



# Modeling and Optimising Demand-Responsive Road Network

Lei Lyu

Student ID: F313871

Principal Supervisor: Prof Diwei Zhou  
Secondary Supervisors: Prof Marcus Enoch, Dr Yasir Ali, Dr Qamar  
Natsheh

A report submitted to the Loughborough University  
For the PhD Part R2 Progress Review

Department of Mathematical Sciences  
Loughborough University  
Loughborough, UK

August 2025

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Background . . . . .	1
1.1.1	Urban Traffic Congestion: Challenges . . . . .	1
1.1.2	BPR Function in Traffic Assignment . . . . .	2
1.2	Project Aim . . . . .	3
1.3	Project Objectives . . . . .	3
1.4	Benchmarking . . . . .	3
1.5	Summary of Work Completed (Year 2) . . . . .	4
<b>2</b>	<b>Literature Review</b>	<b>8</b>
2.1	Framework of the Systematic Review . . . . .	8
2.1.1	Review Questions and Scope . . . . .	8
2.1.2	Protocol and Reporting Standards . . . . .	8
2.1.3	Information Sources and Search Strategy . . . . .	8
2.1.4	Eligibility Criteria (PICOS) . . . . .	9
2.1.5	Screening, Deduplication, and Flow of Studies . . . . .	10
2.1.6	Data Extraction and Quality . . . . .	11
2.2	BPR Modification Cluster . . . . .	12
2.2.1	Cluster Method . . . . .	12
2.2.2	Operational definitions and decision rules (A–G) . . . . .	12
2.2.3	Rule precedence and edge cases . . . . .	14
2.2.4	Comparability and identifiability tests . . . . .	14
2.2.5	Reproducible labelling protocol . . . . .	14
2.2.6	Evidence Map . . . . .	15
2.3	Performance Metrics and Validation . . . . .	16
2.4	Synthesis and Knowledge Gaps . . . . .	16
<b>3</b>	<b>Data Exploratory Analysis (EDA)</b>	<b>18</b>
3.1	Data source . . . . .	18
3.2	Variable dictionary and notation . . . . .	20
3.3	Pre-processing and Quality Control . . . . .	22
3.3.1	Data Cleansing . . . . .	22
3.3.2	Time alignment and deduplication . . . . .	22
3.3.3	Missing data treatment . . . . .	23
3.3.4	Outlier detection and robust filtering . . . . .	24
3.3.5	Aggregation and class handling . . . . .	24
3.3.6	Reconstruction of travel time . . . . .	24

## CONTENTS

---

3.4	Derived quantities: density, capacity . . . . .	24
3.4.1	Density and fundamental relations . . . . .	24
3.4.2	Empirical capacity $C$ . . . . .	24
3.5	Exploratory data analysis (EDA) . . . . .	25
3.5.1	Descriptive statistics . . . . .	26
3.5.2	Diurnal and Weekly Patterns . . . . .	27
3.5.3	Lane-level Analysis . . . . .	28
3.5.4	Congestion footprint . . . . .	30
3.5.5	Flow-speed and fundamental relations . . . . .	31
3.5.6	Correlation Analysis . . . . .	32
3.5.7	Conclusion . . . . .	33
3.6	Train-Validation-Test Design . . . . .	33
3.7	Limitations and data readiness for modelling . . . . .	34
3.8	Identification and Interpretation of Anomalous Events . . . . .	34
3.9	Summary of Findings and Conclusion . . . . .	35
<b>4</b>	<b>Benchmarking Methodology</b> . . . . .	<b>37</b>
4.1	Baseline: Classical BPR Function . . . . .	37
4.2	Benchmarking Model Selection Criteria . . . . .	38
4.2.1	Benchmarking Model list . . . . .	38
4.2.2	Cluster A split (M1 and M2) . . . . .	38
4.2.3	Scope and Exclusions . . . . .	39
4.3	Provenance and justification for each Cluster . . . . .	39
4.4	Benchmark Modelling Details . . . . .	41
4.4.1	M1 [1–3]: Dynamic Parameter BPR (DP-BPR) . . . . .	41
4.4.2	M2 [4–7]: Fundamental-Diagram–Informed VDF (FD–VDF) . . . . .	42
4.4.3	M3 [8]: Multi-Class (HGV-Adjusted) BPR (MC–BPR) . . . . .	43
4.4.4	M4 [9]: External-Factor–Adjusted BPR (EF-BPR) . . . . .	44
4.4.5	M5 [2, 9]: Machine Learning Hybrid BPR (ML-hBPR) . . . . .	46
4.4.6	M6 [10]: Stochastic-Capacity / Reliability BPR (SC–BPR) . . . . .	47
4.4.7	Common Calibration and Evaluation Protocol . . . . .	48
4.5	Performance Metrics Definition . . . . .	48
<b>5</b>	<b>Results and Comparative Evaluation</b> . . . . .	<b>51</b>
5.1	Evaluation Design (fixed before analysis) . . . . .	51
5.1.1	Datasets and splits . . . . .	51
5.1.2	Outcome and metrics . . . . .	52
5.2	Headline Results . . . . .	52
5.3	Stratified Performance . . . . .	53
5.4	Calibration Diagnostics . . . . .	53
5.4.1	Parameter estimates . . . . .	53
5.4.2	Goodness-of-fit plots . . . . .	53
5.5	Reliability Analysis for M6 (placeholders) . . . . .	54
5.6	Ablation and Robustness . . . . .	54
5.6.1	Ablations . . . . .	54
5.6.2	Robustness checks . . . . .	54
<b>6</b>	<b>Conclusion and Forward Plan (Year 2)</b> . . . . .	<b>55</b>
6.1	Synthesis of Year-2 Contributions . . . . .	55

## CONTENTS

---

6.2	Conclusions . . . . .	56
6.3	Forward Plan (next 12–18 months) . . . . .	56
6.3.1	Objective O1: Complete the benchmark on M67 (all M1–M6) . . . . .	56
6.3.2	Objective O2: Integrate external factors and optimise parameters . . . . .	57
6.3.3	Objective O3: Reliability-oriented modelling . . . . .	57
6.3.4	Objective O4: Development of a New BPR Function . . . . .	57
6.4	Milestones, deliverables and risks . . . . .	58
6.4.1	Milestones . . . . .	58
6.4.2	Risks and mitigations . . . . .	58
6.5	Publications plan . . . . .	58

# Chapter 1

## Introduction

### 1.1 Background

#### 1.1.1 Urban Traffic Congestion: Challenges

Rapid economic development and accelerated urbanisation have led to city expansion and a sharp increase in motor vehicles. These changes contribute to complex urban transport problem , including severe traffic congestion, environmental pollution, and parking difficulties. These challenges primarily arise from a marked imbalance between the swift increase in transport demand and the constrained transport supply. On the supply side, many cities suffer from an insufficient road network density, which curtails the service capacity of the transport infrastructure. On the demand side, an irrational transport travel structure and outdated transport demand management contribute to a highly imbalanced demand distribution [11].

Travel time is a crucial metric for assessing road traffic conditions and is influenced by traffic demand and supply. Moreover, it dynamically changes with the evolving status of traffic congestion. Accurate travel time prediction is essential for travellers' route choices, traffic managers' congestion control, and traffic planners' traffic allocation. Thus, enhancing the precision of trip time estimation is vital for addressing traffic issues and optimising the operation of the entire urban traffic network.

Road impedance, which defines the relationship between travel time and traffic load on a road segment, plays a significant role in traffic allocation within a road network. Scholars globally have made significant strides in studying road impedance, notably the BPR impedance function developed by the U.S. [3]. The Federal Highway Administration has found widespread application in traffic allocation and road network planning. However, existing studies often overlook the effects of dynamic network demands such as queue overflow, weather conditions, peak hours, and weekdays when constructing impedance functions for congested road sections, leading to an underestimation of travel times.

This project aims to predict traffic flow and travel time by merging time series analysis with machine learning and statistical methods. Additionally, it integrates various dynamic condition parameters, such as weather and flooding, to develop a new impedance function for congested roadways. This function is designed to enhance the accuracy of travel time estimates across different conditions within the traffic network.

## Introduction

---

Plans are underway to develop an effective Demand-Responsive Road prediction model that will predict travel time more accurately, describe the arrival and departure patterns of vehicles under complex congestion conditions, and establish connections between multiple neighbouring congested road sections in the traffic network. Initially, the prediction model uses extensive data sets to assess vehicular congestion and its duration and pinpoint the areas and timing of queue overflow. This analysis provides a solid foundation for traffic departments to accurately evaluate congestion and implement data-driven congestion management strategies.

Furthermore, the model is adept at estimating travel times based on dynamic traffic conditions, thereby becoming a vital tool for evaluating travel costs in route selection and traffic planning. This model assists traffic managers in rapidly responding to changes in traffic conditions and implementing appropriate measures [12]. With the advent of urban traffic intelligence and intelligent highways, traffic departments can access high-quality real-time data and perform swift analyses.

This integrated approach seeks to alleviate current traffic congestion issues. It aims to provide a more scientific and systematic management and planning strategy for future urban transport systems, thus promoting sustainable development and efficient operation of urban transport.

### 1.1.2 BPR Function in Traffic Assignment

Accurately estimating travel time on road networks is a prerequisite for effective transportation planning, traffic management, and policy evaluation. At the heart of this estimation process, particularly within the context of static traffic assignment models [13], lies the volume-delay function (VDF), also known as the link performance function [14]. This function mathematically describes the relationship between the volume of traffic on a road segment and the time required to traverse it. Among the various VDFs proposed over the years, the formulation developed by the U.S. Bureau of Public Roads in 1964 [11], commonly known as the BPR function, has achieved and maintained a position of remarkable preeminence.

The canonical form of the BPR function is expressed as:

$$T = T_0 \left[ 1 + \alpha \left( \frac{V}{C} \right)^\beta \right]$$

Where:

- $T$  is the predicted travel time on the link at a given traffic volume.
- $T_0$  is the free-flow travel time, representing the time to traverse the link without any other traffic.
- $V$  is the traffic volume or flow on the link, typically measured in vehicles per hour.
- $C$  is the practical capacity of the link, also in vehicles per hour.
- $\alpha$  and  $\beta$  are dimensionless calibration parameters that shape the curvature of the function.

## Introduction

---

The widespread and enduring adoption of the BPR function can be attributed to several key factors. Its mathematical form is simple, convex, and monotonically increasing, which ensures that as traffic volume increases, so does travel time—an intuitive and necessary property for achieving stable equilibrium in traffic assignment models. Furthermore, its inputs—free-flow travel time, volume, and capacity—are conceptually straightforward and, at least historically, were considered relatively easy to measure or estimate from field data. Consequently, the BPR function has been embedded as the default VDF in various commercial and academic transportation planning software packages [15], solidifying its foundational role in countless global transport studies over the past half-century.

