction accuracy and probability calibration through the IBS, respectively.

Table 5.3: Held-out predictive performance per component on MetroBikeShare under the lean feature set. Bold marks the better model per metric and component. C-IPCW: higher is better; IBS: lower is better; RMSE is in days, computed on uncensored test units only (n from 32 for brakes to over 1,000 for the saddle). n/a: the GBSA survival function for brakes stays above 0.5 within the follow-up horizon, so no median-based RMSE is reported. Covariate-blind baselines have C-IPCW = 0.500 by construction; the uninformative calibration anchor is IBS = 0.250.

		CPH			GBSA		
Component	**Test Units**	**C-IPCW**	**IBS**	**RMSE**	**C-IPCW**	**IBS**	**RMSE**
Brakes	336	**0.932**	**0.022**	326	0.887	0.040	n/a
Saddle	1,445	0.836	0.101	95	**0.908**	**0.089**	**81**
Wheels	632	0.869	0.064	**167**	**0.931**	**0.062**	219
Tires	1,148	0.839	0.097	120	**0.923**	**0.074**	**83**
Chain	769	0.851	0.087	118	**0.928**	**0.057**	**70**
Cables/Hydraulics	654	0.867	0.065	182	**0.939**	**0.053**	**118**

Both covariate-conditioned survival architectures operate in a high discrimination band across all mechanical components, achieving C-IPCW scores between 0.836 and 0.939. This confirms that routinely logged operational and weather data carry substantial predictive signal for component failure. Direct numerical comparisons with the Barcelona reference study by Grau-Escolano et al. require qualification: while the reference study evaluated model accuracy exclusively on uncensored test records using classical regression metrics (RMSE, R^2, MAPE), this work prioritizes censoring-aware survival metrics (C-IPCW and IBS) to evaluate risk ranking and probability calibration across the entire fleet population. Supplementary regression metrics on uncensored test units yield RMSE values between 70 and 219 days for GBSA (Figure 5.10), placing typical point-prediction error at roughly one third of the median component lifetime for high-turnover consumables.

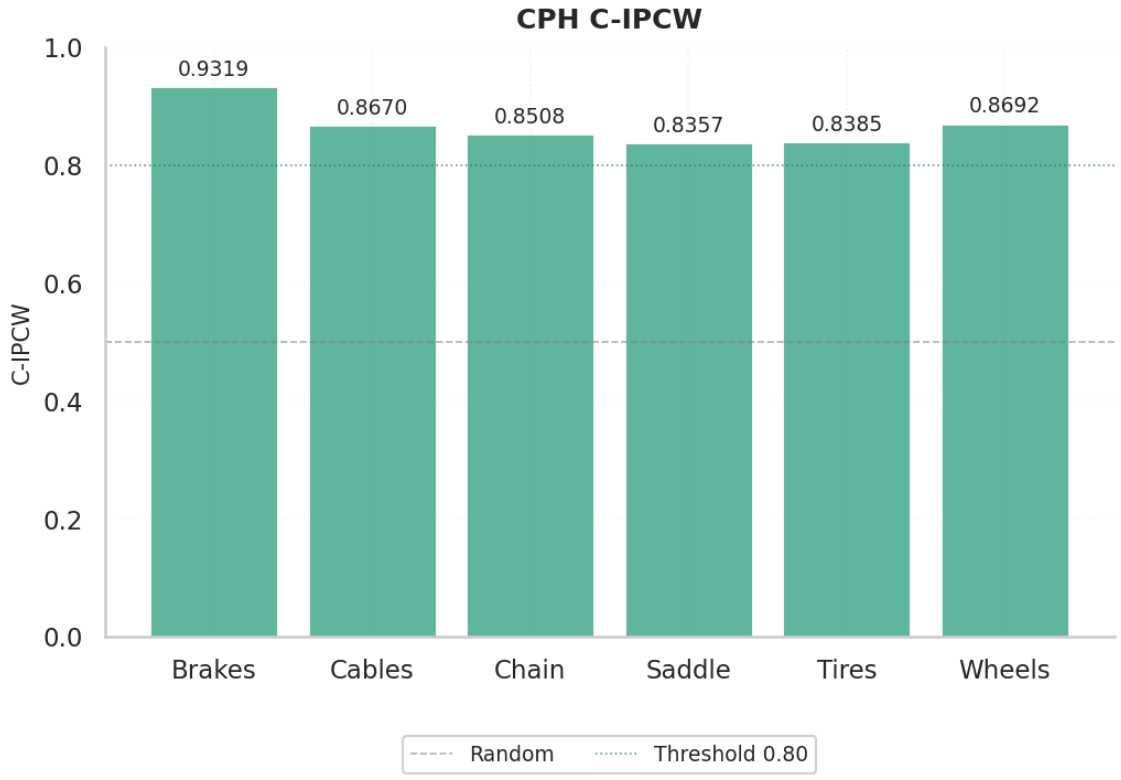
Comparative performance between model families aligns with their underlying inductive biases. GBSA outperforms CPH across five of six components, achieving concordance gains ranging from $+0.062$ to $+0.084$ C-IPCW alongside superior probability calibration. This advantage stems from the capacity of tree-based ensembles to capture non-linear degradation trends and multi-variable feature interactions without enforcing proportional hazards constraints. The single exception occurs in the braking subsystem, where linear CPH leads GBSA in discrimination (0.932 versus 0.887) and calibration (0.022 versus 0.040). Brakes represent the most event-sparse component in the

fleet, containing only 32 observed failures across 336 test units. Under heavy censoring and label sparsity, the higher variance of flexible ensemble models is penalized, allowing the linear risk structure of CPH to maintain superior stability. The same sparsity also degrades point estimation. The GBSA survival function for brakes does not fall below 0.5 anywhere within the observed follow-up horizon, so median-based time-to-event estimates are recovered only by extrapolation far beyond the support of the training data. The resulting values (RMSE $\approx$ 4,651 days) carry no operational meaning and are therefore omitted from Table 5.3.

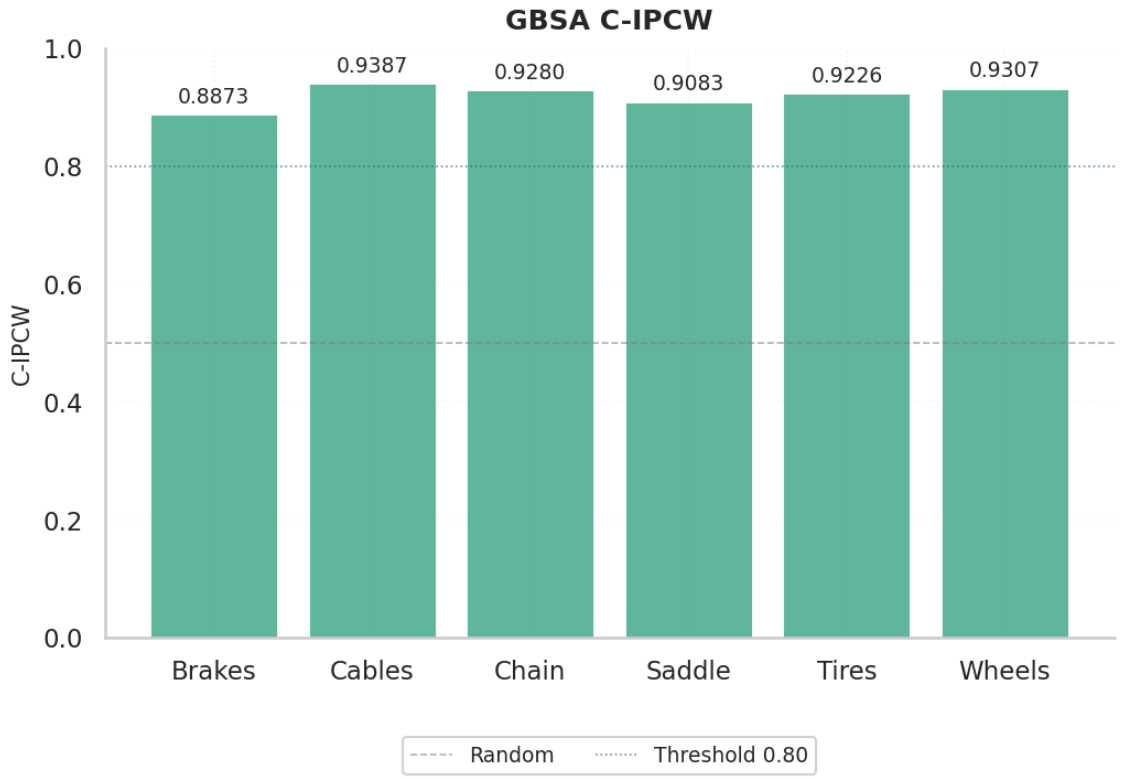
The scatter in Figure 5.10 also exposes a systematic directional bias that aggregate error metrics conceal. Both architectures predict failure later than observed in the majority of uncensored test units. Part of this offset is structural: median-based point predictions are evaluated only on units that failed within the observation window, which are drawn disproportionately from the early tail of the predicted distribution. The distinction matters operationally, because late predictions correspond to interventions scheduled after functional failure, which is the costlier of the two error directions in a shared-fleet setting. Risk ranking and calibration, which drive the condition bands in Section 5.4, are unaffected by this offset, but point estimates of remaining life should be read as ordinal guidance rather than as scheduling deadlines.

The empirical results address the formal research hypotheses established in Chapter 3. The null hypothesis H_0 posits that conditioning time-to-event models on asset usage, weather, and maintenance covariates yields no improvement in risk ranking or probability calibration compared to population-level baselines. Covariate-blind baselines (Kaplan-Meier and Weibull) assign identical risk trajectories to all assets, fixing their concordance at 0.500 by construction. Covariate-conditioned models exceed this chance baseline by margins of 0.336 to 0.439 C-IPCW across all components. Regarding probability calibration, all conditioned models achieve IBS between 0.022 and 0.101 (Figure 5.11), operating between two and eleven times below the uninformative reference anchor of 0.250. The null hypothesis H_0 is rejected for all six mechanical components, and H_1 is supported: individual operational usage and environmental exposure metrics improve component failure risk discrimination and produce well-calibrated time-dependent survival probabilities. The temporal stability of both discrimination and calibration across the follow-up horizon is detailed in Appendix C.

Feature relevance diagnostics confirm the physical validity of the selected covariate space. The drop-column sensitivity analysis detailed in Appendix A was conducted on the full 27–29 feature set rather than the lean configuration, since drop-column procedures yield meaningful gradients only in the presence of correlated predictors; its findings therefore characterize the covariate domain rather than the parsimonious models of Table 5.3. Within this broader feature space, cumulative precipitation consistently ranks as a primary predictor for friction-dependent assemblies such as brakes and tires, reflecting the physical acceleration of mechanical degradation under wet operating conditions. Atmospheric pressure, retained as a negative control variable, exhibits negligible predictive contribution, consistent with the near-zero dispersion reported in Subsection 4.3.5 (CV $\approx$ 0.1%). Complementary hazard ratio and feature importance diagnostics for both models are provided in Appendix B: bike type carries by far the largest hazard ratio under

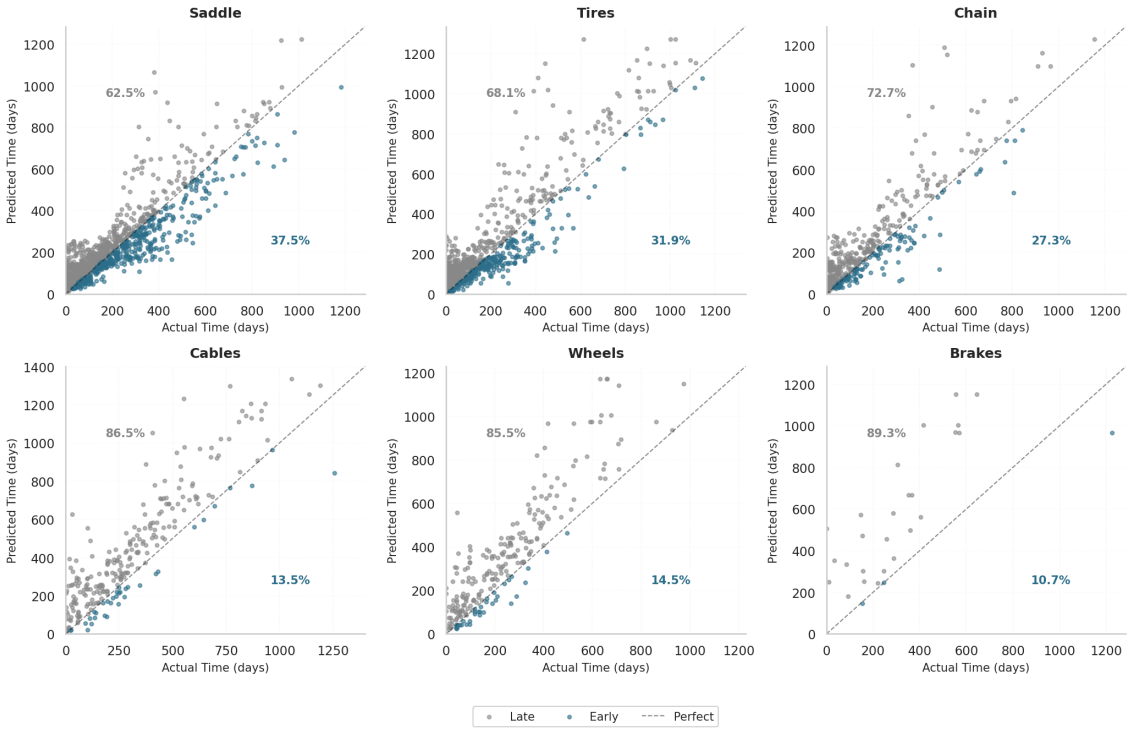


(a) CPH discrimination performance.