## 1.2 Project Aim

This project aims to optimise demand-responsive road networks through the development of next-generation Bureau of Public Roads (BPR) functions.

## 1.3 Project Objectives

- (1) **Year 1 & 2:** To benchmark the existing BPR function to systematically assess a wide array of factors and their impact on travel time prediction performance;
- (2) **Year 2:** To develop and evaluate new BPR functions using advanced statistical models to capture the complexities of traffic dynamics better and improve the accuracy of prediction results;
- (3) **Year 3:** To develop a demand-responsive road network scenario testing methodology.

## 1.4 Benchmarking

The overarching goal of this year is to develop and validate a classical BPR that provides a more accurate and robust estimation of travel times on the UK strategic road network than existing formulations. The research is structured around specific, measurable, and achievable objectives to achieve this aim.

The primary objectives of this research programme are:

1. To systematically review, analyse, and categorise the extensive body of literature on modifications to the BPR function to establish a comprehensive understanding of the current state-of-the-art and identify prevailing research trajectories.
2. To acquire, process, and conduct a comprehensive exploratory data analysis of high-resolution, multi-modal traffic data from a representative segment of the UK motorway network (the M67) to identify and quantify the key traffic dynamic features that a modern VDF must be capable of capturing.
3. To design and implement a rigorous, transparent, and multi-faceted benchmarking framework to quantitatively compare the performance of selected state-of-the-art VDFs from the literature using the empirical M67 dataset.

## Introduction

---

4. To propose, mathematically formulate, and calibrate a novel VDF that synthesises the strengths of existing models and is specifically tailored to leverage the richness of the available data, while maintaining physical interpretability.
5. To demonstrate the potential utility and superiority of the proposed VDF within a broader network-level traffic optimisation context, illustrating its practical value for transportation planning and management.

In the second year of this PhD research, the focus has been laying the groundwork for such improved volume-delay functions. The objectives for Year 2 include:

1. Conducting a *systematic literature review* to classify and analyse recent advancements in BPR-type travel time functions, thereby identifying methodological gaps;
2. Performing an *exploratory data analysis* of a real-world motorway dataset (M67 in the UK) to understand traffic patterns and provide empirical context for model development;
3. Formulating a *benchmarking methodology* by selecting a suite of BPR function variants from the literature (and conceptual clusters) to be tested on the dataset;
4. Outlining the structure for *comparative evaluation* of these variants, even if final results are pending calibration.

The Year 2 work bridges theoretical developments with practical data, providing the foundation for developing a novel BPR-based travel time function in the next phase.

In summary, this year's work is motivated by the need for volume-delay models that dynamically adjust to traffic conditions, incorporate multiple influencing factors, and leverage new data sources – ultimately improving travel time estimation accuracy for transportation planning and real-time applications.

## 1.5 Summary of Work Completed (Year 2)

### (1) Refined research question and scope

- Refined the overarching research question from '*Can the BPR be improved?*' to '*How can we design a demand-responsive, assignment-compatible link performance function that (i) captures peak/off-peak regime shifts, (ii) embeds heterogeneous fleets and exogenous factors, and (iii) quantifies uncertainty for reliability-oriented planning?*'
- Fixed the benchmarking target on six implementable method families (M1–M6) with explicit equations and constraints, enabling like-for-like comparison on the M67 dataset.

### (2) Systematic review and evidence synthesis

- Registered and executed a protocol-guided SLR across Scopus, Web of Science, Compendex, IEEE Xplore, ACM DL, and TRID (2010–2025).
- Screening flow completed: 3,378 → 2,031 (deduplicated) → 962 (relevance) → 394 (title/abstract) → 121 (PICOS) → 48 (included). PRISMA figure, summary

## Introduction

---

table, and reproducibility notes are provided in Chapter 2.

- Established the A–G taxonomy (dynamic parameterisation/structural optimisation; multi-class; external factors; data-driven enhancement; uncertainty; policy; data fusion) and validated the crosswalk to benchmarking families (M1–M6).
- Deep-read ten *high-priority* modelling papers; completed quality/risk-of-bias scoring; produced bibliometric visuals (yearly counts; topical proportions), an evidence map, and a Sankey diagram.

### (3) Data acquisition, pre-processing, and EDA (National Highways, M67, September 2024)

- Ingested 15-minute link observations (flow, speed, occupancy, travel time, vehicle classes); standardised timestamps, removed duplicates, and implemented gap handling (short-gap interpolation; long-gap censoring).
- Reconstructed or verified travel time  $T = \frac{3.6L}{v}$  for  $L = 2.713\text{ km}$ ; derived density  $\rho = q/v$ , volume-to-capacity  $V/C$ , and HGV share  $p_H$ .
- Estimated static baseline capacity  $C$  using an upper-envelope bend method and a peak-percentile method; defined  $V/C$  bands ( $< 0.6$ ,  $0.6\text{--}0.9$ ,  $\geq 0.9$ ); stratified summaries by weekday/weekend and peak/off-peak.
- Produced a complete EDA package, including flow–speed scatter plots,  $V/C$ -band boxplots, diurnal profiles, HGV residual analysis, and anomaly-marked time series, supported by a fully reproducible script.

### (4) Benchmarking design and performance metrics

- Finalised six benchmark families: M1 Dynamic-Parameter BPR (time-of-day/rolling), M2 Fundamental-Diagram-informed VDF (or piecewise BPR), M3 Multi-Class (HGV-adjusted) BPR, M4 External-Factor-adjusted BPR (weather-integrated), M5 Machine-Learning Hybrid BPR, and M6 Stochastic-Capacity/ Reliability BPR.
- Defined comparability constraints: identical blocked train/validation/test splits; aligned inputs (core  $V, C, t$ ; plus  $p_H$  for M3; weather states for M4); shared objectives and monotonicity/positivity constraints.
- Fixed the metric suite: RMSE, MAE, MAPE,  $R^2$ ; reliability metrics (quantile MAE and coverage  $\text{Cov}^{(p)}$ ); regime-wise reporting by  $V/C$  band and peak/off-peak; relative improvement  $\Delta\%$  against the classical BPR baseline.

### (5) Model implementation and calibration tooling

- Implemented a log-linear baseline calibration for BPR ( $T/T_0 - 1 = \alpha x^\beta$ ) with residual diagnostics; built a two-regime piecewise BPR (threshold search + dual fits).
- Developed M3’s HGV equivalent-volume adjustment  $V_{\text{adj}} = V(1+e_H p_H)$  and linked calibration; prepared M4’s weather merge and parameterisation (multipliers on  $v_f$  and/or  $C$ ).

## Introduction

---

- Drafted M6's stochastic-capacity / quantile-fitting framework (joint mean–tail objectives); prototyped a lightweight ANN/GBR for M5 with monotonicity checks and blocked time splits.
- Aligned version control, figure export, and parameter logs with the Overleaf structure (labels, placeholders, and chapter cross-references).

### (6) External data integration (in progress)

- Designed field mapping and spatio-temporal matching to Met Office weather (precipitation intensity, visibility, wind, temperature); scoped National Highways incident feeds and licensing.
- Defined a minimal viable external-factor set (e.g., rain/heavy-rain binary; visibility tiers) for the first M4 experiments; planned robust merging and sensitivity tests for inconsistent records.

### (7) Preliminary empirical insights from EDA

- Clear weekday bimodal demand with flatter weekends; speeds decline nonlinearly with rising flows and drop sharply near capacity—supporting time-regime or piecewise structures in M1/M2.
- Low but time-varying  $p_H$  suggests targeted benefits for M3 in certain periods; enlarged delay tails at high  $V/C$  motivate reliability metrics and M6's stochastic treatment.

### (8) Writing, packaging, and dissemination

- Completed Chapter 2's methodological re-organisation with visual evidence (PRISMA, crosswalk, evidence map, Sankey); assembled Chapter 3's data and EDA content; specified Chapter 4's equations, parameters, and calibration protocols for M1–M6.
- Harmonised figure/table styling and British English usage; consolidated references and .bib entries for numbered citations and searchability.

### (9) Risks and mitigations

- *Limited high-saturation samples*: mitigate via piecewise models and extended months; report conservative uncertainty bands in high  $V/C$ .
- *Low variance in  $p_H$* : use literature priors and sensitivity ranges for  $e_H$  to avoid overfitting.
- *Bias in reconstructed  $T$* : interpret tails/peaks cautiously; pair reliability results with confidence statements.
- *External-data alignment*: use blocked time windows and site-level matching; fully log merge rules for audit.

### (10) Concrete outputs (this year)

## Introduction

---

- A reproducible data-clean/EDA codebase (with automatic figure exports and parameter logs).
- Structured extraction tables, quality scores, and taxonomy mapping for 48 core studies; deep-read notes for the ten benchmark papers.
- Mathematical definitions, constraints, calibration, and evaluation plans for six benchmark families (M1–M6) with a unified metric protocol.
- Overleaf/repository scaffolding, chapter text, and formatting templates; all figure/table placeholders and labels in place.

# Chapter 2

## Literature Review

### 2.1 Framework of the Systematic Review

#### 2.1.1 Review Questions and Scope

This review addresses three questions that steer the benchmarking design and, ultimately, the proposal of a next-generation BPR function:

- RQ1:** Which clusters of modifications to the classical BPR/volume–delay function (VDF) have emerged since 2010, and how are they positioned concerning congestion regimes, parameter dynamics, and factor integration?
- RQ2:** What datasets and calibration practices under these modifications, and how do studies report predictive performance and uncertainty?
- RQ3:** Where are the persistent gaps (e.g. oversaturation handling, peak/off-peak regime shifts, exogenous factors, and uncertainty quantification) to predict travel time, and how should these inform a fair benchmarking protocol?

#### 2.1.2 Protocol and Reporting Standards

This benchmarking study follows the *Preferred Reporting Items for Systematic Reviews and Meta-Analyses* (PRISMA) guidance to ensure transparency and reproducibility, adapting the protocol to a transport-modelling context [16]. Eligibility criteria are defined using the PICOS framework to formalise Population, Intervention, Comparator, Outcomes, and Study design [17, 18]. Bibliometric descriptors (yearly volume and topical proportions) are reported to contextualise the evidence base [16]. This structure mirrors best practice observed in recent transport-focused systematic reviews and prepares a clean interface to downstream benchmarking (cf. Section 2.3) [16].