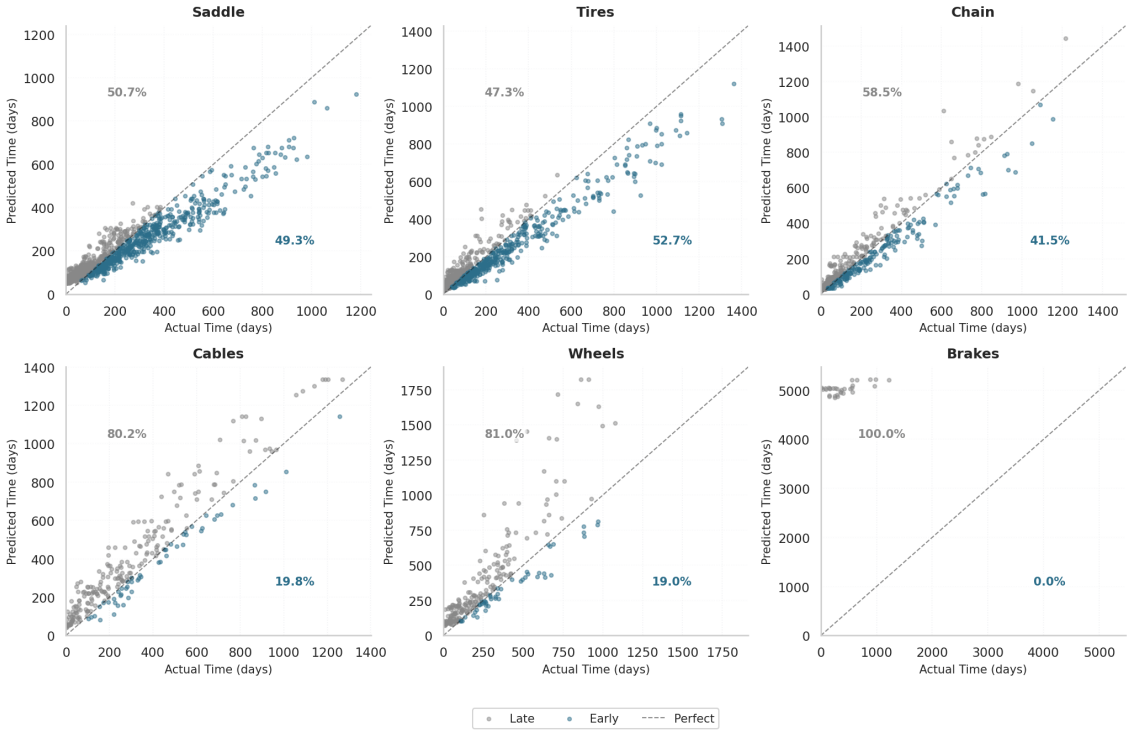


(b) GBSA discrimination performance.

Figure 5.9: Held-out C-IPCW discrimination performance per component under the lean feature configuration: CPH (top) and GBSA (bottom), relative to the chance baseline (0.500).

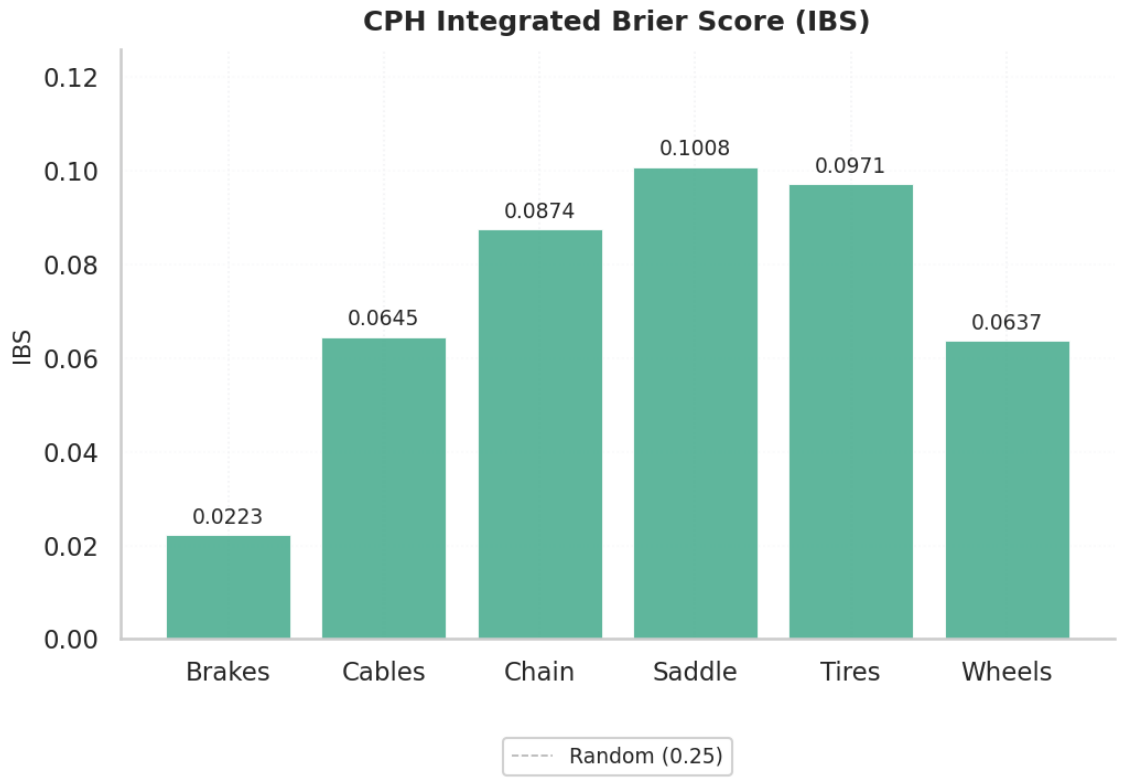


(a) CPH predicted vs. actual time-to-event.

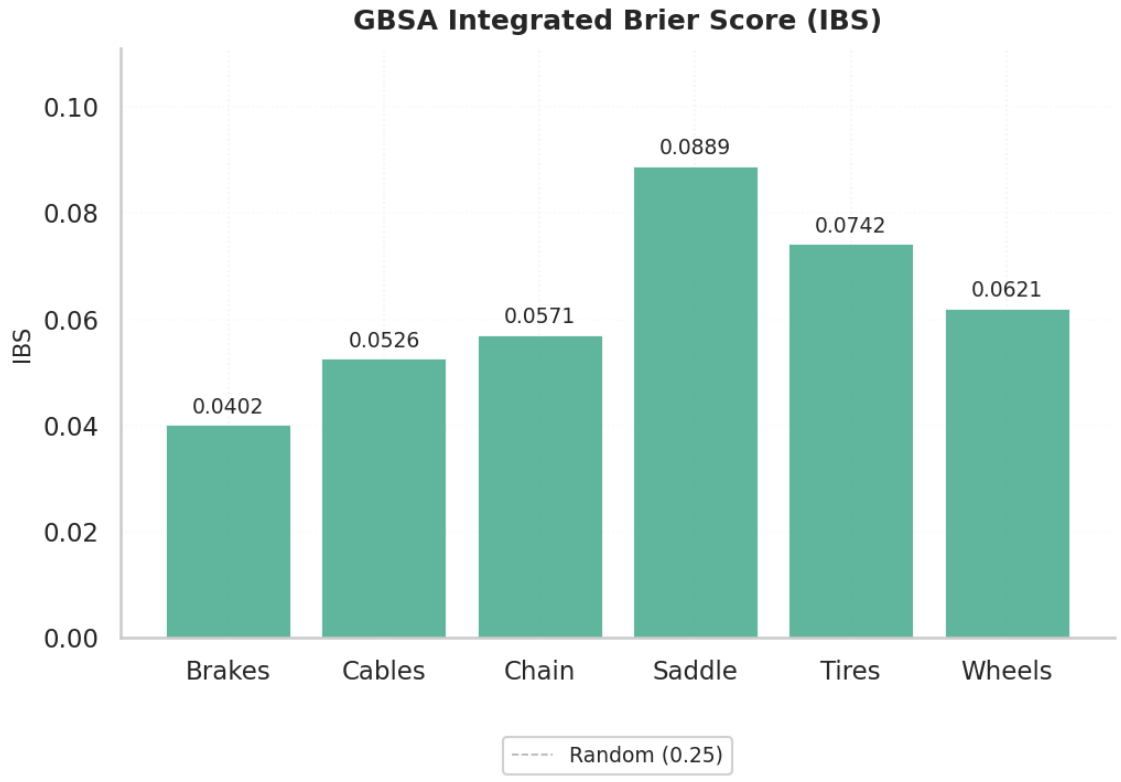


(b) GBSA predicted vs. actual time-to-event.

Figure 5.10: Predicted versus actual time-to-event per component on uncensored held-out units: CPH (top) and GBSA (bottom), relative to the perfect-prediction diagonal. Brake predictions under GBSA are extrapolated beyond the observation horizon and are shown for completeness only.



(a) CPH Integrated Brier Score.



(b) GBSA Integrated Brier Score.

Figure 5.11: Integrated Brier Score per component on the held-out test set: CPH (top) and GBSA (bottom), relative to the uninformative reference anchor (0.250).

CPH, while GBSA instead attributes most predictive weight to precipitation, reflecting the different notions of importance each method captures.

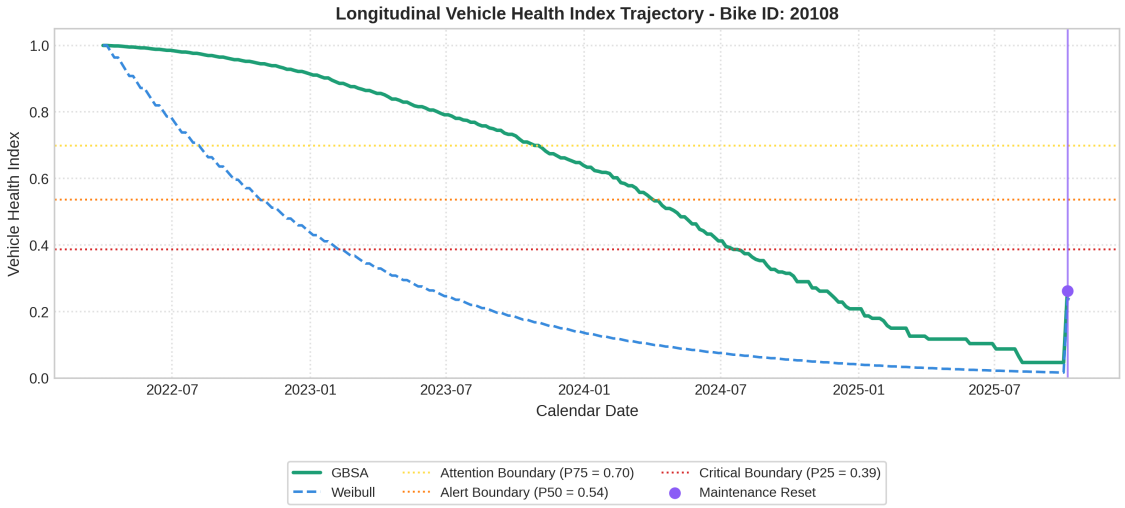
5.4 Health Index Behavior

The monitoring layer transforms time-dependent survival probability functions into continuous asset health trajectories following the formulation in Subsection 3.3.3. Behavioral validation confirms that the derived Health Indices satisfy all operational criteria established in Section 3.6: trajectories remain strictly bounded within $[0, 1]$, decrease as operational usage accumulates within an active interval, and exhibit sharp step-function recoveries following corrective servicing events.

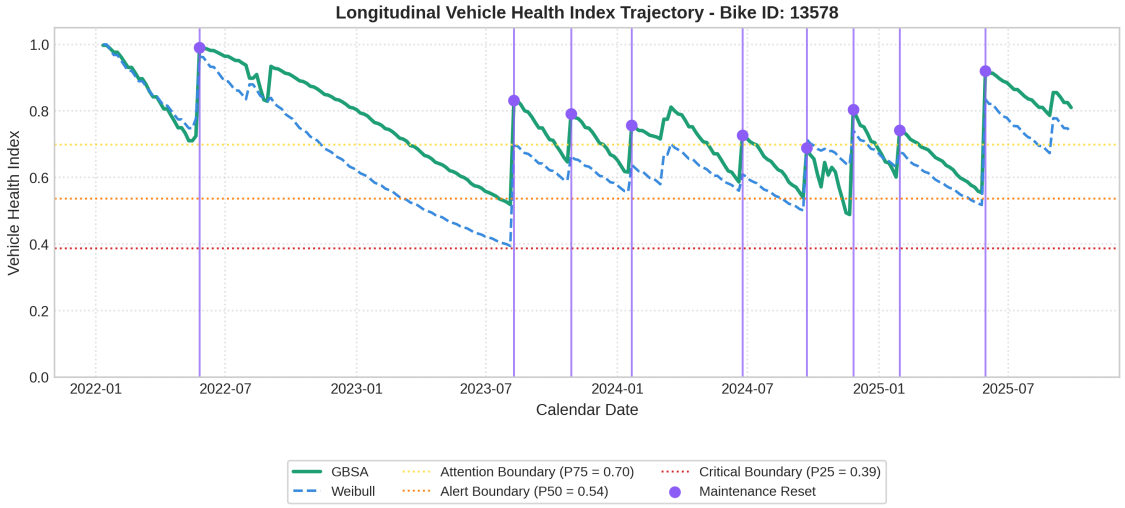
The HI implementation required resolving the trajectory oscillation defect described in Section 4.6. Raw component health series generated on asynchronous date grids created non-physical daily fluctuations when aggregated naively into asset-level Health Indices. Rebuilding component trajectories over a unified daily calendar grid, forward-filling health values across inactive periods, and applying segment-aware smoothing eliminated these structural artifacts. Post-maintenance health recovery jumps are preserved as sharp step-function discontinuities, providing fleet managers with clear visual confirmation of servicing interventions. Figure 5.12 illustrates representative longitudinal Health Index trajectories for individual assets across calendar time, depicting composite vehicle health and step-function recovery jumps for bicycles 20108 and 13578.

A strict conceptual boundary is maintained between quantitative threshold calibration and qualitative multi-interval asset tracking. All quantitative performance metrics and risk-ranking comparisons are computed on the held-out test partition. The empirical Health Index condition boundaries (P_{25}, P_{50} and P_{75}) are likewise derived from the failure events of that partition ($n = 2,720$), and are therefore descriptive cutoffs calibrated on held-out failures rather than boundaries validated on data unseen at calibration time. In contrast, the continuous longitudinal Health Index trajectories presented for individual physical assets evaluate their complete multi-year service history. This continuous visualization reflects the operational inference layer of the framework, wherein frozen, pre-trained survival architectures evaluate daily telemetry streams to demonstrate real-world condition monitoring, monotonicity during active service, and step-function recoveries following recorded workshop interventions. Because no model parameters or decision boundaries are re-estimated during asset-level tracking, evaluating pre-trained models across an asset's full service history introduces no data leakage while providing a complete qualitative assessment of multi-interval vehicle degradation.

Comparing the deterministic Weibull baseline (HI_{det}) against the covariate-conditioned model-based index (HI_{model}) demonstrates the operational value of multi-vector data fusion. While both indices capture general macro-level degradation trends anchored by mileage accumulation, they diverge meaningfully within active intervals. The model-based HI dynamically adjusts asset health trajectories in response to operational intensity and harsh weather exposure that odometry alone cannot detect.



(a) Longitudinal Health Index trajectory for bicycle 20108.



(b) Longitudinal Health Index trajectory for bicycle 13578.

Figure 5.12: Longitudinal Health Index trajectories for bicycles 20108 (top) and 13578 (bottom) across calendar time, depicting component degradation, composite vehicle health, and step-function recovery jumps following recorded maintenance events.