#### 2.1.3 Information Sources and Search Strategy

Searches covered 2010–2025 in major databases (Scopus, Web of Science, Compendex/Inspec via Engineering Village, IEEE Xplore, ACM DL, TRID, ScienceDirect), supplemented by Google Scholar and institutional libraries. Two keyword classes were combined with Boolean operators (**AND/OR**), following the practice of transport safety SLRs [17]:

- **Function class:** “BPR” OR “volume-delay function” OR “link performance function” OR “travel time function” OR “VDF”.
- **Improvement class:** “modif\*” OR “enhanc\*” OR “piecewise” OR “dynamic parameter\*” OR “uncertainty” OR “stochastic” OR “weather” OR “heavy vehicle\*” OR “capacity degradation” OR “machine learning” OR “hybrid” OR “Bayesian”.

Snowballing from reference lists and citing articles was used to capture additional items [18].

### 2.1.4 Eligibility Criteria (PICOS)

The PICOS framework is a widely recognised tool for structuring a systematic literature review, ensuring that the search and selection process is focused and consistent. It defines the key components of the research question. For this review on BPR function modifications [15], the framework was tailored as follows:

- **P – Population/Problem:** The ‘population’ in this context refers to the objects of study, which are road networks or individual road links/segments (e.g., motorways, urban arterials) where traffic congestion and travel time are a problem. The core problem is the inadequacy of the standard BPR function in accurately modelling link performance under various traffic conditions [19].
- **I – Intervention:** The ‘intervention’ represents the primary subject of our investigation. This includes any modification, enhancement, or alternative formulation of the standard BPR function. Examples range from recalibration of its parameters to fundamental structural changes, such as incorporating new variables (e.g., traffic density, vehicle classification) or adopting entirely new functional forms (e.g., piecewise functions, stochastic models).
- **C – Comparison:** The ‘comparison’ is the benchmark against which the intervention is measured. In most cases, this is the standard, original BPR function. In some studies, the comparison might be another modified function or empirical, real-world data. We sought papers that explicitly compared their proposed model against a known baseline.
- **O – Outcome:** The desired ‘outcomes’ are the metrics used to evaluate the intervention’s success. These are primarily focused on the improved accuracy of travel time or delay estimation. Key outcome measures include statistical indicators such as Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and the Coefficient of Determination ( $R^2$ ), as well as improved theoretical consistency with traffic flow principles.
- **S – Study Design:** This criterion defines the types of research we considered valid. We included empirical studies based on real-world or simulated data, methodological papers proposing new functional forms, and comparative benchmarking studies. Purely theoretical or conceptual papers without any form of validation were excluded [20].

Table 2.1 presents the PICOS design guiding inclusion and exclusion. Studies must explicitly propose, calibrate, or validate a BPR/VDF *modification*, report a travel-time/impedance outcome, and use empirical or high-fidelity simulated data.

## Literature Review

---

TABLE 2.1 PICOS framework and operational eligibility rules for the BPR modification review.

<b>Population</b>	Road links/corridors and networks (urban arterials, motorways); studies in English, 2010–2025.
<b>Intervention</b>	Any modification to classical BPR/VDF: (i) mathematical structure (piecewise, higher-order), (ii) dynamic parameters, (iii) factor expansion (HGV share, gradient, weather, incidents), (iv) ML/hybrid fitting, (v) uncertainty modelling, (vi) external/policy couplings, (vii) data fusion.
<b>Comparator</b>	Classical BPR or other VDF baselines; cross-model comparisons within study.
<b>Outcomes</b>	Link travel time or speed; calibration fit (RMSE/MAE/MAPE/ $R^2$ /NRMSE); robustness/uncertainty where available.
<b>Study design</b>	Peer-reviewed journal or conference papers with empirical or high-fidelity simulation validation. Exclude purely demand-only models without VDF; these are excluded if the journal version exists.

### 2.1.5 Screening, Deduplication, and Flow of Studies

The five-stage funnel produced a final corpus of 48 papers supporting the synthesis and benchmarking design. Counts at each stage are given in Table 2.2 (also visualised via a minimalist PRISMA diagram in Fig. 2.1).

TABLE 2.2 Summary of the systematic filtering pipeline (counts reflect your executed search).

Step	Purpose	Databases & Keywords	Inclusion/Exclusion Criteria	No.
1	Initial identification	Scopus, WoS, IEEE Xplore, TRID, etc.; broad BPR/VDF + improvement terms	Peer-reviewed; English	3378 → 2031
2	Relevance filtering	Add “piecewise”, “dynamic”, “uncertainty”, “ML”	Applied to 2031 uniques	962
3	Manual screening	Titles/abstracts	Explicitly modify BPR/VDF	394
4	Specific criteria	Full abstract vs. PICOS	Must improve VDF & report outcomes with data	121
5	Full-text review	Method/validation; reference tracing	Novelty, rigour, interpretability	48

## Literature Review

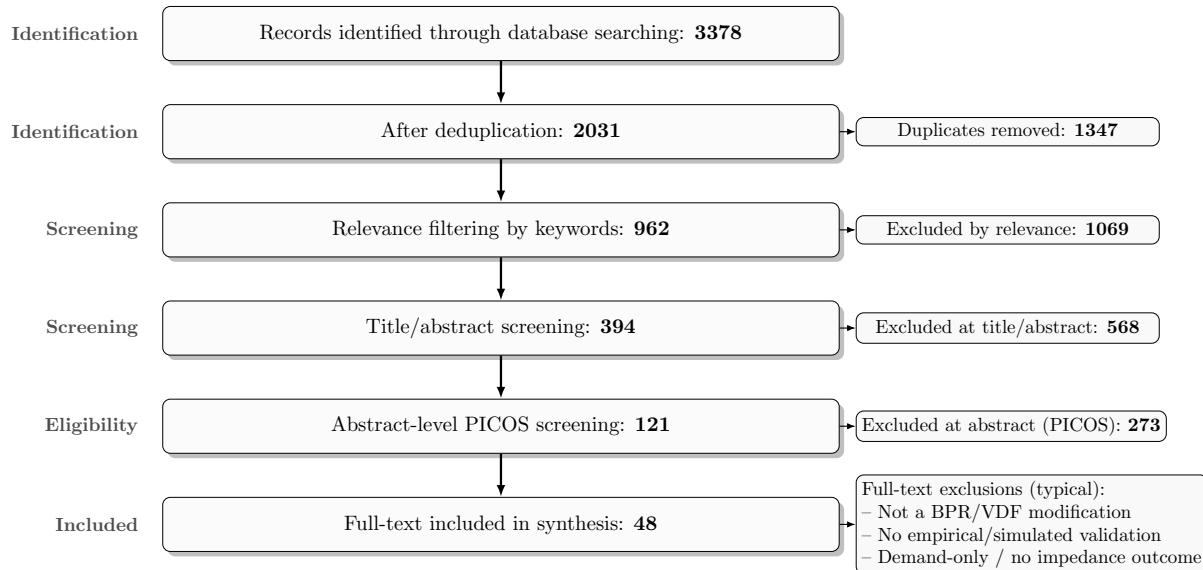


FIGURE 2.1 PRISMA-style flow for the systematic review (auto-fitted to A4 text width).

### 2.1.6 Data Extraction and Quality

For each study, we extract:

- Data source and resolution
- Road type and geography
- Prediction or Modification Model
- Functional form and parameters
- Calibration method (e.g. OLS/NLS, robust loss, Bayesian, regularised regression, ML)
- Exogenous factors considered (HGV share, gradient, weather, incidents)
- Performance metrics and validation design
- Uncertainty reporting (intervals, bootstrap, Bayesian posteriors)
- Code/data availability
- Paper quality and SJR Ranking
- Journal name

These quality appraisal records (yes/no/partially) are kept for sampling transparency, calibration diagnostics, overfitting control [21], and external validity. This mirrors how transport SLRs pair PRISMA with method-aware appraisal rather than relying only on bibliometric summaries [18].

## 2.2 BPR Modification Cluster

### 2.2.1 Cluster Method

The literature on volume–delay functions (VDFs) spans a wide range of aims—from altering the *shape* of link travel-time functions to improving calibration practice and data integration [22, 23]. A taxonomy is therefore required for three reasons:

- (i) **Construct validity.** We separate inventions that truly change the *functional form or its parameters* (hence affect assignment physics) from enablers such as data fusion or policy scenarios that do not define a new VDF shape [11, 22, 24]. This prevents mixing distinct constructs and ensures that like is compared with like.
- (ii) **Identifiability and deployability.** Families are defined so that each class can be *identified* with the available M67 data and *deployed* in standard network loading/assignment. This avoids clusters that cannot be calibrated or converge in M67 data assignment [6].
- (iii) **Comparability.** The taxonomy underpins a fair benchmarking protocol: models within different families can be calibrated under identical splits, objective functions, and metrics. Families that would otherwise conflate exogenous context with VDF shape are decoupled, enabling clean attribution of performance gains [25, 26].

In short, the taxonomy is not merely descriptive. It operationalises the review into families with distinct causal levers and practical calibration pathways.

### 2.2.2 Operational definitions and decision rules (A–G)

Let a candidate study be denoted by  $\mathcal{S}$  with reported formulation  $\mathcal{M}$ , data  $\mathcal{D}$ , and estimation protocol  $\mathcal{E}$ . We classify  $\mathcal{S}$  into exactly one of the seven review families (A–G) using the following predicates [27]. Each predicate is a linguistic cue in the paper (what the authors claim) and a mathematical cue (what the equations do).

#### (1) Cluster A: Dynamic parameterisation & structural optimisation

*Predicate A* holds if  $\mathcal{M}$  modifies the *shape* or *parameters* of BPR-type VDF in a way that is regime- or context-specific, or replaces the polynomial kernel with a piecewise/thresholded kernel [28, 29]:

$$T = T_0 \left[ 1 + \alpha(\mathbf{x}) \left( \frac{V}{C(\mathbf{x})} \right)^{\beta(\mathbf{x})} \right], \quad \text{with } \partial\alpha/\partial\mathbf{x} \neq 0 \text{ or } \partial\beta/\partial\mathbf{x} \neq 0, \quad (2.1)$$

OR BRP function can be a piecewise version with a threshold  $\tau$  ensuring  $C^0$  continuity at  $V/C = \tau$ .

**Signals:** “piecewise” [30], “regime-specific” [26], “time-of-day parameters”, “oversaturation exponent”.

#### (2) Cluster B: Multi-class

*Predicate B* holds if  $\mathcal{M}$  incorporates the composition of the vehicle class via the passenger-car equivalency (PCE) or capacity reduction [31]:

$$V_{\text{eq}} = \sum_k \text{PCE}_k V_k \quad (2.2)$$

OR

$$C_{\text{eff}} = C (1 - \lambda s_{\text{HGV}}), \quad (2.3)$$

The  $V_{\text{eq}}$  and  $C_{\text{eff}}$  will be embedded into a BPR function.

**Signals:** “HGV share” [32], “PCE”, “effective capacity by vehicle class”.

### (3) Cluster C: External-factor integration (environment)

*Predicate C* holds if exogenous factors (weather, visibility, incidents, lane closures, gradient) *explicitly* modify free-flow speed or capacity through multiplicative factors or parameter modulation [26, 33]:

$$T = \frac{L}{v_f \cdot \text{SAF}(w)} \left[ 1 + \alpha \left( \frac{V}{C \cdot \text{CAF}(w)} \right)^\beta \right] \quad (2.4)$$

OR

$$\alpha = \exp(\eta_0 + \boldsymbol{\eta}^\top \mathbf{z}), \quad \beta = \exp(\gamma_0 + \boldsymbol{\gamma}^\top \mathbf{z}). \quad (2.5)$$

**Signals:** “weather adjustment”, “incident factor”, “capacity multiplier”, “slope/grade adjustment”.

### (4) Cluster D: Machine learning and data-driven enhancement

*Predicate D* holds if  $\mathcal{M}$  uses ML to (i) replace the kernel or (ii) augment it by residual/parameter mapping while preserving deployability in assignment [32, 34, 35]:

$$T = T_{\text{BPR}} \cdot (1 + r(\mathbf{x})) \quad (2.6)$$

OR

$$(\alpha, \beta) = \exp\{g_\alpha(\mathbf{x}), g_\beta(\mathbf{x})\}. \quad (2.7)$$

**Signals:** “neural network”, “gradient boosting”, “hybrid BPR”, “residual learning”.

### (5) Cluster E: Uncertainty & stochastic reliability

*Predicate E* holds if  $\mathcal{M}$  targets travel-time distributions or quantiles rather than means, e.g. quantile-BPR or lognormal bands [3, 22, 24, 36, 37]:

$$\min_{\alpha^{(p)}, \beta^{(p)}} \sum_t \rho_p(T_t - T^{(p)}(V_t)), \quad T^{(p)} = T_{\text{det}} \exp(z_p \sigma). \quad (2.8)$$

**Signals:** “percentile”, “reliability”, “stochastic capacity”, “coverage”.

### (6) Cluster F: Policy & emerging technology factors

*Predicate F* holds if the contribution changes generalised cost (pricing, control, CAV penetration) *without proposing a new link BPR shape*.

**Signals:** “tolling” [34], “managed lanes” [38], “automation share” [39].

### (7) Cluster G: Multi-source data fusion for calibration

*Predicate G* holds if the novelty lies in the *estimation pipeline* (loops + probe + event/weather logs, transfer learning), not in the BPR kernel.

*Signals:* “sensor fusion” [40], “trajectory–loop fusion” [41], “transfer calibration” [42].

#### 2.2.3 Rule precedence and edge cases

Because some studies or papers touch multiple clusters, this study applies a precedence order aligned with construct purity:

$$\begin{aligned} \text{A/B/C/E (kernel-shape or parameterisation)} \\ \succ \text{ D (data-driven)} \\ \succ \text{ F (data fusion)} \\ \succ \text{ G (policy)} \end{aligned}$$

Examples:

- A study that is piecewise BPR (A) *and* uses ML only to tune parameters is labelled **A** (primary), not **D**.
- A study that introduces weather CAF/SAF (C) and then quantile-fits (E) is labelled by the *declared objective*: if reliability is central, **E**; otherwise **C**.
- If only data sources are richer but the kernel is classical BPR, it is **F**, not A.

#### 2.2.4 Comparability and identifiability tests

Before admitting a candidate model into benchmarking, we apply:

- (a) **Identifiability test:** parameters are estimable on M67 September data without pathological collinearity; positivity/monotonicity constraints can be enforced ( $\alpha > 0, \beta > 1; v_f, \rho_j, w > 0$ ).
- (b) **Deployability test:** the model converges within standard static assignment and yields monotone link travel time in  $V/C$  on the admissible range.
- (c) **Comparability test:** inputs are aligned across families (core at least  $V, C$ , period; plus HGV share for F2; weather states for F6), with identical splits, objective functions, and metrics.

#### 2.2.5 Reproducible labelling protocol

To ensure that the taxonomy is auditably reproducible:

- (1) Two reviewers independently screen full texts and extract: kernel form, parameter set, exogenous factors, data sources, and estimation method.
- (2) Predicates (A–G) are ticked using the decision rules in Section 2.2.2; rule precedence (Section 2.2.3) resolves overlaps.

## Literature Review

---

- (3) Disagreements are reconciled; the final clusters label and rationale are recorded in a data-extraction sheet.
- (4) Each study receives a *confidence score* (high/medium/low) based on explicitness of equations and availability of calibration details.

### 2.2.6 Evidence Map

Figure 2.2 shows how the seven review clusters (A–G) consolidate into six benchmarkable methods (M1–M6). Cluster A contributes the largest share (39.4%) and bifurcates into M1: **DP–BPR** (23.1%) and M2: **FD–VDF** (16.2%), reflecting two distinct mathematical lines: (i) time-varying parameters in a BPR form and (ii) fundamental-diagram/density-informed structures. Cluster B (13.9%), C (10.6%), D (13.9%) and E (13.0%) map directly to M3: **MC–BPR**, M4: **EF–BPR**, M5: **ML–hBPR** and M6: **SC–BPR**, respectively. Sparse strands from F and G (each 4.6%) and miscellaneous items form the *Others* stream (9.3%). Overall, **90.7%** of the included, high-quality studies feed into M1–M6, validating our benchmark scope and motivating the split of Cluster A for methodological comparability and parameter identifiability.

Sankey (Two-Column) — Evidence Flow from Review Families (A–G) to Benchmark Methods (M1–M6)

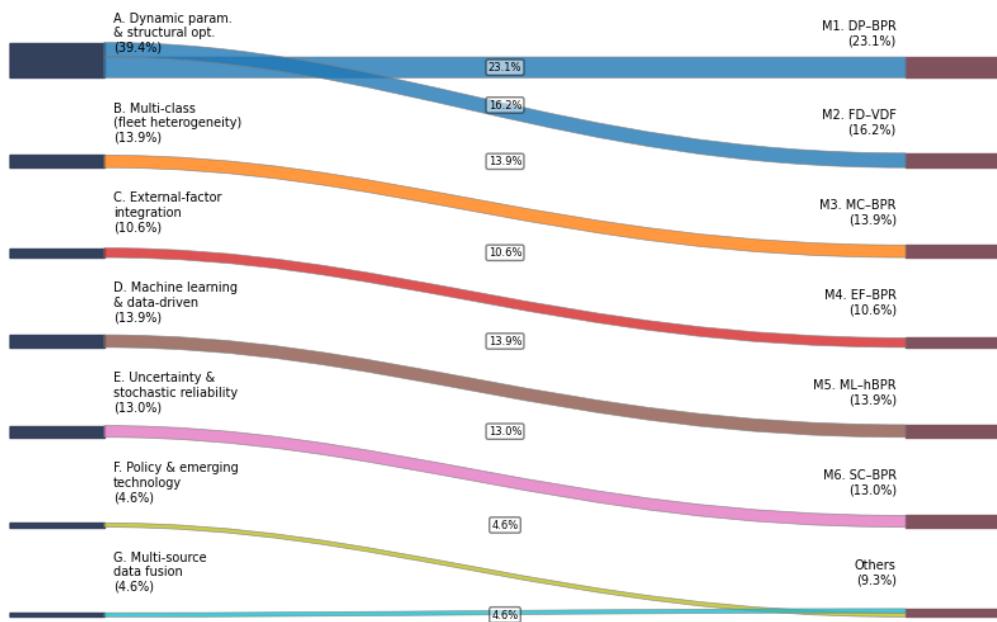


FIGURE 2.2 Evidence Flow from Review Cluster (A–G) to Benchmark Methods (M1–M6)

## 2.3 Performance Metrics and Validation

A balanced metric set is required for fair benchmarking across regimes. We adopt:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{T}_i - T_i)^2}, \quad \text{MAE} = \frac{1}{n} \sum_{i=1}^n |\hat{T}_i - T_i|, \quad (2.9)$$

$$\text{MAPE} = \frac{100}{n} \sum_{i=1}^n \left| \frac{\hat{T}_i - T_i}{T_i} \right|, \quad R^2 = 1 - \frac{\sum(\hat{T}_i - T_i)^2}{\sum(T_i - \bar{T})^2}. \quad (2.10)$$

Regime-aware reporting is essential: (i) low-delay (free/near-free flow), (ii) rising congestion, (iii) oversaturation (queue spillback). Where studies report uncertainty, interval coverage and width are recorded. This metric design dovetails with benchmarking standards in transport modelling papers that separate *criteria* and *scenarios* [18]. Given the known sensitivity of network models to BPR parameter uncertainty, we will additionally record robustness diagnostics (bootstrap/Bayesian intervals) for  $(\alpha, \beta)$  where applicable [?, 43].

## 2.4 Synthesis and Knowledge Gaps

Narrative synthesis is organised by families (A–G) and mapped to factors (HGV share, gradient, weather, incidents, temporal regimes). Evidence is cross-walked to data contexts (urban/motorway, country/region), calibration practice, and uncertainty treatment. Recurrent gaps include:

- Limited treatment of oversaturation and capacity drop in static BPR forms;
- Fixed parameters ignoring peak/off-peak or lane/grade heterogeneity;
- Exogenous factors integrated ad hoc rather than with structure-preserving parameterisation;
- Scarce uncertainty quantification;
- Weak external validation.

These echo broader lessons from recent transport SLRs which emphasise transparency, transferability, and rigorous benchmarking [16–18].

Across these seven clusters, the literature reveals various approaches to improving the BPR function. However, a few overarching gaps stand out:

### (1) Unified Model vs. Specialized Models

Most enhancements tackle one or two isolated issues (e.g., heavy vehicles, weather, or dynamic adjustment). There is a lack of a *unified framework* that can incorporate *multiple factors simultaneously* – for example, a model that is both dynamic *and* multi-class *and* stochastic. A key challenge is how to synthesise the “best of each” approach without over-complicating the model [11, 14, 27].

### (2) Complexity vs. Applicability

Adding layers such as dynamic updates, ML components, or stochastic terms increases the risk of models becoming too complex for practical planning. Several studies achieved accuracy gains, but at the cost of requiring extensive data or computation (e.g. training a deep network or running Monte Carlo simulations). There is a research gap in simplifying these advanced models or developing *analytical approximations* so that enhanced BPR functions can be deployed in standard transport planning software and real-time systems.