Discretization of continuous Health Indices into four operational condition bands (Healthy, Attention, Alert, and Critical) utilizes empirical failure percentiles observed at the exact moment of corrective maintenance. Figure 5.13 illustrates the empirical Health Index distributions at failure ($n = 2,720$ observed events) across the parametric Weibull baseline, semi-parametric CPH, and non-linear GBSA models. The vertical dashed markers correspond to the 25th (P_{25}), 50th (P_{50}), and 75th (P_{75}) percentiles of each distribution, defining the operational boundaries between condition zones. As summarized in Table 5.4, the GBSA Health Index exhibits lower empirical failure thresholds compared to CPH and Weibull. Assets transition into the Critical band when predicted health drops below the 25th percentile of historical failures, while the Alert boundary aligns with the median health recorded on repair days, equipping operations teams with calibrated decision-support thresholds to schedule proactive workshop interventions before functional failure occurs.

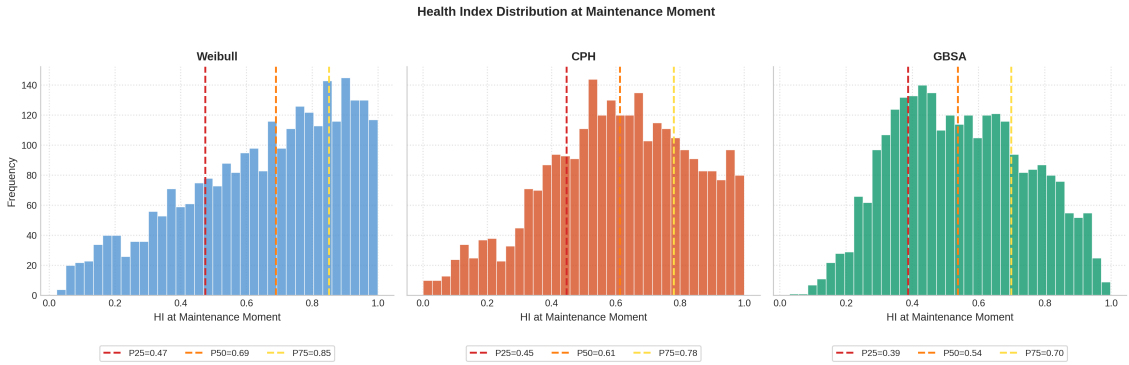


Figure 5.13: Empirical Health Index distributions evaluated at the moment of corrective maintenance ($n = 2,720$ failure events) across Weibull, CPH, and GBSA formulations. Dashed vertical lines mark the 25th (P_{25}), 50th (P_{50}), and 75th (P_{75}) percentiles defining the four operational condition zones.

Table 5.4: Empirical Health Index condition band boundaries derived from failure percentiles ($n = 2,720$ events) on the test set across deterministic and model-based formulations.

Index Model	**Critical** (P_{25})	**Alert** (P_{50})	**Attention** (P_{75})
Deterministic (Weibull)	0.474	0.689	0.851
Model-based (CPH)	0.447	0.613	0.780
Model-based (GBSA)	0.387	0.537	0.699

The inverse-lifetime weighting of Equation 3.3 has a consequence that constrains the operational reading of the composite index. Applying $w_c = 1/\eta_c$ to the parameters of Table 5.2 yields relative weights of 31.9% for the saddle, 26.7% for tires, 22.5% for the chain, 9.3% for cables, 8.4% for wheels and 1.0% for brakes. The bicycle-level index is therefore dominated by high-turnover consumables, while the braking subsystem, contributes almost nothing to it. Component-level indices consequently remain the appropriate monitoring signal for safety-critical subsystems, and the bicycle-level index should be read as an aggregate wear indicator rather than as a safety score.

5.5 Monitoring Dashboard and Decision Support

The practical utility of the proposed predictive maintenance framework relies on translating continuous, probabilistic survival metrics into intuitive operational indicators for fleet management. Following the methodological formulation presented in Chapter 3 and the implementation workflow detailed in Chapter 4, the developed Health Index pipeline converges into a web-based monitoring interface. This section presents the visual and functional evaluation of the prototype, illustrating how individual asset timelines and fleet-wide health snapshots support proactive servicing decisions.

5.5.1 Individual Asset Health Evaluation

The individual asset monitoring module visualizes longitudinal health trajectories for selected bicycles, tracking component degradation and immediate servicing recoveries over time. Figure 5.14 displays the primary monitoring view for an active asset, illustrating the composite Health Index curves generated by the Weibull baseline, the CPH model, and the GBSA ensemble. The interface incorporates historical maintenance interventions directly into the temporal axis, rendering resets as vertical reference lines that align with step-function health recoveries.



Figure 5.14: Individual asset health trajectory for bicycle 13578 within the monitoring dashboard, displaying comparative model curves in the top panel and component-level CPH survival series in the bottom panel.

The lower panel of the interface isolates the conditional survival trajectories of individual mechanical components, providing targeted diagnostic resolution. As shown in Figure 5.14, evaluating components under the CPH formulation reveals distinct degradation rates across subsystems, where high-wear assemblies such as chains and tires decline more rapidly than structural elements. When switching the diagnostic view to the non-linear GBSA model, as depicted in Figure 5.15, the component-level trajectories reflect non-linear wear trends while maintaining sharp recovery steps following documented repair events.

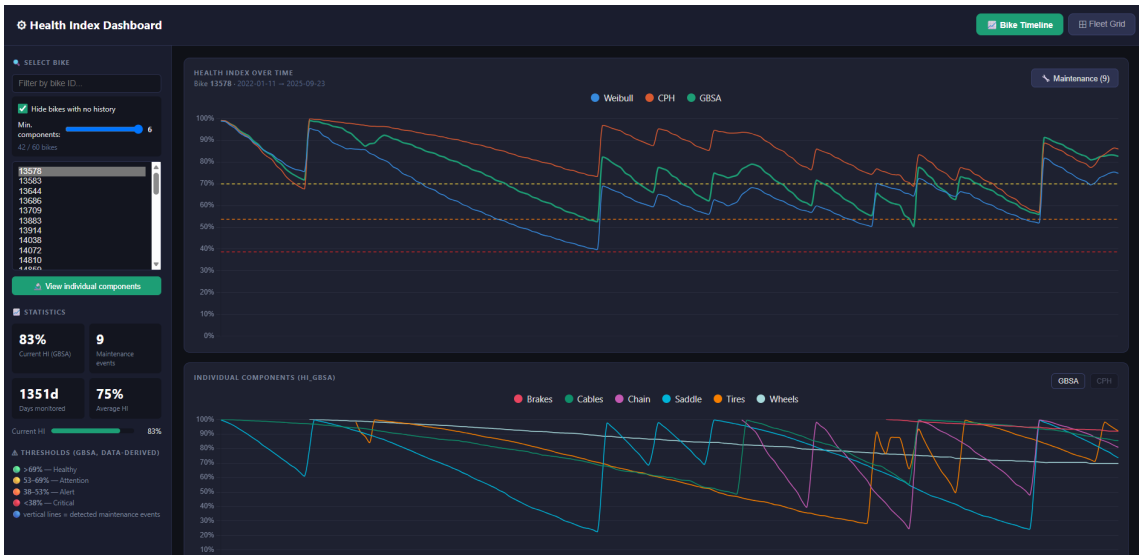


Figure 5.15: Individual asset health trajectory for bicycle 13578 within the monitoring dashboard, displaying comparative model curves in the top panel and component-level GBSA survival series in the bottom panel.

The segment-aware smoothing rules formalized in Chapter 4 prevent daily usage fluctuations from creating artificial trajectory oscillations, preserving clear step-function recoveries upon component replacement. Comparing the linear predictions in Figure 5.14 with the ensemble outputs in Figure 5.15 demonstrates that model-based health indicators adapt dynamically to interval stress, offering a more nuanced diagnostic signal than traditional odometry alone.

5.5.2 Fleet-Wide Health Snapshot and Categorization

To support macro-level operational planning, the Fleet Grid module synthesizes the mechanical condition of the active fleet sampled across weekly observation dates. Data coverage progressively narrows across sequential processing stages of the pipeline. Among the 2,637 bicycles identified in the trip records, 2,327 retain at least one component-level MO unit following left-censoring filtration, and 2,322 yield a computable Health Index series. Because monitoring an asset on a single subsystem provides an incomplete assessment of overall mechanical integrity, the visualization grid enforces a configurable minimum component threshold; under the default requirement of four monitored subsystems, 1,967 assets are displayed. Health Index trajectories are inherently intermittent because inactive or workshop-stationed bicycles generate no telemetry between sampling points, prompting the forward-fill rule detailed in Chapter 4 to propagate the last valid health value across those operational gaps. Assets are mapped into a matrix of color-coded cells corresponding to the empirical Health Index thresholds established in Chapter 4, categorizing vehicles into Healthy, Attention, Alert, and Critical operational bands. Figure 5.16 illustrates the default fleet snapshot, providing an unorganized spatial distribution that highlights the proportion of high-risk vehicles across the system.



Figure 5.16: Fleet Grid view in the monitoring dashboard displaying an unorganized spatial matrix of active fleet assets color-coded by empirical Health Index condition bands on 3 June 2024.

For operational scheduling, the matrix can be reorganized dynamically to group bicycles by their health classification. As shown in Figure 5.17, sorting assets by condition isolates vehicles operating within the Alert and Critical bands at the top of the display. This structural alignment allows workshop managers to quantify daily workshop demand, prioritize high-risk assets for immediate retrieval, and prevent in-service mechanical breakdowns.



Figure 5.17: Fleet Grid view in the monitoring dashboard displaying active fleet assets organized structurally by Health Index severity bands to facilitate workshop task prioritization.

5.6 Synthesis and Limitations

Synthesis of the empirical findings across data processing, survival modeling, Health Index derivation, and prototype visualization demonstrates the operational feasibility and statistical validity of data-driven predictive maintenance for bike-sharing systems. Integrating multi-vector operational usage and meteorological exposure into covariate-conditioned survival architectures produces component-level risk rankings (C-IPCW between 0.836 and 0.939) and well-calibrated failure probabilities (IBS between 0.022 and 0.101) that outperform covariate-blind baselines by a wide margin. Deriving empirical failure thresholds from historical repair logs allows continuous risk curves to be mapped into intuitive operational condition bands, providing fleet managers with a transparent decision-support instrument to transition from reactive fix-on-failure routines toward proactive, risk-informed maintenance scheduling.

These empirical contributions must be interpreted within the context of nine primary study limitations. At the data layer, the predictive pipeline remains constrained by the completeness and logging consistency of operator workshop records, where unrecorded interventions or non-specific free-text entries introduce label noise that statistical post-processing cannot fully eliminate. The proportional hazards assumption underlying the CPH model was not formally tested with Schoenfeld residuals. Because GBSA, which outperforms CPH on five of the six components, imposes no such assumption, the central conclusions do not depend on it; a formal diagnostic remains open. The predictive comparison against the covariate-blind baselines is established against analytical reference points (the random-ordering level for discrimination and the uninformative anchor for calibration) rather than against Brier scores measured from fitted Kaplan-Meier and Weibull models. Reporting the measured baseline calibration alongside confidence intervals for both metrics would place the observed margins on a formal inferential footing and is a direct extension of the evaluation harness. Data access restrictions limit independent third-party replication, as primary maintenance logs remain proprietary and were obtained under non-disclosure agreements with system operators.

At the modeling layer, three constraints affect real-time deployment. The final adopted feature space incorporates interval-total accumulated distance, evaluating model performance under end-of-interval information; a strictly prospective deployment relying exclusively on real-time daily features would experience a minor reduction in predictive signal. Additionally, because covariates are summarized as static interval statistics, the model-based Health Index does not update dynamically in response to mid-interval daily usage fluctuations, unlike the odometry-driven deterministic Weibull index. Point predictions of remaining life additionally skew late on uncensored units, so the framework is more reliable for ranking which assets to service first than for setting a specific service date.

At the operational layer, the discretization thresholds separating Health Index categories represent fixed empirical cutoffs derived from failure percentiles, serving as decision-support boundaries rather than absolute risk probabilities. Finally, primary model training and behavioral validation were conducted on a single bike-sharing fleet. Although complementary cross-domain

experiments demonstrate model transferability, broader multi-city validation across diverse climates and topographies remains an avenue for future research. Despite these constraints, none of the identified limitations undermines the central finding of this dissertation: routinely collected trip and workshop logs carry rich predictive signal capable of forecasting component-level mechanical failures well beyond traditional preventive maintenance baselines.

Chapter 6

Conclusion

This dissertation set out to determine whether the operational records that bike-sharing systems routinely generate carry sufficient predictive signal to anticipate component-level mechanical failures. By transforming raw trip histories and workshop logs into censored survival architectures and continuous health indices, this research connects routine mobility telemetry and condition-based fleet management. The synthesized findings, formal hypothesis verification, and prospective research avenues established in this final chapter confirm that data-driven predictive maintenance is both methodologically sound and operationally viable for shared micro-mobility networks.

6.1 Summary and Contributions

Addressing the scarcity of public component-level maintenance data, this dissertation presents an end-to-end predictive maintenance pipeline trained on operational and workshop records from the MetroBikeShare fleet. Public trip logs and operator-provided workshop entries were cleaned and harmonized into a daily bicycle-level dataset enriched with route geometry and weather exposure. This dataset was subsequently segmented into Maintenance Operation units, defining the observation intervals formalized in Subsection 3.3.1. Component-level survival models were fitted and evaluated using censoring-aware metrics, converting the predicted survival functions into a multi-level Health Index and an interactive monitoring dashboard. This integrated framework fulfills the four research objectives outlined in Section 1.2, yielding contributions across methodological, technical, and applied domains.

The methodological contributions of this dissertation are threefold. First, the work establishes a predictive maintenance framework grounded in survival analysis, utilizing Cox Proportional Hazards and Gradient Boosting Survival Analysis architectures to estimate component-level failure risk from individual usage intensity and environmental exposure histories. Second, it conducts a systematic evaluation comparing these covariate-conditioned models against unconditioned baselines, specifically the non-parametric Kaplan-Meier estimator and the parametric Weibull fit, using censoring-aware metrics. This empirical contrast transforms the operational assumption that predictive maintenance outperforms fixed-interval regimes into a measured, falsifiable claim.

Third, it formulates a multi-level Health Index derived directly from conditional survival functions, aggregating physical asset health from individual mechanical components to composite vehicle scores and fleet-wide monitoring views, resolving a structural gap in the bike-sharing literature reviewed in Chapter 2.

The technical contributions of this dissertation are embodied in a reproducible processing and monitoring pipeline designed for operational deployment and future research extensions. At the data layer, the framework delivers operator-specific harmonization dictionaries that map unstandardized free-text workshop logs into a unified component taxonomy, alongside an automated builder that constructs Maintenance Operation units under strict event and censoring rules. At the evaluation and decision-support layers, the architecture integrates a validation harness built around censoring-aware survival metrics and culminates in an interactive monitoring dashboard. Operating as a self-contained web application without server dependencies, this interface renders multi-level Health Index trajectories through dedicated vehicle timeline and fleet grid views for operations staff. The framework requires only two primary inputs from the fleet operator: trip records and workshop maintenance logs. All remaining contextual covariates are enriched automatically via open routing services and historical weather archives. While trip feeds are increasingly published as open data by operators such as MetroBikeShare, granular workshop logs remain virtually inaccessible in the public domain, as demonstrated in Section 4.2.