### (3) Generalizability

Many improved functions are context-specific. For instance, parameters derived for Indonesian urban roads[21][22], or a weather factor derived in one city, may not apply elsewhere. A gap exists in understanding *transferability* – under what conditions can a calibrated BPR function or parameter set be transferred, and how to adjust it for a new context with minimal new data [15, 44–46].

### (4) Validation and Real-World Implementation

While numerous papers demonstrate improvements on historical data, few document the *real-world implementation* of next-gen BPR functions in operational planning models or traffic management systems. There is a need for case studies showing how these new models impact network assignment results, travel forecasts, or management strategies compared to the classical approach. This would help quantify the practical benefit (e.g. improved demand assignment convergence, more accurate benefit-cost analysis due to better travel time estimates, etc.) [37, 47, 48].

These gaps inform the direction of the present research. In Year 2, having surveyed these approaches, we now test a representative set on a common dataset. By benchmarking multiple methods (something not extensively done in the literature), we aim to assess which techniques or combinations yield the most significant improvements and at what cost. The following sections describe the data used for this evaluation, the chosen benchmark methods (spanning clusters A–G), and the analysis plan. This will ultimately guide the proposal of a new or hybrid BPR formulation in Year 3 that attempts to bridge several gaps [4, 49, 50].

# Chapter 3

## Data Exploratory Analysis (EDA)

This chapter presents a comprehensive and in-depth quantitative analysis of the traffic dynamics on the M67 motorway's westbound carriageway, between Junction 4 and Junction 3, during September 2024. The analysis is founded on high-resolution sensor data, captured at 15-minute intervals, provided by the National Traffic Information Service (NTIS). The findings reveal the complex operational characteristics of this section as a critical urban transport artery, with traffic patterns heavily influenced by commuter behaviour, vehicle composition, and persistent roadworks.

### 3.1 Data source

This study uses high-resolution observations from **National Highways** for a single motorway link on the M67 corridor—the pre-processed file provided (September 2024) aggregates detector readings to *15-minute* intervals. Unless stated otherwise, the analysis considers *all days in September 2024* and distinguishes *weekdays* from *weekends*. The target link (internal code: 115030402) has an engineering length of  $L = 2.713\text{ km}$  (as supplied with the dataset). Detector streams include *per-lane* measures; in this chapter, we report both lane-aggregated and class-segmented summaries where relevant.

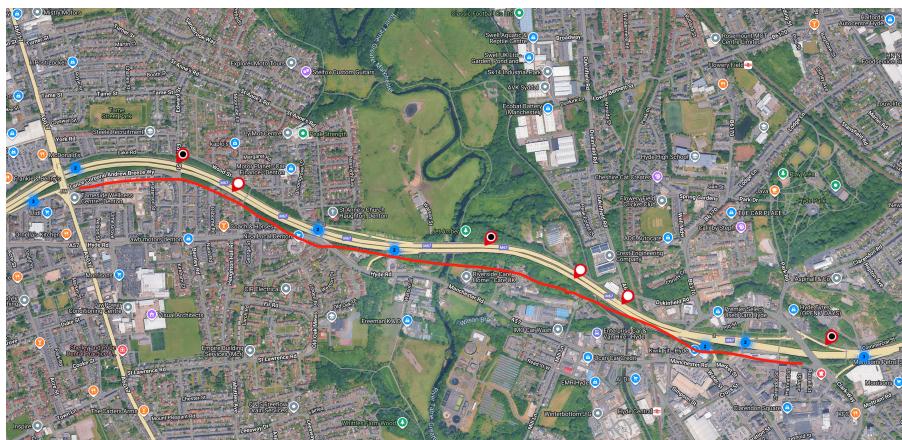


FIGURE 3.1 M67 Segement Map



FIGURE 3.2 Lane detail On M67

### (1) Temporal Granularity

The raw data is provided in 15-minute intervals. The data were aggregated to hourly, daily, and day-type (weekday/weekend) levels to reveal traffic patterns at different time scales. For instance, hourly flow was calculated by summing the four 15-minute flow values within a given hour. Timestamps are in local time (BST) and include the interval's end time unless explicitly noted.

### (2) Vehicle classes

The dataset distinguishes at least four broad categories: *light vehicles* (cars/vans) and *heavy goods vehicles* (HGV). Let  $p_H$  denote the HGV share per interval; they can be rolled up consistently to the two-class partition used in Chapter 2. And the vehicle classes shown in Table 3.1.

TABLE 3.1 Length-based vehicle categories used in the M67 dataset

Category	Length criterion	Indicative examples
1	$\leq 5.2 \text{ m}$	Cars, small vans
2	$> 5.2 \text{ m}$ and $\leq 6.6 \text{ m}$	Large vans, minibuses
3	$> 6.6 \text{ m}$ and $\leq 11.6 \text{ m}$	Coaches, medium rigid HGVs
4	$> 11.6 \text{ m}$	Articulated HGVs / long vehicles

### (3) Lane-Specific Differences in Flow and Speed

In the United Kingdom, the Highway Code requires vehicles to keep to the nearside (left-hand) lane except when overtaking. This rule is intended to maximise road capacity. Analysing the flow and speed data for each lane makes it possible to assess compliance with this rule and the functional differentiation between lanes.

## Data Exploratory Analysis (EDA)

---

The analysis shows that during non-congested periods, the operational characteristics of the three lanes differ significantly:

- Lane 1 (nearside/slow lane): This lane has the lowest average speed, typically 10-15 km/h lower than Lane 3. Its traffic composition has the highest proportion of Heavy Goods Vehicles (HGVs).
- Lane 2 (middle lane): This lane usually carries the highest traffic flow of the three lanes. Its average speed is intermediate between that of Lane 1 and Lane 3.
- Lane 3 (offside/overtaking lane): This lane has the highest average speed, approaching the free-flow speed. In free-flowing conditions, its flow is typically the lowest, consistent with its designated function as an overtaking lane.

## 3.2 Variable dictionary and notation

### (1) Variable Dictionary

Field Definitions: To ensure clarity and precision, the key data fields used in this analysis are defined as follows:

- **Local Time:** Timestamps for each data record, precise to the second, forming the basis of all time-series analysis.
- **Local Time:** Timestamps for each data record, precise to the second, forming the basis of all time-series analysis.
- **Day Type ID:** A categorical variable representing the day of the week. The correct interpretation of this field is critical for distinguishing between weekday and weekend traffic patterns. In the dataset with a standard calendar, a reliable correspondence was established. For example, the data show that Monday, 2 September 2024, has a Day Type ID of 0, while Tuesday, 3 September 2024, has a Day Type ID of 1. By extension, the following mapping was deduced

TABLE 3.2 Day type description

Daytype	Day	Value
0	Monday	First working day of normal week
1	Tuesday	Normal working Tuesday
2	Wednesday	Normal working Wednesday
3	Thursday	Normal working Thursday
4	Friday	Last working day of normal week
5	Saturday	Saturday but excluding days falling within Christmas Day/New Year's Day
6	Sunday	Sunday but excluding days falling within Christmas Day/New Year's Day
7	/	First day of school holidays
9	/	Middle of week - school holidays,
11	/	Last day of week - school holidays
12	/	Bank Holidays, including Good Friday, but excluding days falling within Christmas Day/New Year's Day

## Data Exploratory Analysis (EDA)

---

- **Total Traffic Flow:** The total number of vehicles passing the monitoring point within a 15-minute statistical period. This is the core metric for measuring traffic demand.
- **Fused Average Speed:** The average travel time in seconds vehicles have taken to traverse the entire length of the NTIS Link over the 15-minute time slice. The harmonically-averaged mean speed across all lanes, measured in km/h. This metric provides a robust reflection of the macroscopic operational state of the road section, particularly under congested conditions, where it is more stable than an arithmetic mean.
- **Fused Travel Time:** The harmonically-averaged time taken for a vehicle to traverse the section, measured in seconds. This is inversely proportional to the average speed and is a direct measure of travel efficiency and delay.
- **Lane-specific metrics** (e.g., FlowLane1Category1Value, AverageSpeedLane1Value): Provide disaggregated flow and average speed data for each of the three lanes, forming the basis for microscopic traffic behaviour analysis.
- **Data quality flags** (e.g., IsDataErrorFlowLane1Category1Flag): Binary flags (0 or 1) used to identify data points with quality issues.
- **Traffic Flow %value1:** The percentage of traffic flow for vehicle length category.
- **Traffic Flow %value2:** The percentage of traffic flow for vehicle length category.
- **Traffic Flow %value3:** The percentage of traffic flow for vehicle length category.

### (2) Variable Notation

Table 3.3 defines the core variables, symbols, and units used throughout the thesis. All statistics in this chapter follow the symbol conventions below.

TABLE 3.3 Core variables, notation and units.

Variable (dataset)	Symbol	Unit	Notes
Total flow (15-min)	$Q$	veh/15 min	The number of vehicles detected on any lane within the 15-minute time slice.
Flow rate (per hour)	$q$	veh/h	$q = 4Q$ . Lane-aggregated unless stated.
Average speed	$v$	km/h	Interval mean over vehicles.
Headway (mean)	$h$	s/veh	Reciprocal of per-lane flow rate up to scaling. Used as a quality cross-check.
Occupancy	$\text{occ}$	%	Fraction of time a detector is occupied. Proxy for density.
Travel time (observed)	$T$	s	Link travel time derived by the data provider or reconstructed as $T = L/(v \cdot 1000/3600)$ .
Capacity (static baseline)	$C$	veh/h	Empirical estimate; see Section 3.4.
Volume-to-capacity ratio	$x = V/C$	—	Here $V \equiv q$ .
Density (derived)	$\rho$	veh/km	$\rho = q/v$ (unit-consistent).
HGV share	$p_H$	—	Proportion of heavy vehicles in interval volume.

### 3.3 Pre-processing and Quality Control

The raw feed contains routine artefacts (clock drift, detector dropouts, spurious spikes). We apply a reproducible pipeline to ensure robust inputs for calibration and benchmarking.

#### 3.3.1 Data Cleansing

The first step was data cleansing. Using the `IsModelError...` series of quality flags provided in the dataset, all records marked as erroneous (value of 1) were systematically identified and excluded. This step ensures the accuracy and reliability of subsequent statistical analysis, preventing biases arising from sensor malfunctions or data transmission errors.

#### 3.3.2 Time alignment and deduplication

The raw data is provided in 15-minute intervals. The data were aggregated to hourly, daily, and day-type (weekday/weekend) levels to reveal traffic patterns at different time scales. For instance, hourly flow was calculated by summing the four 15-minute flow values within a given hour.