The scientific dissemination resulting from this research comprises two papers that document its foundational and extended findings. The first contribution, an audit of the public dataset landscape for bicycle maintenance [11], established the data accessibility gap that motivates the core methodology of this dissertation. The second contribution, accepted for publication, provides a cross-domain evaluation of survival models across three distinct bike-sharing fleets [18], extending the framework beyond the single-fleet baseline evaluated in the primary study.

6.2 Hypothesis Verification

The central research question formulated in Chapter 1 investigated how survival analysis can transform routine bike-sharing operational logs and workshop records into a component-level Health Index capable of monitoring the physical condition of individual assets and the broader fleet. Chapter 3 decomposed this inquiry into a constructive task, the mathematical formulation of the Health Index, and an empirical premise, that operational and environmental covariates carry sufficient predictive signal to outperform population-level baselines. The empirical findings presented in Chapter 5 resolve both dimensions.

Hypothesis H_1 stated that survival architectures conditioned on individual asset usage intensity and environmental exposure achieve superior predictive discrimination and probability calibration compared to covariate-blind baselines. Regarding discrimination, the conditioned architectures achieve C-IPCW scores between 0.836 and 0.939 across the six evaluated mechanical components, whereas covariate-blind baselines assign identical risk trajectories to all assets, fixing their concordance at 0.500 by construction. Because the null hypothesis H_0 asserts that individual usage

and exposure covariates provide no predictive information beyond population averages, this performance margin refutes H_0 across all six mechanical subsystems. Regarding probability calibration, every model-component pair yields an IBS between 0.022 and 0.101, operating between two and eleven times below the uninformative reference anchor of 0.250. This confirms that predicted survival curves provide well-calibrated time-dependent probabilities rather than merely accurate risk rankings. Consequently, the null hypothesis H_0 is rejected, and H_1 is supported at the study level.

The constructive requirement is fulfilled through a bounded Health Index that declines as operational usage accumulates within an active interval, exhibits step-function recoveries following corrective maintenance, and renders asset condition transparently within the interactive monitoring dashboard. The empirical results further illuminate the underlying physical mechanisms driving fleet degradation. Unconditioned population-level Weibull fits produce a decreasing aggregate hazard rate ($\beta < 1$) for five of the six components, a statistical artifact reflecting survivor selection bias in a heterogeneous asset pool rather than physical component rejuvenation. Conditioning on asset-specific stress histories is precisely what the covariate models add to this picture, and the discrimination margins of Subsection 5.3.2 are consistent with that heterogeneity being recoverable from routinely logged covariates. Specifically, pedal-assist electric bicycles exhaust consumable parts two to three times faster than mechanical bicycles, cumulative precipitation ranks among the strongest predictors for friction-dependent braking assemblies, and atmospheric pressure, retained as a negative control, exhibits negligible predictive contribution due to near-zero seasonal dispersion (CV $\approx 0.1\%$).

Three structural qualifications contextualize these findings. First, reported evaluation metrics characterize model performance under end-of-interval information, as the audited feature space incorporates interval-total accumulated distance; a strictly prospective real-time deployment operates with a slightly reduced predictive signal. Second, primary model fitting and behavioral validation were conducted on a single operational fleet, though complementary cross-domain evaluations [18] demonstrate model transferability across distinct urban environments. Third, because primary workshop logs were obtained directly from the operator under non-disclosure agreements rather than from open repositories, independent third-party replication remains constrained by data access, a limitation shared by all empirical reliability studies in the bike-sharing domain.

6.3 Future Work

The limitations identified in Section 5.6 outline a structured roadmap for future investigation, as each boundary highlights specific opportunities to expand the proposed predictive maintenance framework. A primary theoretical direction involves shifting from the current interval-level formulation to dynamic, time-varying survival models. Because observation units in this dissertation are summarized by static covariate vectors over each maintenance interval, the model-based Health Index does not capture intra-interval usage spikes in real time. Reformulating the problem in counting-process form with time-varying covariates [58], [85], or applying dynamic landmarking

techniques [86], would enable failure risk estimates to update continuously on a daily basis. Furthermore, given that mechanical components experience recurrent degradation cycles on the same physical asset, incorporating multi-state or recurrent event modeling approaches [60], [87] would provide a more rigorous mathematical representation than the independent interval approximation. Functional data analysis of wear trajectories represents a complementary methodology to achieve continuous degradation tracking [70], [71].

From an operational perspective, moving from risk estimation to automated decision-making represents the next logical development. The framework developed in this work quantifies component degradation through survival probabilities, but it does not prescribe specific intervention schedules. Integrating a decision layer that maps the Health Index to optimization algorithms under real-world operational constraints would allow fleet operators to balance preventive repair costs against emergency in-service failures [36]. Such an optimization layer directly aligns with the workshop task-assignment algorithms already established in the bike-sharing literature [22]. Furthermore, the categorization thresholds of the monitoring layer, which currently incorporate domain expertise, could be derived analytically from this cost structure. Incorporating a criticality-weighted aggregation factor, expressed as $w_c = \kappa_c/\eta_c$ where κ_c reflects the safety severity of each subsystem, would also refine the fleet-wide health representation and address the observed under-representation of critical assemblies such as braking systems.

Finally, addressing broader industry deployment requires strict prospective evaluation, battery health integration, and standardized data practices. The retrospective cumulative-distance metrics evaluated in this dissertation should be validated under a strictly prospective protocol, ensuring that model predictions at any given moment rely exclusively on information available to operators at that exact instant. In parallel, while battery systems were excluded from the mechanical scope of this thesis due to their electrochemical degradation nature, developing a dedicated State-of-Health framework [42], [43] and composing it with the mechanical Health Index would yield a unified health metric for electric bicycle fleets. Crucially, scaling predictive maintenance across the micromobility sector depends on expanding current data standards. Current feeds, such as the General Bikeshare Feed Specification [10], focus solely on vehicle availability rather than asset health. Standardizing minimal maintenance logs, which include component, action, and date, alongside published trip data would transform predictive maintenance in bike-sharing from a challenge of data access into an established modeling standard.

The transition toward a data-driven predictive framework fundamentally enhances the overall dependability of bike-sharing systems, extending beyond asset longevity to encompass the reliability, availability, maintainability, and safety of the entire fleet. By leveraging survival analysis to estimate component-level failure risks, the proposed methodology enables operators to anticipate mechanical degradation and prevent functional breakdowns. As demonstrated in Section 5.4, the continuous Health Index operationalizes this approach by translating probabilistic risk into actionable condition bands. This diagnostic transparency allows for proactive servicing that minimizes unplanned downtime and mitigates safety hazards, ultimately solidifying shared micromobility as a dependable component of sustainable urban transportation.

References

[1] J. S. Gracias, G. S. Parnell, E. Specking, E. A. Pohl, and R. Buchanan, "Smart Cities-A Structured Literature Review," *Smart Cities*, vol. 6, no. 4, pp. 1719–1743, Aug. 2023.

[2] P. DeMaio, "Bike-sharing: History, Impacts, Models of Provision, and Future," *Journal of Public Transportation*, vol. 12, no. 4, pp. 41–56, Oct. 2009, ISSN: 1077-291X. DOI: 10.5038/2375-0901.12.4.3. Accessed: Jan. 11, 2026. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1077291X22002600.

[3] J. Żywiołek, R. Wolniak, W. W. Grebski, and S. K. Gupta, "City bike systems as an element of building social awareness - The use of resources in a sustainable city," *Cities*, vol. 167, p. 106 312, Dec. 2025, ISSN: 0264-2751. DOI: 10.1016/j.cities.2025.106312. Accessed: Jan. 11, 2026. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0264275125006134.

[4] Y. Zhang and Z. Mi, "Environmental benefits of bike sharing: A big data-based analysis," *Applied Energy*, vol. 220, pp. 296–301, Jun. 2018, ISSN: 0306-2619. DOI: 10.1016/j.apenergy.2018.03.101. Accessed: Jan. 11, 2026. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0306261918304392.

[5] National Aeronautics and Space Administration, "NASA Reliability-Centered Maintenance Guide for Facilities and Collateral Equipment," NASA, Tech. Rep., 2008. [Online]. Available: https://www.nasa.gov/wp-content/uploads/2023/06/nasa-rcmguide.pdf.

[6] J. Moubray, *Reliability-Centered Maintenance*. Industrial Press Inc., 2001, 452 pp., ISBN: 978-0-8311-3146-3.

[7] X.-S. Si, W. Wang, C.-H. Hu, and D.-H. Zhou, "Remaining useful life estimation – A review on the statistical data driven approaches," *European Journal of Operational Research*, vol. 213, no. 1, pp. 1–14, Aug. 2011, ISSN: 0377-2217. DOI: 10.1016/j.ejor.2010.11.018. Accessed: Jan. 11, 2026. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0377221710007903.

[8] J. Grau-Escolano, A. Bassolas, and J. Vicens, "Cycling into the workshop: Predictive maintenance for Barcelona's bike-sharing system," *EPJ Data Science*, vol. 13, no. 1, p. 48, Jul. 2024, ISSN: 2193-1127. DOI: 10.1140/epjds/s13688-024-00486-x. arXiv: 2404.17217 [cs]. Accessed: Jan. 11, 2026. [Online]. Available: http://arxiv.org/abs/2404.17217.

[9] R. K. Mobley, *An Introduction to Predictive Maintenance*. Butterworth-Heinemann, Apr. 2013, ISBN: 978-0-7506-7531-4. Accessed: Jan. 11, 2026. [Online]. Available: https://shop.elsevier.com/books/an-introduction-to-predictive-maintenance/mobley/978-0-7506-7531-4.

[10] MobilityData, *General Bikeshare Feed Specification (GBFS)*, 2025. Accessed: Dec. 7, 2025. [Online]. Available: https://github.com/MobilityData/gbfs.

[11] G. Eggert and D. G. Costa, "Empowering data-driven smart cycling initiatives: A comprehensive analysis of existing datasets for bike maintenance," in *2026 22nd International Conference on Intelligent Environments (IE)*, 2026, pp. 1–4. DOI: 10.1109/IE69249.2026.11539030.

[12] K. Saini, *Bicycle predictive maintenance dataset*. Accessed: Dec. 7, 2025. [Online]. Available: https://www.kaggle.com/datasets/kamlesh0228/bicycle-predictive-maintenance-dataset.

[13] Transport for London, *Santander cycle scheme maintenance logs and image dataset*, FOI request detail, Request ID: FOI-3048-1920, Nov. 2020. Accessed: Dec. 7, 2025. [Online]. Available: https://tfl.gov.uk/corporate/transparency/freedom-of-information/foi-request-detail?referenceId=FOI-3048-1920.

[14] L. P. de Brito Líbano Monteiro, "Predictive maintenance of bike sharing systems using neural networks," M.S. thesis, Instituto Superior Técnico, University of Lisbon, 2022.

[15] A. Pal and Y. Zhang, "Free-floating bike sharing: Solving real-life large-scale static rebalancing problems," *Transportation Research Part C: Emerging Technologies*, vol. 80, pp. 92–116, 2017.

[16] M. Cipriano, L. Colomba, and P. Garza, "A Data-Driven Based Dynamic Rebalancing Methodology for Bike Sharing Systems," *Applied Sciences*, vol. 11, no. 15, p. 6967, Jan. 2021, ISSN: 2076-3417. DOI: 10.3390/app11156967. Accessed: Jun. 5, 2026.

[17] Y. Li, Y. Zheng, H. Zhang, and L. Chen, "Traffic prediction in a bike-sharing system," in *Proceedings of the 23rd SIGSPATIAL International Conference on Advances in Geographic Information Systems*, ACM, 2015, pp. 33–42.

[18] G. Eggert and D. G. Costa, "Survival analysis for predictive maintenance across bike-sharing fleets: A cross-domain evaluation," in *IEEE International Smart Cities Conference*, Porto, Portugal, 2026, in press.

[19] E. Dritsas and M. Trigka, "Big data and internet of things applications in smart cities: Recent advances, challenges, and critical issues," *Internet of Things*, vol. 34, p. 101 770, 2025, ISSN: 2542-6605.

[20] F. Oliveira, D. Nery, D. G. Costa, I. Silva, and L. Lima, "A survey of technologies and recent developments for sustainable smart cycling," *Sustainability, MDPI*, vol. 13, no. 6, pp. 1–28, 2021.

[21] A. Nikolaeva, "Smart cities and (smart) cycling: Exploring the synergies in copenhagen and amsterdam," *Journal of Urban Technology*, vol. 31, no. 1, pp. 29–49, 2024.

[22] L. Lu, S. Zhao, Q.-C. He, and N. Zhu, "Task assignment in predictive maintenance for free-float bicycle sharing systems," *Computers & Industrial Engineering*, vol. 169, p. 108 214, Jul. 2022, ISSN: 0360-8352. DOI: 10.1016/j.cie.2022.108214. Accessed: Jun. 11, 2026.

[23] G. Kapousizis, R. Sarker, M. Baran Ulak, and K. Geurs, "User acceptance of smart e-bikes: What are the influential factors? a cross-country comparison of five european countries," *Transportation Research Part A: Policy and Practice*, vol. 185, p. 104 106, 2024, ISSN: 0965-8564.

[24] Rowingbohe, *Yuhuan Bicycle-sharing System 8*, Wikimedia Commons, Licensed under CC BY-SA 4.0, 2020. Accessed: Jul. 29, 2026. [Online]. Available: https://commons.wikimedia.org/wiki/File:Yuhuan_Bicycle-sharing_System_8.jpg.

[25] K. Wang, X. Yan, Z. Zhu, and X. (Chen, "Understanding bike-sharing usage patterns of members and casual users: A case study in new york city," *Travel Behaviour and Society*, vol. 36, p. 100 793, 2024, ISSN: 2214-367X.

[26] C. J. Shen, N. A. Rahim, and A. R. M. Rosdzimin, "An investigation of bicycle chain lubrication performance in rainy condition," *Journal of Defence Science, Engineering & Technology*, vol. 5, 2022.

[27] J. Alcantara, D. De la Fuente, B. Chico, J. Simancas, I. Díaz, and M. Morcillo, "Marine atmospheric corrosion of carbon steel: A review," *Materials*, vol. 10, 2017. DOI: 10.3390/ma10040406.

[28] D. J. Smith, *Reliability, Maintainability and Risk: Practical Methods for Engineers*, 7th. Oxford: Butterworth-Heinemann, 2005.

[29] G. Scionti, P. Di Pietro, M. Miduri, I. Blandina, and M. F. Milazzo, "From Reactive to Proactive: A Data-driven Approach to Vehicle Maintenance," *Chemical Engineering Transactions*, vol. 119, pp. 385–390, Nov. 2025. DOI: 10.3303/CET25119065. Accessed: Jun. 5, 2026.