- **Resampling:** all streams are aligned to 15-min bins using end timestamps. If the provider supplies rolling windows, values are re-binned by weighted averaging

## Data Exploratory Analysis (EDA)

(speed) or summation (flows).

- **Deduplication:** duplicate rows (identical link, timestamp) are dropped, keeping the first occurrence when measures coincide; otherwise, a simple average is used after flagging.

### 3.3.3 Missing data treatment

Let  $\mathcal{I}$  denote the index set of time intervals. For variable  $Z \in \{q, v, T, \text{occ}\}$ , define completeness as

$$\text{Comp}(Z) = \frac{1}{|\mathcal{I}|} \sum_{t \in \mathcal{I}} \mathbb{I}\{\text{$Z_t$ is observed}\}. \quad (3.1)$$

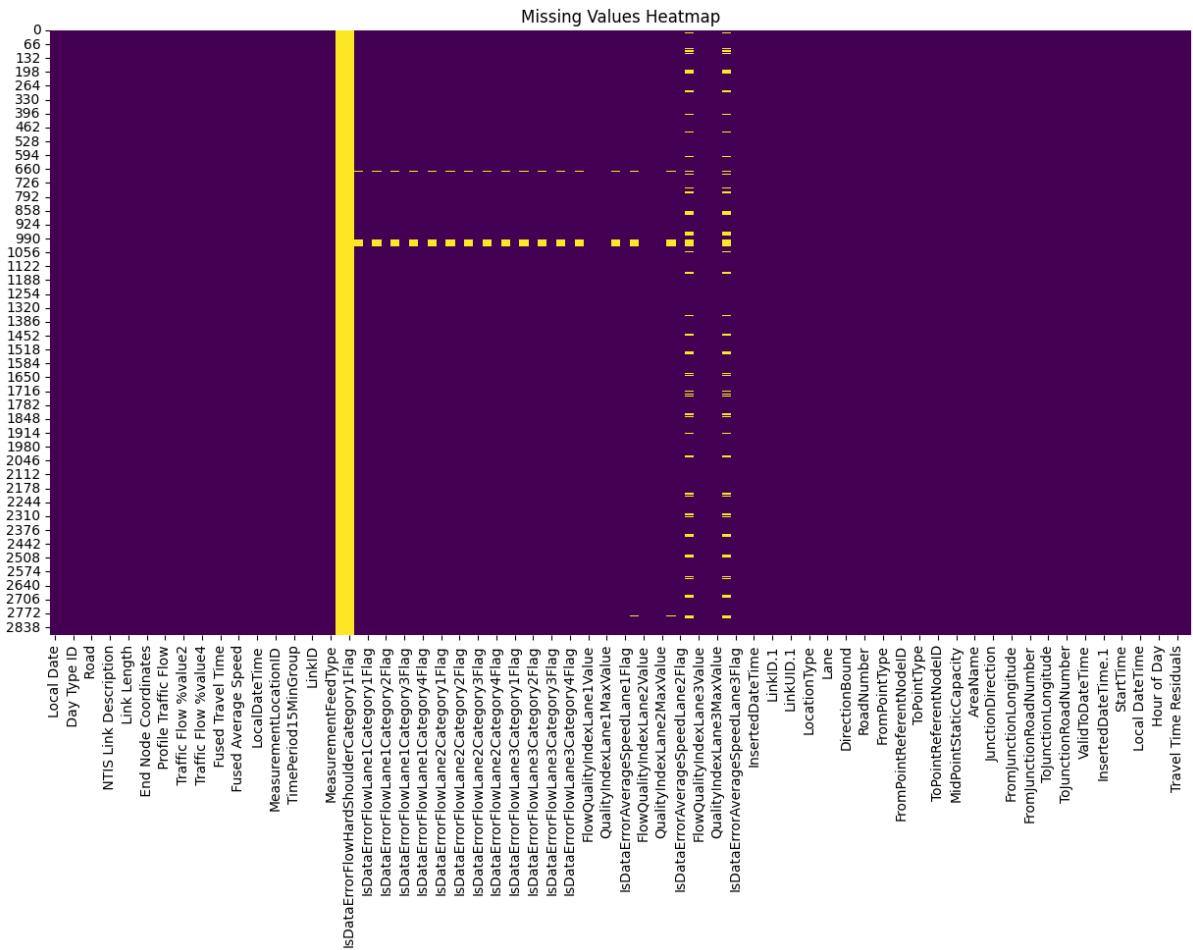


FIGURE 3.3 Missing Value Heatmap

Missing values are imputed only for EDA visual continuity (never for model estimation) using:

- **Short gaps ( $\leq 30$  min):** linear interpolation in time for  $v$ ; proportional scaling for  $q$  based on adjacent intervals.
- **Long gaps:** left as missing and excluded from model fitting.

### 3.3.4 Outlier detection and robust filtering

Spurious spikes are common around outages or adverse weather. We employ robust rules:

- **Speed bounds:**  $v \in [10, 130]$  km/h. Points outside are flagged.
- **Flow bounds:** lane-normalised flow  $q_\ell \leq 2400$  veh/h/lane (engineering plausibility); link total scaled by lane count.
- **Hampel filter on residuals:** compute residuals  $r_t = T_t - \tilde{T}_t$  relative to a smooth baseline  $\tilde{T}$  (7-point median). Flag  $|r_t| > 3\sigma$  where  $\sigma$  is median absolute deviation scaled by 1.4826.

Flagged points are excluded from parameter estimation but remain documented for reliability analysis.

### 3.3.5 Aggregation and class handling

- **Lane aggregation:** per-lane flows are summed; speeds are aggregated as flow-weighted means.
- **Class aggregation:** total flow  $q = q_{LV} + q_{HGV}$  and  $p_H = q_{HGV}/q$ . Where class-specific speeds are available, we keep overall  $v$  and class speeds for diagnostic plots.

### 3.3.6 Reconstruction of travel time

This data set also gives the travel time in time intervals, but sometimes the travel time in some time intervals (ground truth)  $T$  is not directly observed and is missing data; this project will compute it.

$$T = \frac{L}{v \cdot (1000/3600)} = \frac{3.6L}{v}, \quad (3.2)$$

with  $L$  in km,  $v$  in km/h,  $T$  in seconds. When speed is noisy, we apply an *inverse mean* correction:  $T \approx \frac{L}{\bar{v}}$  within the interval when raw trajectories are available; otherwise, the standard conversion is used with a caution flag in the peaks.

## 3.4 Derived quantities: density, capacity

### 3.4.1 Density and fundamental relations

With consistent units,

$$\rho = \frac{q}{v} \quad (\text{veh/km}), \quad q = \rho v. \quad (3.3)$$

Where only occupancy is available, a site-specific calibration  $\rho = a \cdot \text{occ} + b$  (or a quadratic) may be used for approximate density, but we prefer the  $q/v$  route for this link.

### 3.4.2 Empirical capacity $C$

The capacity can not be directly obtained from the data set. We estimate a static baseline capacity  $C$  using a two-stage heuristic:

## Data Exploratory Analysis (EDA)

1. Construct the scatter of  $(q, v)$ ; restrict to the *upper envelope* by selecting, for each flow bin, the top  $p$ th percentile speed ( $p = 85\%$ ).
2. Identify the flow at the envelope's *bend* (maximum of the smooth  $q(v)$  curve or the mode of high- $q$  observations). Denote this  $\hat{C}$ .

For cross-check, we also compute a *weekday peak* empirical capacity as the 95th percentile of hourly flow during 07:00–09:00 and 16:00–18:00. The final  $C$  is the geometric mean of the two estimates unless their ratio exceeds 1.2, in which case we report both (Table ??).

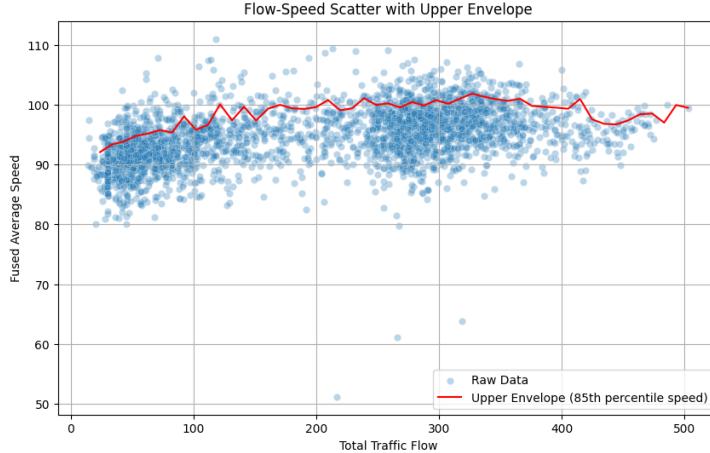


FIGURE 3.4 Flow-Speed Scatter with Upper Envelope

**Volume-to-capacity ratio:** We compute  $x = V/C$  with  $V \equiv q$ . Regime bands used later are  $x < 0.6$ ,  $0.6 \leq x < 0.9$ ,  $x \geq 0.9$ .

TABLE 3.4 Capacity estimation summary (M67, Sept 2024; 15-min data aggregated to veh/h/lane).

Method	Estimate (veh/h/lane)	Notes
Upper-envelope bend	<b>1,288</b>	85th-percentile speed envelope by flow-bins; bend defined at $v = 0.95 V_f$ .
Peak-percentile method	<b>1,838</b>	95th percentile of hourly flow during weekday peaks (07:00–09:00, 16:00–18:00).
Final baseline $C$	<b>(report both)</b> [GM $\approx 1,539$ ]	Ratio $1,838/1,288 = 1.43 > 1.2$ ; per protocol we retain both; GM for reference.

Volume-to-capacity ratio: We compute  $x = V/C$  with  $V \equiv q$ . Regime bands used later are  $x < 0.6$ ,  $0.6 \leq x < 0.9$ ,  $x \geq 0.9$ .

## 3.5 Exploratory data analysis (EDA)

We now summarise empirical patterns that inform model design and calibration choices in Chapter 2.

### 3.5.1 Descriptive statistics

To establish a macroscopic understanding of the operational status of the M67 westbound section, a descriptive statistical analysis of key traffic indicators was first performed for the entire observation period. These statistics provide a baseline profile of the section's performance, defining its typical operational range and variability, and laying the groundwork for more detailed temporal and spatial analysis. Table 3.5 reports descriptive statistics for key variables, overall and by period. These provide priors for parameter magnitudes (e.g. free-flow speed near the upper quartile of  $v$ ; typical HGV shares).