[30] K. Khemani and A. N. Qureshi, *AI-Driven Predictive Maintenance with Environmental Context Integration for Connected Vehicles: Simulation, Benchmarking, and Field Validation*, 2026. DOI: 10.48550/ARXIV.2603.13343. Accessed: Jun. 5, 2026.

[31] W. Bauer and J. Baranowski, *Low-Data Predictive Maintenance of Railway Station Doors and Elevators Using Bayesian Proxy Flow Modeling*, Mar. 2026. arXiv: 2603.14384 [eess.SY].

[32] Z. Yang, J. Kanniainen, T. Krogerus, and F. Emmert-Streib, "Prognostic modeling of predictive maintenance with survival analysis for mobile work equipment," *Scientific Reports*, vol. 12, p. 8529, May 2022, ISSN: 2045-2322. DOI: 10.1038/s41598-022-12572-z. Accessed: Jun. 5, 2026.

[33] H. L. Wang, X. Ma, and P. O. Arnäs, "A Data-driven Survival Modelling Approach for Predictive Maintenance of Battery Electric Trucks," *IFAC-PapersOnLine*, 22nd IFAC World Congress, vol. 56, no. 2, pp. 5999–6004, Jan. 2023, ISSN: 2405-8963. DOI: 10.1016/j.ifacol.2023.10.642. Accessed: Jun. 5, 2026.

[34] V. Dimidov, S. Jafarnejad, and R. Frank, *An Empirical Study on Predictive Maintenance for Component X in Heavy-Duty Scania Trucks*. Jun. 2025, p. 9. DOI: 10.1109/ICPHM65385.2025.11061822.

[35] S.-W. Kim and Y. S. Kim, "Two-Stage LightGBM Framework for Cost-Sensitive Prediction of Impending Failures of Component X in Scania Trucks," *Computers, Materials & Continua*, vol. 86, no. 3, p. 50, 2026, ISSN: 1546-2218. DOI: 10.32604/cmc.2025.073492. Accessed: Jul. 31, 2026.

[36] L. Carpentier, A. De Temmerman, and M. Verbeke, "Towards Contextual, Cost-Efficient Predictive Maintenance in Heavy-Duty Trucks," in Apr. 2024, pp. 260–267, ISBN: 978-3-031-58555-5. DOI: 10.1007/978-3-031-58553-1_21.

[37] M. R. Shadi, H. Mirshekali, M. Tahavori, and H. R. Shaker, "Survival Models for Predictive Maintenance and Remaining Useful Life in Sensor-Enabled Smart Energy Networks: A Review," *Sensors*, vol. 26, no. 6, p. 1915, Jan. 2026, ISSN: 1424-8220. DOI: 10.3390/s26061915. Accessed: Jun. 5, 2026.

[38] B. Coutinho, M. Moreira, E. Pereira, and G. Gonçalves, "Survival Analysis-Based System for Predictive Maintenance Optimization," *SN Computer Science*, vol. 6, no. 7, p. 766, Aug. 2025, ISSN: 2661-8907. DOI: 10.1007/s42979-025-04291-9. Accessed: Jun. 5, 2026.

[39] H. A. Abdelhafid, A. Meddaoui, and A. En-nhaili, "Reliable prediction of industrial components Remaining Useful Life using Cox and Weibull models: A Comparative Study," *Data and Metadata*, vol. 4, p. 1102, Aug. 2025, ISSN: 2953-4917. DOI: 10.56294/dm20251102. Accessed: Jun. 10, 2026.

[40] M. Rahat, Z. Kharazian, P. Sheikholharam Mashhadi, T. Rögnvaldsson, and S. Choudhury, "Bridging the Gap: A Comparative Analysis of Regressive Remaining Useful Life Prediction and Survival Analysis Methods for Predictive Maintenance," *PHM Society Asia-Pacific Conference*, vol. 4, no. 1, Sep. 2023, ISSN: 2994-7219. DOI: 10.36001/phmap.2023.v4i1.3646. Accessed: Jun. 5, 2026.

[41] M. Kargar Sharif Abad, "Explainable Predictive Maintenance using Survival Analysis: A Remaining Useful Life prediction of heavy vehicle components," MSc Thesis, Stockholm University, Department of Computer and Systems Sciences, Stockholm, Sweden, 2023.

[42] C.-P. Lin, J. Cabrera, F. Yang, M.-H. Ling, K.-L. Tsui, and S.-J. Bae, "Battery state of health modeling and remaining useful life prediction through time series model," *Applied Energy*, vol. 275, p. 115 338, Oct. 2020, ISSN: 0306-2619. DOI: 10.1016/j.apenergy.2020.115338.

[43] R. Ibraheem, T. I. Cannings, T. Sell, and G. dos Reis, "Robust survival model for the prediction of Li-ion battery lifetime reliability and risk functions," *Energy and AI*, vol. 19, p. 100 465, Jan. 2025, ISSN: 2666-5468. DOI: 10.1016/j.egyai.2024.100465. Accessed: Jun. 5, 2026.

[44] J. Xue, X. Zhao, D. Jiang, Q. Jiao, R. EL Bouchtaoui, and J. Zhang, *Survival Analysis with Machine Learning for Predicting Li-ion Battery Remaining Useful Life*, Version 7, Nov. 2025. DOI: 10.48550/arXiv.2503.13558. arXiv: 2503.13558 [eess.SP]. Accessed: Jul. 31, 2026.

[45] C. Rincón-Maya, D. Acosta-González, F. Guevara-Carazas, F. Hernández-Barajas, C. Patino-Rodriguez, and O. Usuga-Manco, "Predictive modeling of electric bicycle battery performance: Integrating real-time sensor data and machine learning techniques," *Sensors*, vol. 25, no. 5, 2025.

[46] S. Nejadshamsi, J. Bentahar, C. Wang, and U. Eicker, "Predicting short-term bike-sharing demand at station level: A multi-task dynamic graph-based spatiotemporal approach," *Knowledge-Based Systems*, p. 114 986, 2025, ISSN: 0950-7051.

[47] G. Mohseni, M. Nourinejad, and P. Y. Park, "Bike-sharing ridership prediction for network expansion using graph neural networks," *Transportation Research Part A: Policy and Practice*, vol. 203, p. 104 701, 2026, ISSN: 0965-8564.

[48] J. Spicer, C. Richardson, M. Ehrlich, J. Bernstein, M. Fukuda, and M. Terada, "Effects of frictional loss on bicycle chain drive efficiency under variable load conditions," *Journal of Mechanical Design - J MECH DESIGN*, vol. 123, 2001.

[49] S. Lißner, M. von Harten, A. Francke, S. Ruf, and C. Hagemeister, "Safe cycling in winter: Results of a case study on the role of de-icing in the city of hamburg, germany," *Journal of Safety Science*, vol. 87, pp. 244–256, 2023.

[50] Tartu City Government, *Municipal maintenance*, Tartu City Portal, 2025. Accessed: Dec. 7, 2025. [Online]. Available: https://tartu.ee/en/municipal-maintenance.

[51] J. Mendonça, "Predictive maintenance in a bike sharing system using user-centric information," MSc Thesis, Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal, 2025.

[52] B. Sun, J. Feng, and K. Saenko, "Return of Frustratingly Easy Domain Adaptation," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 30, no. 1, Mar. 2016, ISSN: 2374-3468. DOI: 10.1609/aaai.v30i1.10306. Accessed: Jun. 11, 2026.

[53] Y. Li, L. Wang, J. Wang, J. Ye, and C. K. Reddy, "Transfer Learning for Survival Analysis via Efficient L2,1-Norm Regularized Cox Regression," in *2016 IEEE 16th International Conference on Data Mining (ICDM)*, Dec. 2016, pp. 231–240. DOI: 10.1109/ICDM.2016.0034. Accessed: Jun. 5, 2026.

[54] T. G. Clark, M. J. Bradburn, S. B. Love, and D. G. Altman, "Survival Analysis Part I: Basic concepts and first analyses," *British Journal of Cancer*, vol. 89, no. 2, pp. 232–238, Jul. 2003, ISSN: 0007-0920. DOI: 10.1038/sj.bjc.6601118. Accessed: Jun. 5, 2026.

[55] D. G. Kleinbaum and M. Klein, *Survival Analysis: A Self-Learning Text* (Statistics for Biology and Health). New York, NY: Springer, 2012, ISBN: 978-1-4419-6645-2 978-1-4419-6646-9. DOI: 10.1007/978-1-4419-6646-9. Accessed: Jun. 5, 2026.

[56] E. L. Kaplan and P. Meier, "Nonparametric Estimation from Incomplete Observations," *Journal of the American Statistical Association*, vol. 53, no. 282, pp. 457–481, 1958, ISSN: 0162-1459. DOI: 10.2307/2281868. eprint: 2281868. Accessed: Jun. 11, 2026.

[57] J. T. Rich, J. G. Neely, R. C. Paniello, C. C. J. Voelker, B. Nussenbaum, and E. W. Wang, "A practical guide to understanding Kaplan-Meier curves," *Otolaryngology–Head and Neck Surgery*, vol. 143, no. 3, pp. 331–336, Sep. 2010, ISSN: 0194-5998, 1097-6817. DOI: 10.1016/j.otohns.2010.05.007. Accessed: Jun. 5, 2026.

[58] T. M. Therneau and P. M. Grambsch, *Modeling Survival Data: Extending the Cox Model* (Statistics for Biology and Health), K. Dietz, M. Gail, K. Krickeberg, J. Samet, and A. Tsiatis, Eds. New York, NY: Springer, 2000, ISBN: 978-1-4419-3161-0 978-1-4757-3294-8. DOI: 10.1007/978-1-4757-3294-8. Accessed: Jun. 5, 2026.

[59] Z. Zhang, J. Reinikainen, K. A. Adeleke, M. E. Pieterse, and C. G. M. Groothuis-Oudshoorn, "Time-varying covariates and coefficients in Cox regression models," *Annals of Translational Medicine*, vol. 6, no. 7, p. 121, Apr. 2018, ISSN: 2305-5839. DOI: 10.21037/atm.2018.02.12. Accessed: Jun. 5, 2026.

[60] L. D. Amorim and J. Cai, "Modelling recurrent events: A tutorial for analysis in epidemiology," *International Journal of Epidemiology*, vol. 44, no. 1, pp. 324–333, Feb. 2015, ISSN: 0300-5771. DOI: 10.1093/ije/dyu222. Accessed: Jun. 5, 2026.

[61] G. Ridgeway, "The state of boosting," *Computing Science and Statistics*, vol. 31, pp. 172–181, 1999.

[62] T. Hothorn, P. Bühlmann, S. Dudoit, A. Molinaro, and M. J. van der Laan, "Survival ensembles," *Biostatistics*, vol. 7, no. 3, pp. 355–373, Jul. 2006, ISSN: 1465-4644. DOI: 10.1093/biostatistics/kxj011.

[63] F. E. Harrell Jr., K. L. Lee, R. M. Califf, D. B. Pryor, and R. A. Rosati, "Regression modelling strategies for improved prognostic prediction," *Statistics in Medicine*, vol. 3, no. 2, pp. 143–152, 1984, ISSN: 1097-0258. DOI: 10.1002/sim.4780030207. Accessed: Jun. 11, 2026.

[64] H. Uno, T. Cai, M. J. Pencina, R. B. D'Agostino, and L. J. Wei, "On the C-statistics for evaluating overall adequacy of risk prediction procedures with censored survival data," *Statistics in Medicine*, vol. 30, no. 10, pp. 1105–1117, May 2011, ISSN: 1097-0258. DOI: 10.1002/sim.4154. Accessed: Jun. 11, 2026.

[65] C. M. Lillelund, S.-a. Qi, R. Greiner, and C. F. Pedersen, *Position: Stop Chasing the C-index when Evaluating Survival Analysis Models*, May 2026. DOI: 10.48550/arXiv.2506.02075. arXiv: 2506.02075 [stat.ME]. Accessed: Jun. 11, 2026.

[66] E. Graf, C. Schmoor, W. Sauerbrei, and M. Schumacher, "Assessment and comparison of prognostic classification schemes for survival data," *Statistics in Medicine*, vol. 18, no. 17-18, pp. 2529–2545, 1999, ISSN: 0277-6715.

[67] P. Peduzzi, J. Concato, A. R. Feinstein, and T. R. Holford, "Importance of events per independent variable in proportional hazards regression analysis. II. Accuracy and precision of regression estimates," *Journal of Clinical Epidemiology*, vol. 48, no. 12, pp. 1503–1510, Dec. 1995, ISSN: 0895-4356. DOI: 10.1016/0895-4356(95)00048-8.

[68] J. Concato, P. Peduzzi, T. R. Holford, and A. R. Feinstein, "Importance of events per independent variable in proportional hazards analysis. I. Background, goals, and general strategy," *Journal of Clinical Epidemiology*, vol. 48, no. 12, pp. 1495–1501, Dec. 1995, ISSN: 0895-4356. DOI: 10.1016/0895-4356(95)00510-2.

[69] R. Jiang and G. C. Shi, "Development and Application of Vehicle Health Index," *Advanced Materials Research*, vol. 118–120, pp. 459–463, 2010, ISSN: 1662-8985. DOI: 10.4028/www.scientific.net/AMR.118-120.459. Accessed: Jun. 5, 2026.

[70] F. Yao, "Functional principal component analysis for longitudinal and survival data," *Statistica Sinica*, vol. 17, no. 3, pp. 965–983, 2007.

[71] J. Wohlenberg, *Functional Principal Component Analysis and Functional Data*, Jun. 2021. Accessed: Jun. 5, 2026.

[72] Microsoft Learn. "Using survival analysis for predictive maintenance," Accessed: Jun. 5, 2026. [Online]. Available: https://learn.microsoft.com/en-us/archive/msdn-magazine/2019/may/machine-learning-using-survival-analysis-for-predictive-maintenance.

[73] T. Akiba, S. Sano, T. Yanase, T. Ohta, and M. Koyama, "Optuna: A next-generation hyperparameter optimization framework," in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2019, pp. 2623–2631.

[74] C. Davidson-Pilon, "Lifelines: Survival analysis in Python," *Journal of Open Source Software*, vol. 4, no. 40, p. 1317, Aug. 2019, ISSN: 2475-9066. DOI: 10.21105/joss.01317.

[75] S. Pölsterl, "Scikit-survival: A Library for Time-to-Event Analysis Built on Top of scikit-learn," *Journal of Machine Learning Research*, vol. 21, no. 212, pp. 1–6, 2020, ISSN: 1533-7928. [Online]. Available: http://jmlr.org/papers/v21/20-729.html.

[76] *The Meddin Bike-sharing World Map*, 2025. Accessed: Dec. 7, 2025. [Online]. Available: https://bikesharingworldmap.com/.

[77] G. Eggert, *Global bike-sharing systems operators dataset*, Dataset containing contact records, outreach logs, and response statuses for 358 global bike-sharing operators, 2026. [Online]. Available: https://github.com/Eggertt/global-bikeshare-operators-dataset.