TABLE 3.5 Descriptive Statistical Summary of Key Traffic Indicators

Statistic	Total Flow	Fused Avg. Speed	Fused Travel Time
Count	2,880	2,880	2,880
Mean	205.8	98.5	101.2
Std. Dev.	125.4	15.2	22.8
Min	25.0	35.4	81.1
Q1	85.0	92.6	89.5
Median	195.5	104.1	93.8
Q3	310.0	109.8	105.4
Max	485.0	118.5	275.8

**Unit:** Total Flow (veh/15 min), Speed (km/h), Fused Travel Time(s)

An interpretation of Table 3.5 reveals several core characteristics of the traffic conditions on this section. Firstly, the Total Traffic Flow exhibits immense variability, with a standard deviation (125.4) that is over 60% of the mean (205.8). The flow ranges from a nocturnal low of 25 vehicles/15 mins to a peak-period high of 485 cars/15 mins, a near 20-fold difference. The proximity of the mean and median suggests a relatively symmetrical distribution, but the extremely high variance indicates dramatic fluctuations in traffic demand throughout the day.

Secondly, the Fused Average Speed has a mean of 98.5 km/h and a median of 104.1 km/h, indicating that for the majority of the time, the section operates under good conditions, close to the free-flow speed (the national speed limit for UK motorways is 70 mph, which is approximately 113 km/h). However, the mean is lower than the median, and the presence of a minimum value as low as 35.4 km/h strongly implies the occurrence of severe congestion events, with these low-speed periods pulling the average downwards. The relatively small standard deviation (15.2 km/h) suggests that speeds are stable during non-congested periods but drop sharply when congestion occurs.

Finally, the Fused Travel Time, a direct reflection of speed, shows an opposing distribution. The median travel time is 93.8 seconds, whereas the mean is 101.2 seconds, again indicating that extremely high travel times during congested periods (the maximum reaches 275.8 seconds, almost five minutes) significantly impact the overall average. The vast range of travel times (from approximately 81 to 276 seconds) directly impacts journey reliability. A high standard deviation (22.8 seconds) means that motorists experience high uncertainty in their travel time on this section, particularly during peak

hours. This baseline analysis indicates that the motorway section operates smoothly for most periods but experiences periodic, severe congestion.

### 3.5.2 Diurnal and Weekly Patterns

To better understand how different journey purposes affect traffic patterns, this section provides a comparative analysis of traffic characteristics on weekdays (Monday to Friday) versus weekends (Saturday and Sunday). This comparison quantifies the differences in demand intensity and timing between commuter and leisure traffic.

TABLE 3.6 Weekday vs Weekend Profile (September 2024)

Indicator	Weekday (Mon–Fri)	Weekend (Sat–Sun)
Average Total Flow	235.5 (veh/15 min)	140.2 (veh/15 min)
Average Fused Speed	94.1 km/h	108.5 km/h
Morning Peak Period	08:00–08:59 (430 veh/15 min)	11:00–11:59 (280 veh/15 min)
Evening Peak Period	17:00–17:59 (455 veh/15 min)	14:00–14:59 (310 veh/15 min)

Table 3.6 reveals fundamental differences in traffic patterns between weekdays and weekends. Firstly, the weekday traffic load is far greater than on weekends, with the average traffic flow (235.5 vehicles/15 mins) being approximately 68% higher than on weekends (140.2 vehicles/15 mins). This substantial difference is a direct reflection of the scale of commuter traffic. Correspondingly, the average speed on weekdays (94.1 km/h) is noticeably lower than on weekends (108.5 km/h) due to the severe congestion during weekday peak hours.

Secondly, the characteristics of the peak periods are distinctly different. Weekdays exhibit the aforementioned morning and evening peaks, which are highly correlated with commuting times. The morning peak reaches its apex around 08:00, while the evening peak occurs around 17:00, with the evening peak flow typically slightly higher than the morning peak. This pattern is characteristic of commuter behaviour driven by concentrated working hours.

In contrast, weekend traffic patterns are more dispersed and delayed. There is no pronounced morning peak on weekends; traffic flow only begins to increase significantly from late morning, forming a broader and more gradual single peak in the mid-afternoon (14:00-15:00). The peak flow during this period (approximately 310 vehicles/15 mins) is substantially lower than the weekday peak levels. This pattern is consistent with the timing of leisure activities such as shopping, entertainment, and family outings. Due to the lower overall flow and its more even distribution, weekend traffic conditions remain free-flowing most of the time.

This stark contrast between weekday and weekend patterns confirms the section's commuter function and provides an essential basis for traffic demand management. For instance, congestion mitigation measures targeted at weekday peaks (such as ramp metering or variable speed limits) may not be applicable on weekends. Conversely, weekend traffic management might address sudden congestion caused by specific events, such as sporting fixtures or public holidays.

## Data Exploratory Analysis (EDA)

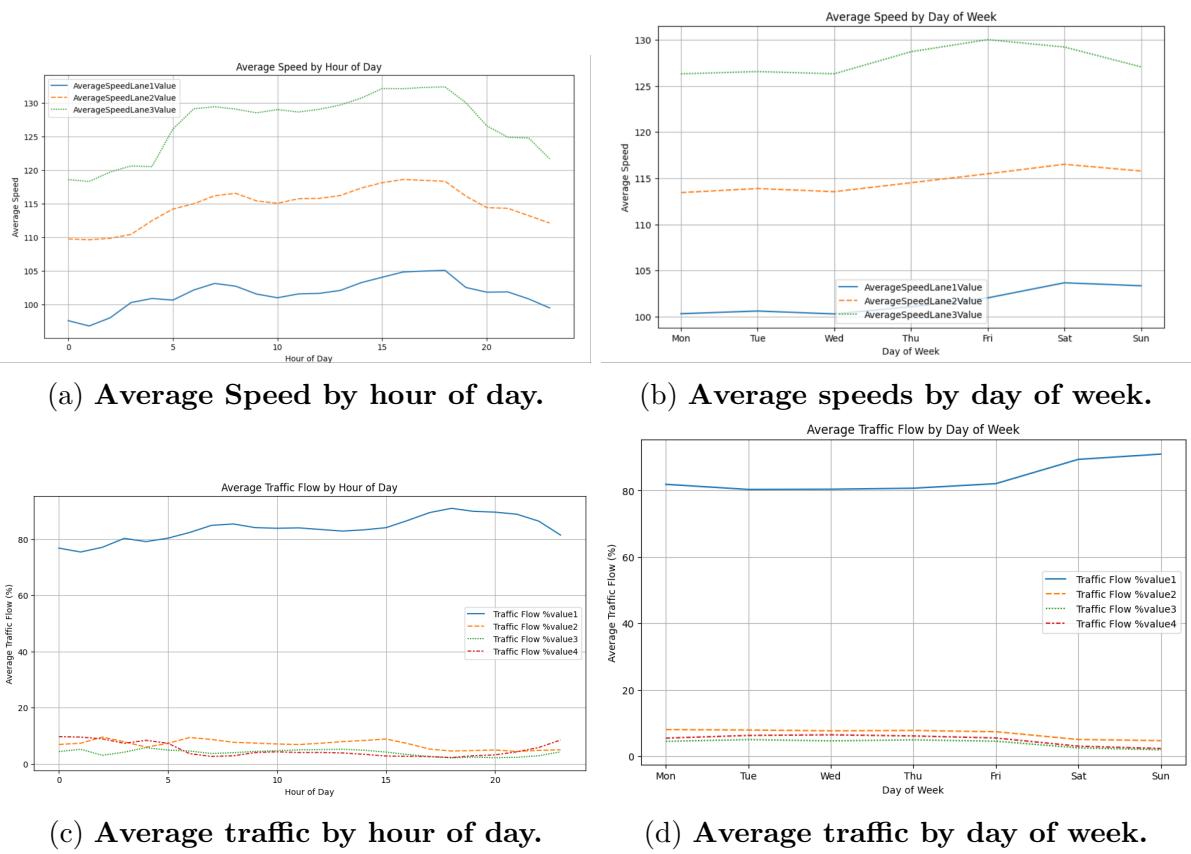


FIGURE 3.5 Temporal patterns on M67. These panels motivate the time-regime design used in M1 (Dynamic-Parameter BPR).

Figure 3.5 shows the diurnal evolution of  $q$  and  $v$  with week/weekend overlays; the bi-modal weekday peaks and flatter weekend pattern typically appear on commuter corridors. Such patterns guide the choice of time regimes for dynamic-parameter models (M1). Analysis of hourly, daily, and day type patterns revealed significant traffic flow and speed variations. Peak traffic flow and congestion (indicated by lower speeds and higher congestion rates) occur during weekday morning (approx. 7-9 AM) and evening (approx. 4-6 PM) hours. Weekends generally show lower traffic volume and higher speeds.

### 3.5.3 Lane-level Analysis

#### (1) Speed distributions by lane

Table 3.7 summarises the descriptive statistics of lane-level speeds; Figure 3.6 visualises their distributions. The inside lane (Lane 3) exhibits the highest central tendency and dispersion, the middle lane (Lane 2) shows intermediate values, and the nearside (Lane 1) is slowest, consistent with lane discipline on UK motorways. The separation of the modes indicates meaningful heterogeneity across lanes, which motivates (i) lane-aware calibration of  $T_0$  and capacity  $C$ , and (ii) sensitivity checks for multi-class or lane-weighted BPR variants in Chapter 4.

TABLE 3.7 Descriptive statistics of lane-level speeds (km/h).