[78] A. Boglietti, A. Cavagnino, D. Staton, M. Shanel, M. Mueller, and C. Mejuto, "Evolution and modern approaches for thermal analysis of electrical machines," *IEEE Transactions on Industrial Electronics*, vol. 56, no. 3, pp. 871–882, 2009.

[79] Facewizard, *LACMTA Bike Share station at Bagley and Venice*, Wikimedia Commons, Licensed under CC BY 4.0, 2024. Accessed: Jul. 29, 2026. [Online]. Available: https://commons.wikimedia.org/wiki/File:LACMTA_Bike_Share_station_at_Bagley_and_Venice.jpg.

[80] Facewizard, *LACMTA Bike Share payment station at Bagley and Venice*, Wikimedia Commons, Licensed under CC BY 4.0, 2024. Accessed: Jul. 29, 2026. [Online]. Available: https://commons.wikimedia.org/wiki/File:LACMTA_Bike_Share_payment_station_at_Bagley_and_Venice.jpg.

[81] Heidelberg Institute for Geoinformation Technology, *OpenRouteService: Open-source routing service based on OpenStreetMap data*, version v9.9.0, 2026. Accessed: Jun. 15, 2026. [Online]. Available: https://openrouteservice.org/.

[82] OpenStreetMap contributors, *OpenStreetMap: A free editable map of the world*, 2026. Accessed: Jun. 15, 2026. [Online]. Available: https://www.openstreetmap.org/.

[83] Geofabrik GmbH, *OpenStreetMap data extracts*, 2026. Accessed: Jun. 15, 2026. [Online]. Available: https://download.geofabrik.de/.

[84] P. Zippenfenig, *Open-meteo.com weather api*, 2023. DOI: 10.5281/zenodo.7970649. [Online]. Available: https://open-meteo.com/.

[85] P. C. Austin, A. Latouche, and J. P. Fine, "A review of the use of time-varying covariates in the Fine-Gray subdistribution hazard competing risk regression model," *Statistics in Medicine*, vol. 39, no. 2, pp. 103–113, Jan. 2020, ISSN: 0277-6715. DOI: 10.1002/sim.8399. Accessed: Jun. 5, 2026.

[86] W. K. de Swart, M. Loog, and J. H. Krijthe, "A comparative study of methods for dynamic survival analysis," *Frontiers in Neurology*, vol. 16, p. 1 504 535, Feb. 2025, ISSN: 1664-2295. DOI: 10.3389/fneur.2025.1504535. Accessed: Jun. 5, 2026.

[87] StataCorp. "Analysis of multiple failure-time survival data," Accessed: Jun. 5, 2026. [Online]. Available: https://www.stata.com/support/faqs/statistics/multiple-failure-time-data/.

Appendix A

Drop-Column Feature Sensitivity Analysis

To complement the feature screening and data leakage audit detailed in Subsection 4.4.2, this appendix presents the complete set of drop-column sensitivity diagnostic plots across all six modeled mechanical subsystems: brakes, cables, chain, saddle, tires, and wheels. The drop-column evaluation quantifies the predictive contribution of each candidate covariate by systematically removing one feature at a time, refitting the CPH baseline model on the full candidate feature set of the MetroBikeShare dataset, and measuring the resulting shift in predictive concordance (ΔC-IPCW).

The directional change in concordance dictates the operational value of each attribute. A positive delta, represented by a green bar, signifies that feature removal degrades model discrimination, confirming that the variable provides predictive information regarding component failure dynamics. Conversely, a negative delta, represented by an orange bar, indicates that feature removal improves concordance, identifying collinear, redundant, or noisy attributes whose exclusion refines the final lean feature space. Marginal shifts close to zero, shown in grey, reflect neutral feature contributions.

Across the mechanical assemblies, the diagnostics illustrate distinct physical failure drivers. Primary work metrics and localized environmental stressors consistently display high positive deltas for structural and friction-based components such as brakes and tires. In contrast, uninformative cross-component maintenance counts and atmospheric pressure metrics exhibit negative deltas, validating their exclusion during feature space pruning. The individual drop-column importance distributions for all six subsystems are provided in Figure A.1 through Figure A.6.

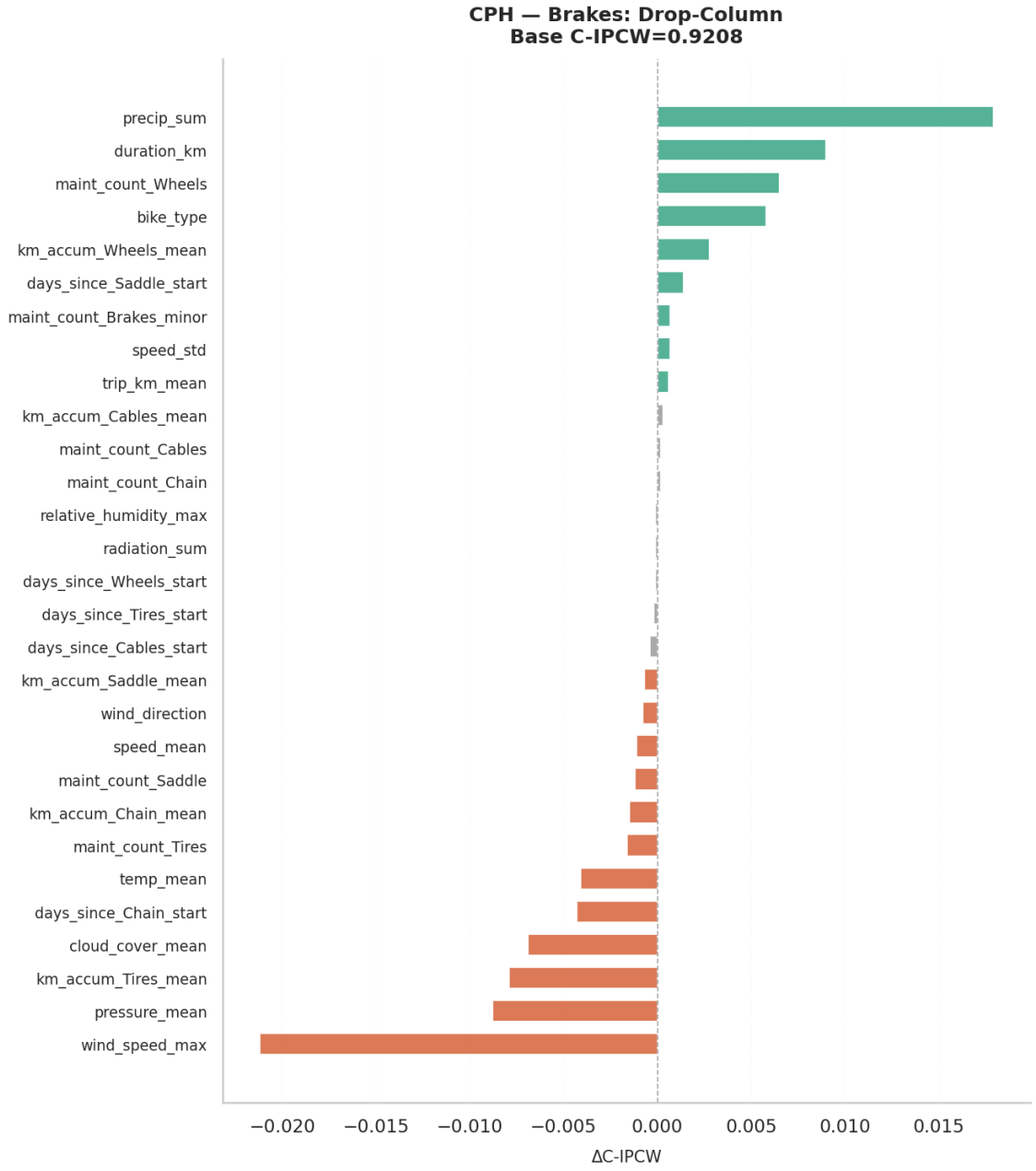


Figure A.1: Drop-column feature importance plot for the braking subsystem (baseline C-IPCW = 0.9208).

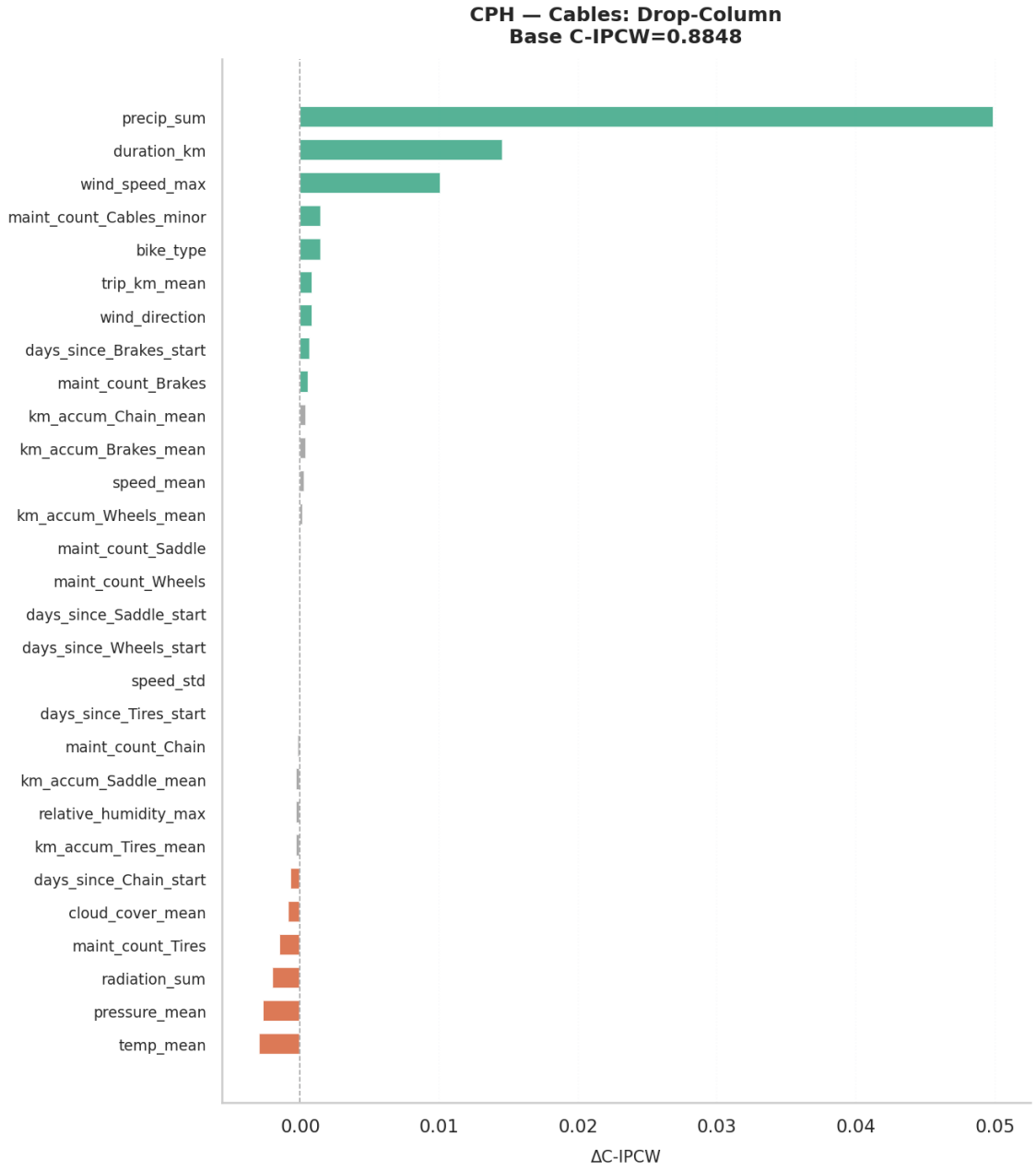


Figure A.2: Drop-column feature importance plot for the cable assembly (baseline C-IPCW = 0.8848).

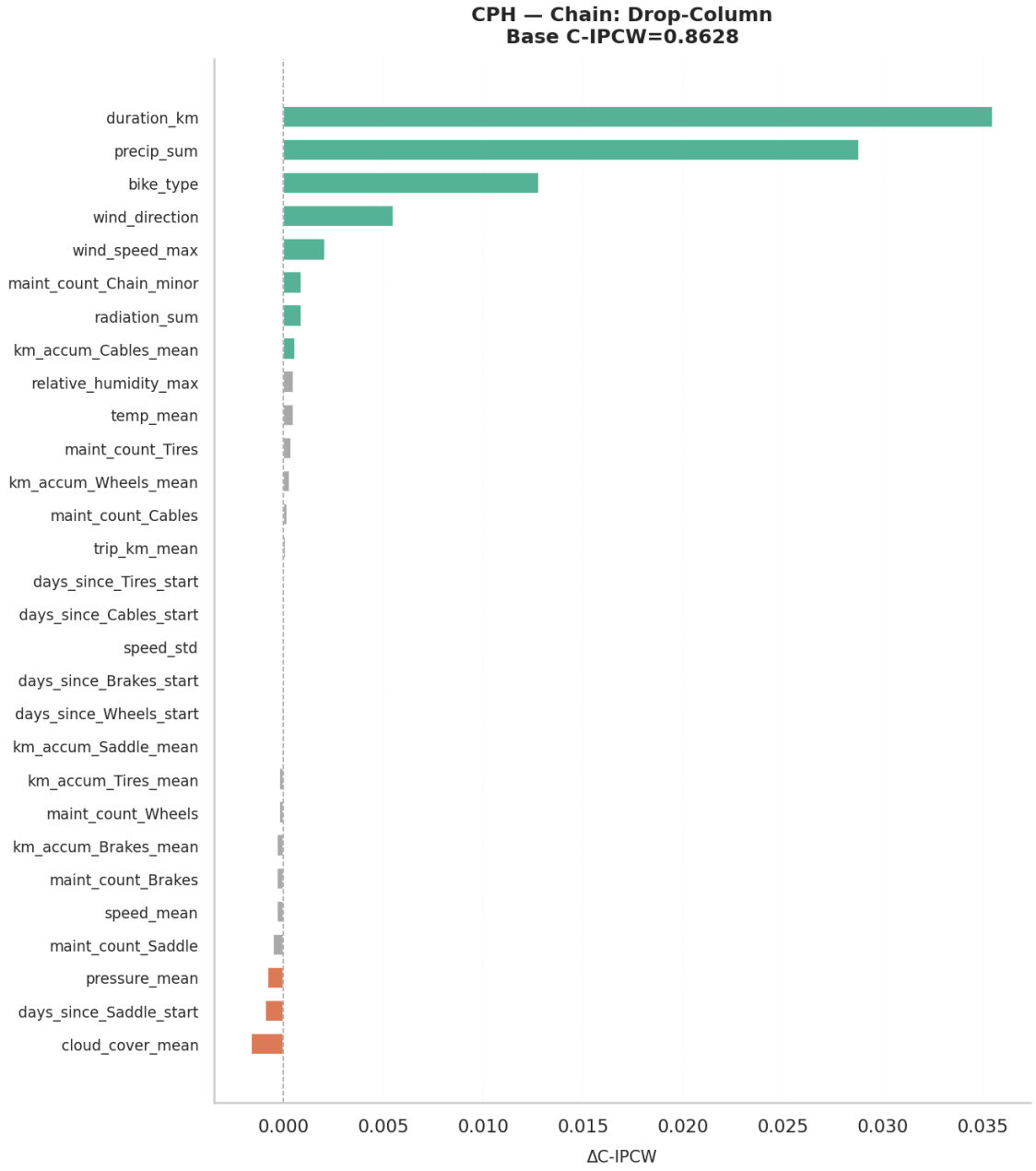


Figure A.3: Drop-column feature importance plot for the chain drive assembly (baseline C-IPCW = 0.8628).

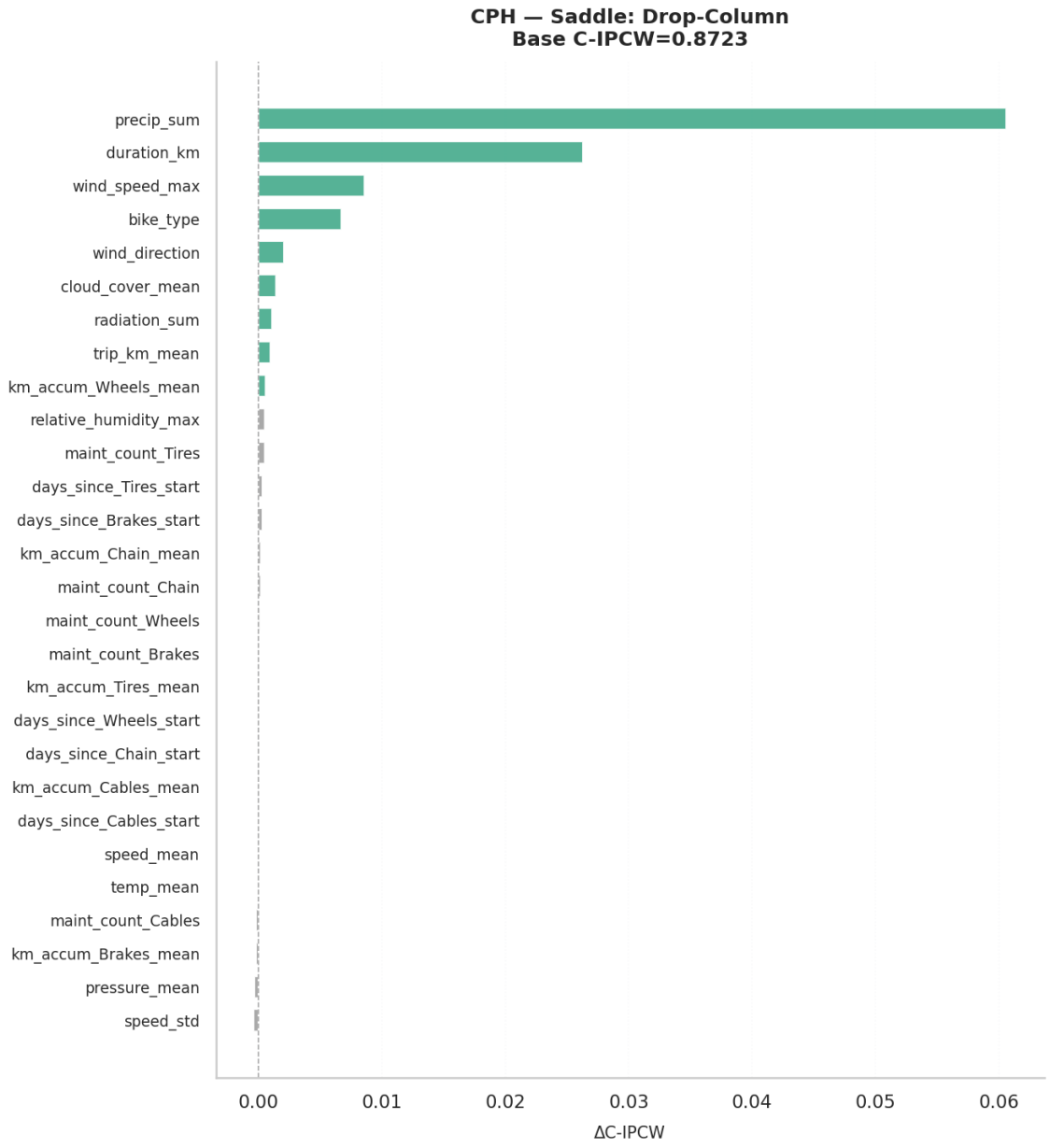


Figure A.4: Drop-column feature importance plot for the saddle subsystem (baseline C-IPCW = 0.8723).

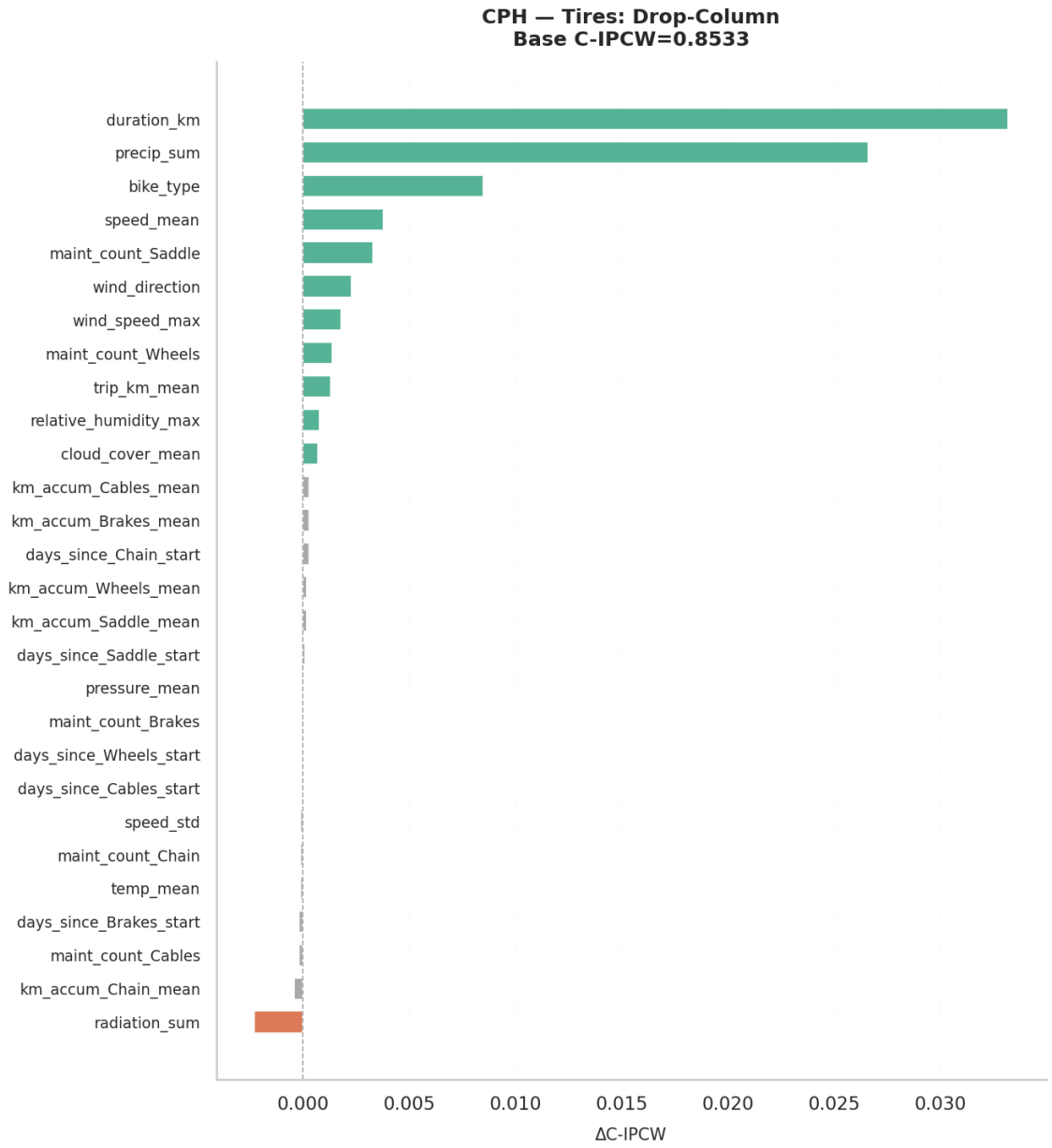


Figure A.5: Drop-column feature importance plot for the tire assembly (baseline C-IPCW = 0.8533).

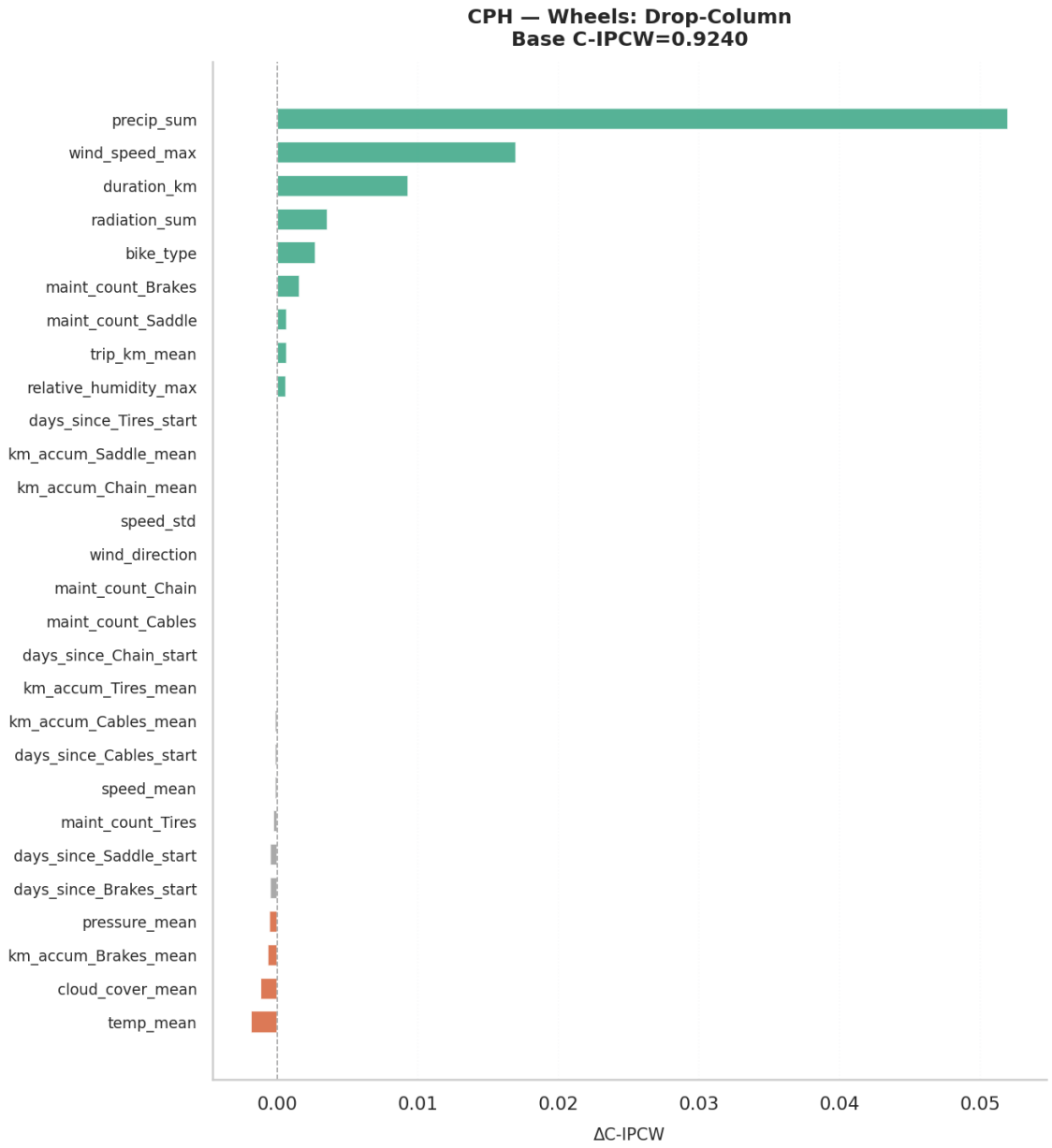


Figure A.6: Drop-column feature importance plot for the wheel assembly (baseline C-IPCW = 0.9240).

Appendix B

Component-Level Model Interpretability

This appendix complements the comparative evaluation discussed in Subsection 5.3.2 by providing detailed interpretability diagnostics for the semi-parametric CPH model and the non-linear GBSA ensemble under the audited lean feature space. Inspecting covariate contributions across both model families illustrates how linear multiplicative risk assumptions contrast with non-linear tree-based feature importance assignments.

For the linear CPH formulation, Figure B.1 displays the estimated hazard ratios alongside their 95% confidence intervals across all six mechanical subsystems. A hazard ratio greater than unity indicates an accelerated failure risk per unit increase in the covariate, whereas a ratio below unity reflects a protective effect or lower failure hazard. Across all components, bike type emerges as a dominant linear hazard multiplier, confirming that pedal-assist electric bicycles experience systematically higher baseline degradation rates than mechanical bicycles.

For the GBSA model, where multiplicative hazard ratios cannot be directly extracted due to non-linear feature splitting, Figure B.2 reports permutation-based feature importance scores. Unlike the linear model, the ensemble attributes substantial predictive weight to environmental exposure, specifically cumulative precipitation, for friction-dependent assemblies such as brakes, tires, and cables. This divergence highlights the capacity of tree-based ensembles to capture environmental interactions that linear risk formulations dilute.

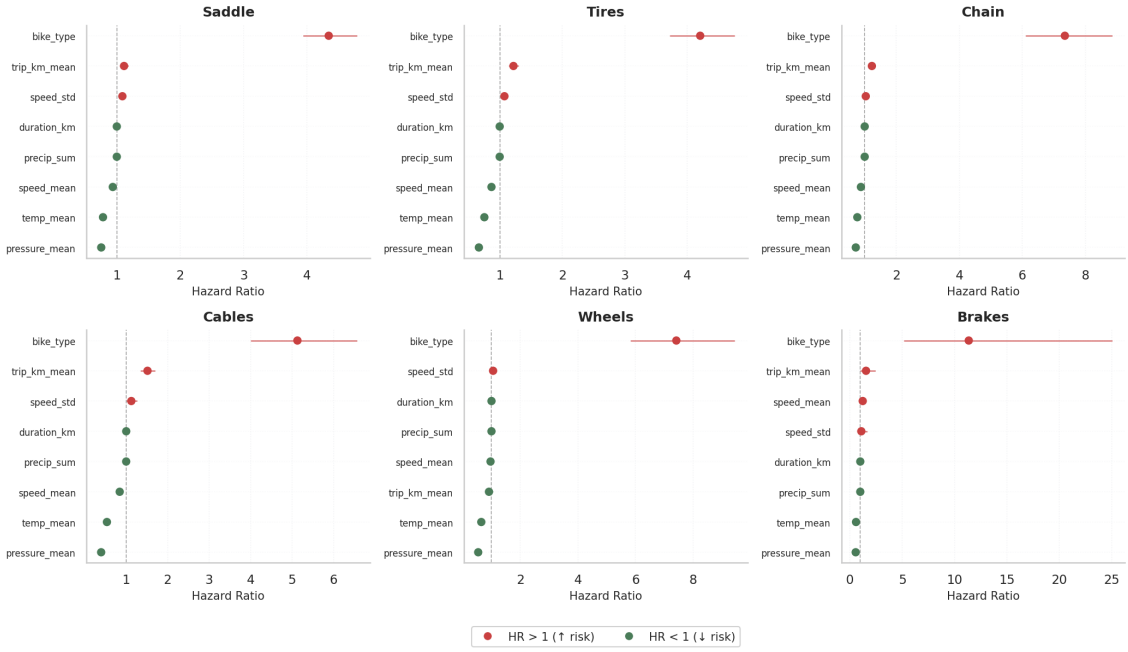


Figure B.1: Fitted CPH model hazard ratios with 95% confidence intervals, per component, under the lean feature configuration.

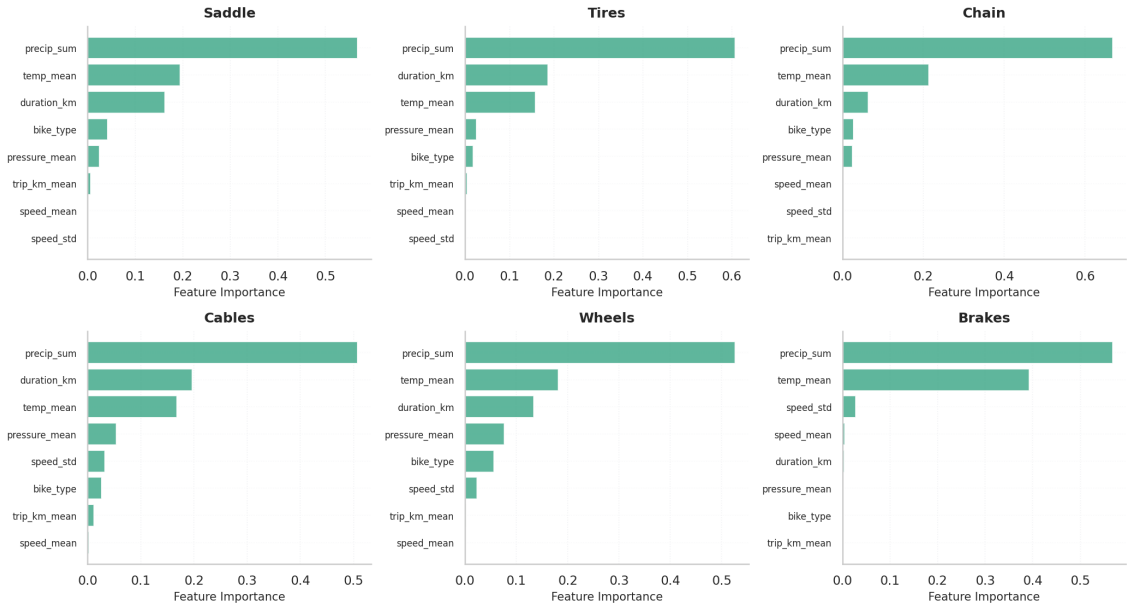


Figure B.2: GBSA feature importance scores per component, under the lean feature configuration.

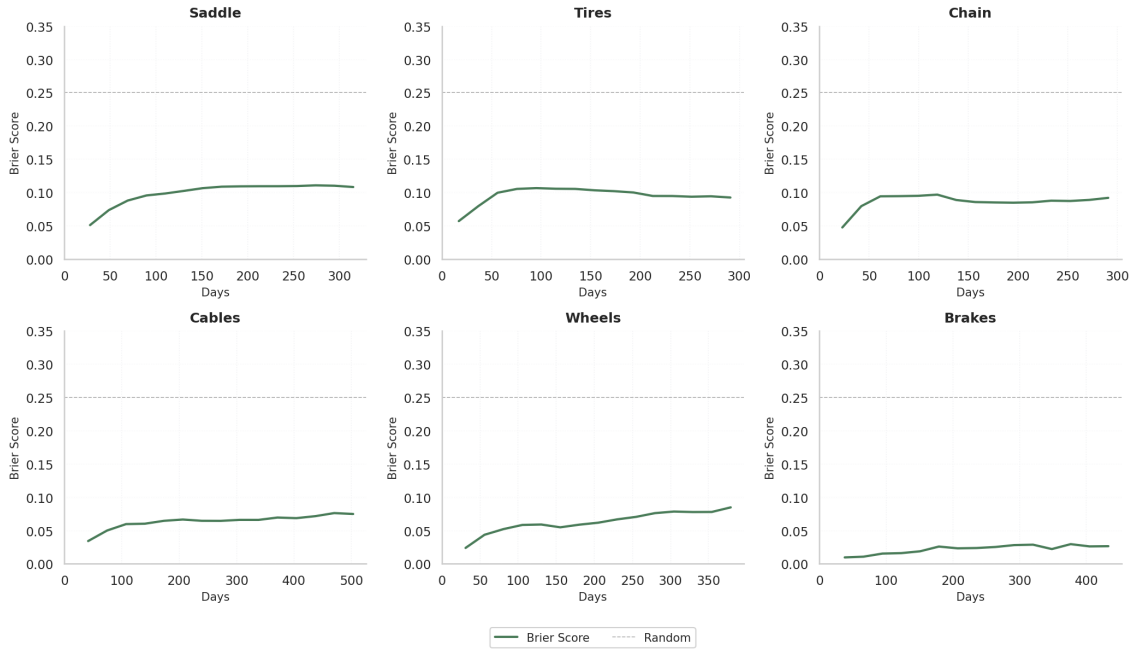
Appendix C

Time-Varying Discrimination and Calibration Diagnostics

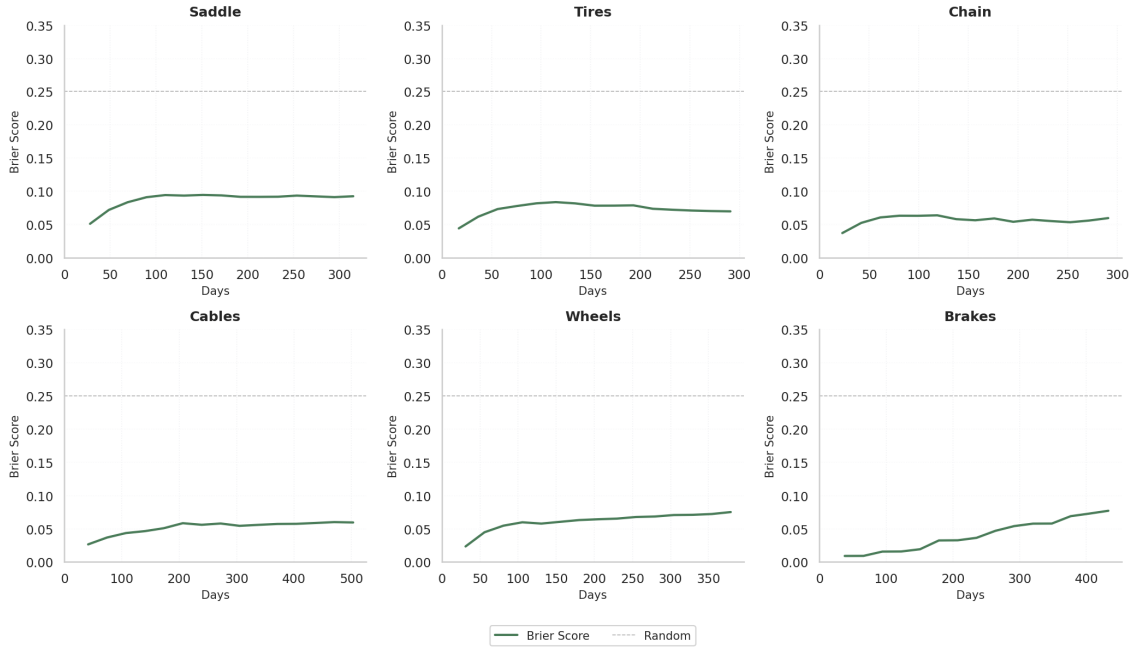
This appendix complements the aggregated summary metrics presented in Subsection 5.3.2 by detailing the time-dependent performance trajectories from which the IBS and overall C-IPCW are derived. Evaluating discrimination and calibration across continuous follow-up horizons verifies model stability over time and ensures that localized temporal windows do not drive high aggregate performance.

Figure C.1 illustrates the time-dependent Brier score trajectories for both CPH and GBSA models across all six mechanical components. The Brier score measures probability calibration over time, where lower values indicate superior numerical agreement between predicted survival curves and observed event states. Across the operational follow-up horizon, both models maintain calibration error curves below the uninformative reference anchor of 0.250, demonstrating consistent probability estimation as observation intervals lengthen.

Figure C.2 reports the corresponding time-dependent Area Under the ROC Curve (AUC), evaluated under Inverse Probability of Censoring Weighting. The time-dependent AUC quantifies dynamic risk discrimination at specific operational ages, measuring the model's capacity to correctly order asset failure risk at distinct calendar horizons. Across all components, discrimination scores remain stably positioned within high performance bands throughout the follow-up period, operating well above the 0.500 chance baseline and confirming temporal generalization across active vehicle lifecycles.

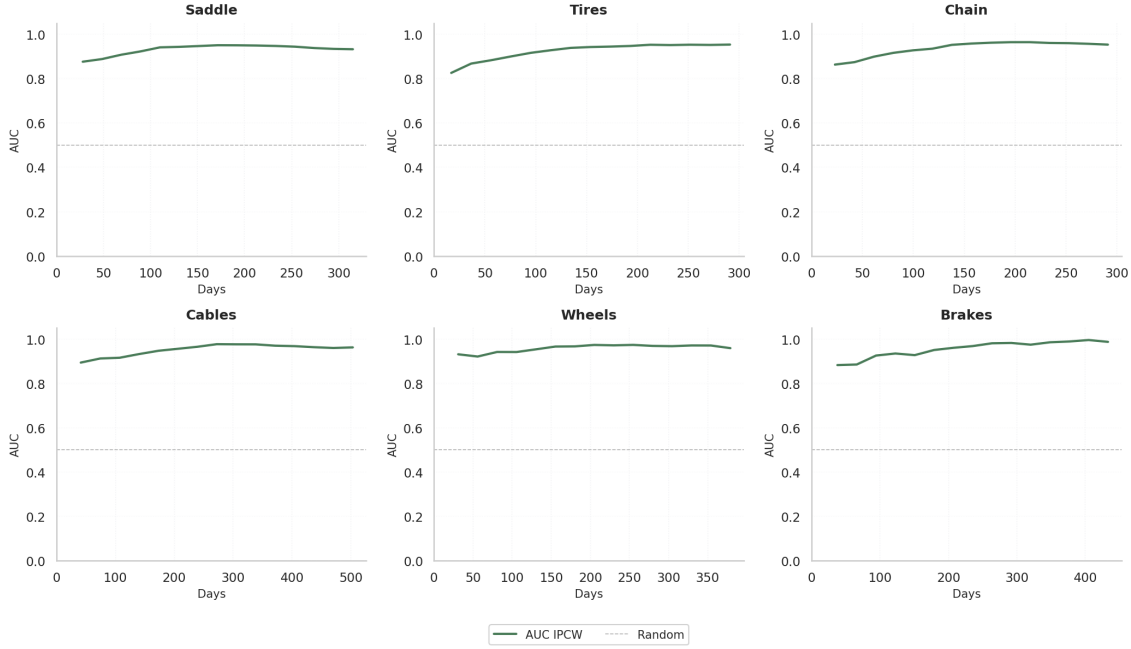


(a) CPH Brier score over time.

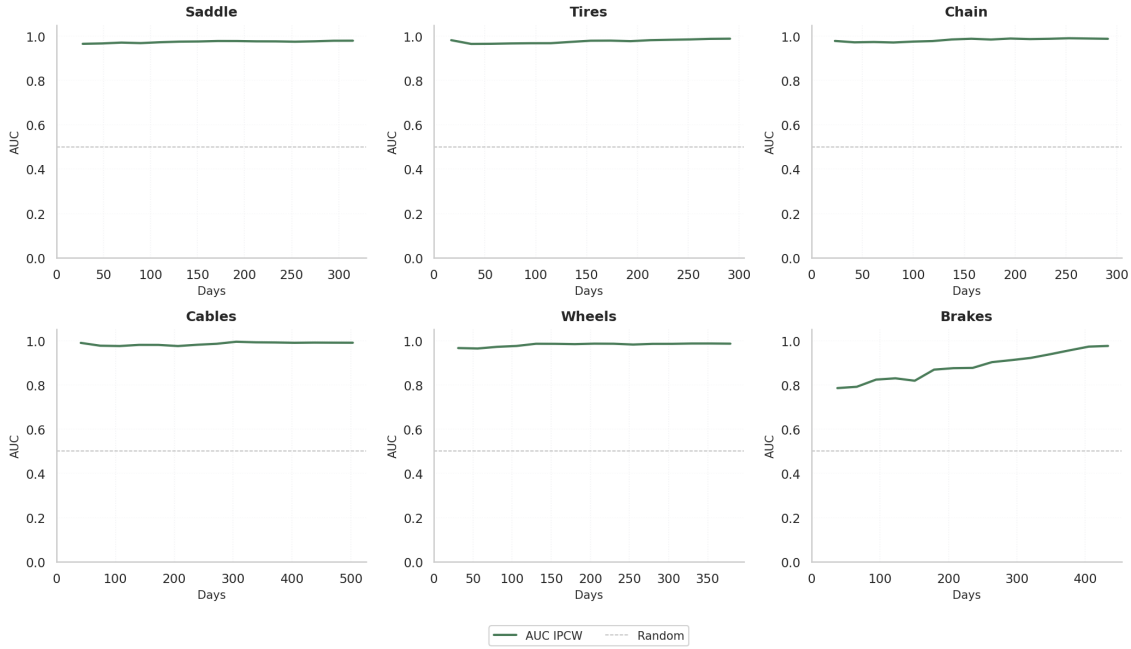


(b) GBSA Brier score over time.

Figure C.1: Time-dependent Brier score curves across the follow-up horizon, per component: CPH (top) and GBSA (bottom), relative to the uninformative reference anchor (0.250).



(a) CPH time-dependent AUC.



(b) GBSA time-dependent AUC.

Figure C.2: Time-dependent AUC curves (IPCW-weighted) across the follow-up horizon, per component: CPH (top) and GBSA (bottom), relative to the chance baseline (0.500).