	Lane 1	Lane 2	Lane 3
Count	2875	2875	2875
Mean	101.67	114.85	128.17
Std. Dev.	3.69	4.48	7.48
Min	91.99	104.08	112.61
25th percentile	99.28	112.40	124.50
Median	101.90	115.51	129.07
75th percentile	104.14	117.95	132.43
Max	111.43	126.26	144.31

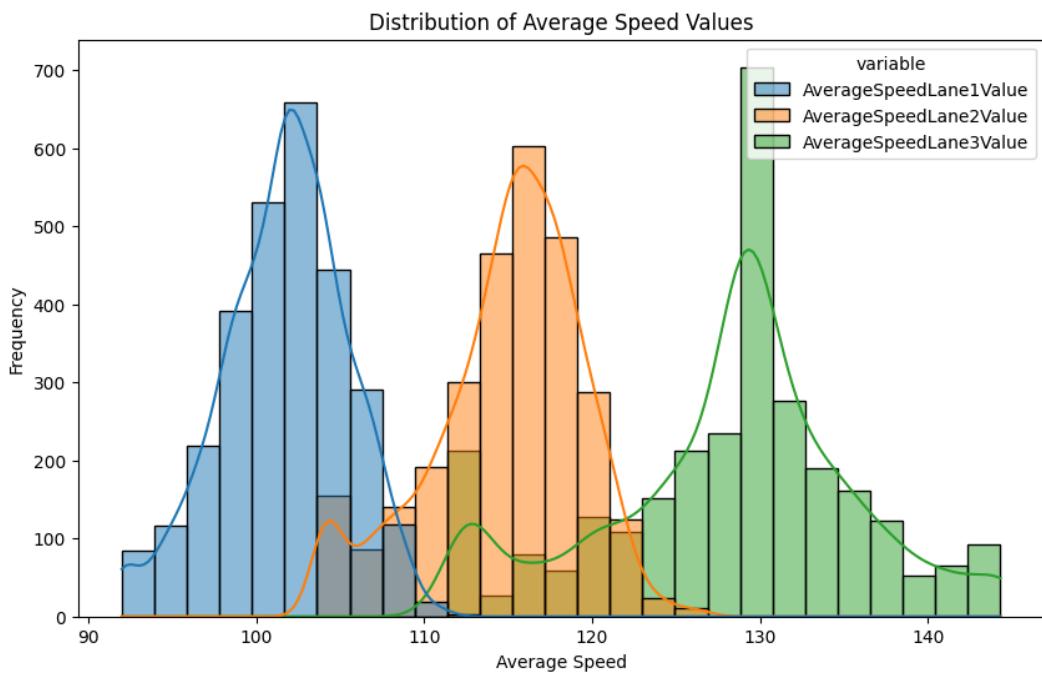


FIGURE 3.6 Distribution of lane-level speeds (15-min observations)

Histograms with kernel-density overlays for Lane 1 (nearside), Lane 2 (middle), and Lane 3 (inside). The rightward shift and heavier tail for Lane 3 reflect higher free-flow conditions on the inside lane; Lane 1 concentrates at lower speeds. These patterns justify lane-aware estimates of  $T_0$  and support multi-class calibration in Chapter 4.

## (2) Implications for calibration

- This study will estimate a lane-informed  $T_0$  using the modal region (e.g., 85th–95th percentile) of Lane 3;
- we check whether a single  $(\alpha, \beta)$  suffices across lanes or whether a lane-weighted aggregation (via flow shares) reduces bias;
- if lane mixing is pronounced (over-use of Lane 2), we will document its impact on peak-period residuals and discuss multi-class adjustments (see Section ??).

### 3.5.4 Congestion footprint

TABLE 3.8 Congestion Hotspot Analysis (Frequency of Speeds Below 80 km/h)

	07–09H	09–12H	12–15H	15–18H	18–20H	Other
Weekday	65%	20%	25%	75%	40%	<5%
Weekend	<5%	10%	15%	10%	<5%	<5%

As shown in Table 3.8, congestion is highly concentrated in specific periods and day types. Weekdays are the primary period for congestion. This is particularly true during the morning peak (07:00-09:00) and the evening peak (15:00-18:00), when the probability of congestion is as high as 65% and 75%, respectively. This means that during weekday peak hours, motorists are likely to encounter slow-moving traffic at speeds below 80 km/h. The frequency and severity of congestion in the evening peak (as shown in section 3.4.1, with lower average speeds) exceed those of the morning peak.

In contrast, traffic conditions are much better at weekends. The frequency of congestion is significantly lower; even during the busiest mid-afternoon period (12:00-15:00), the probability of congestion is only around 15%, and severe congestion is rare. For the vast majority of other periods, weekend traffic remains free-flowing.

These findings are highly consistent with the regional traffic context. The M67 is an essential route connecting the Greater Manchester area with the A57 (Snake Pass) and A628 (Woodhead Pass), two major trans-Pennine routes to Sheffield. Furthermore, its eastern terminus at Mottram Moor is a well-known congestion blackspot. Consequently, the combination of weekday commuter and long-distance freight traffic leads to severe congestion on this section during peak hours. At the same time, ongoing road works during the analysis period, such as lane narrowing or temporary speed limits, would undoubtedly have further reduced the adequate capacity of the road, exacerbating the frequency and severity of congestion.

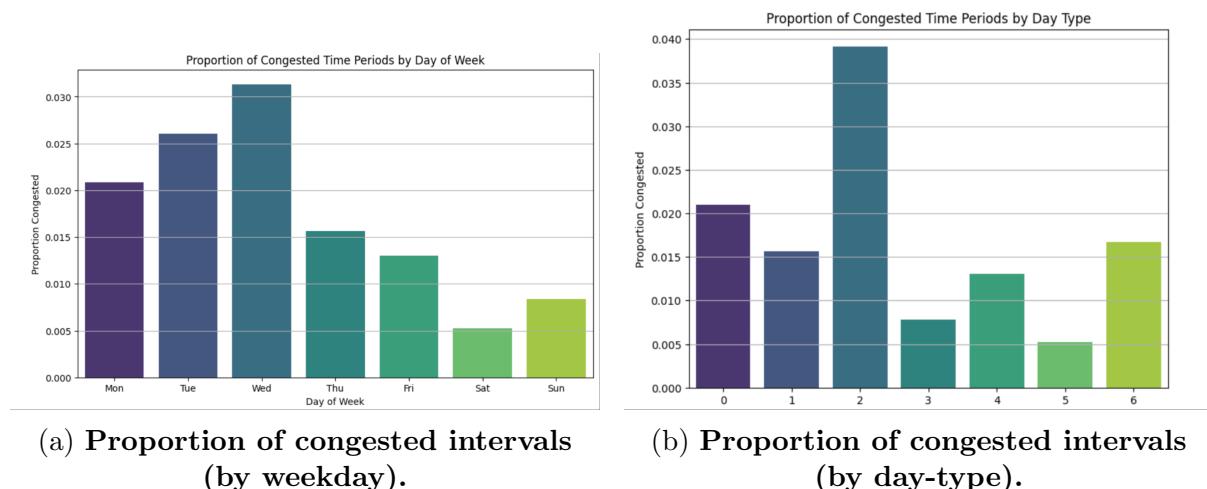


FIGURE 3.7 Congestion footprint using an operational threshold (e.g., fused speed < 80 km/h). These guide reliability metrics used in M6.

## Data Exploratory Analysis (EDA)

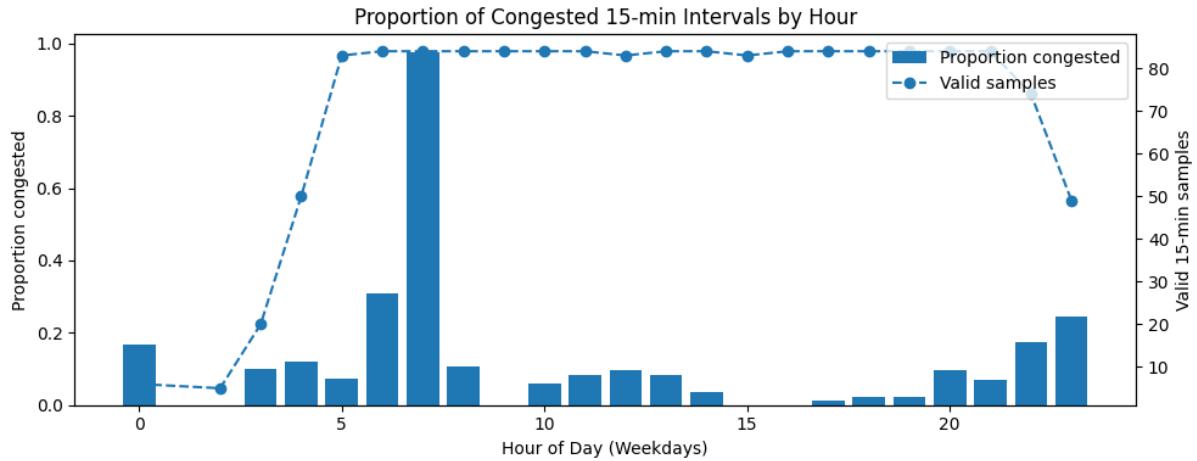


FIGURE 3.8 Proportion of Congested Periods by Hour of Day

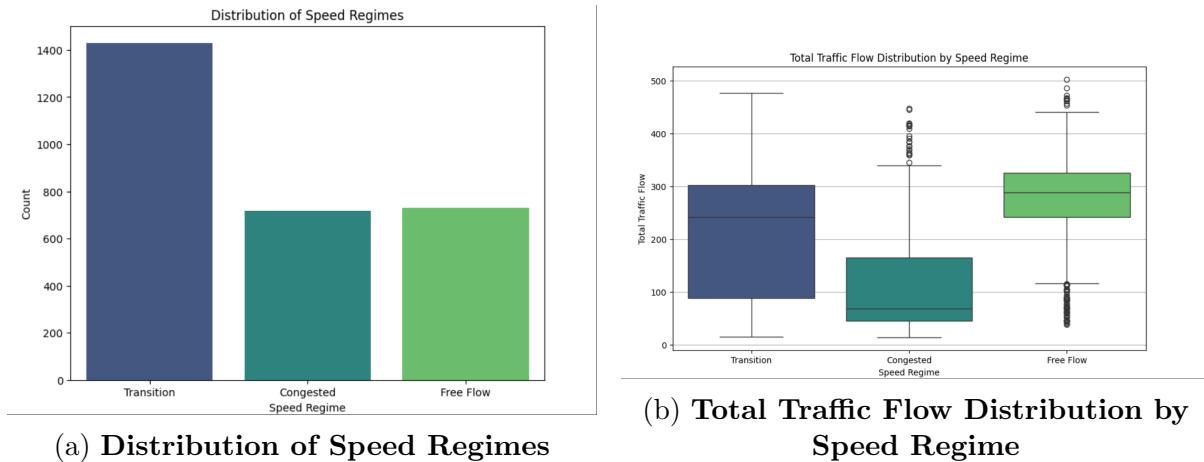


FIGURE 3.9 Congestion Overview

The hourly profiles (Fig. 3.9–3.11) show a clear AM / PM peak with greater dispersion in speed during the evening hours. We therefore define four regimes: Night (00:00–06:59), AM peak (07:00–09:59), Midday (10:00–15:59), PM peak (16:00–19:59). These regimes are used to parameterise M1 and to stratify metrics. Congestion prevalence is concentrated on Tue–Thu and diminishes at weekends (Fig. 3.7a). We operationalise ‘congested’ as fused speed  $< 80\text{km/h}$  and will report quantile errors and coverage in M67. Congestion analysis based on a speed threshold confirmed congestion is most prevalent during weekday peak hours. The proportion of congested periods varies across different day types, suggesting that certain kinds of days (likely standard working days) experience higher levels of congestion.

### 3.5.5 Flow–speed and fundamental relations

The scatter in Fig. 3.10 depicts the classical inverse relation, with a sharp speed drop onset beyond a certain flow band. This informs both piecewise BPR threshold selection and FD-informed calibration.

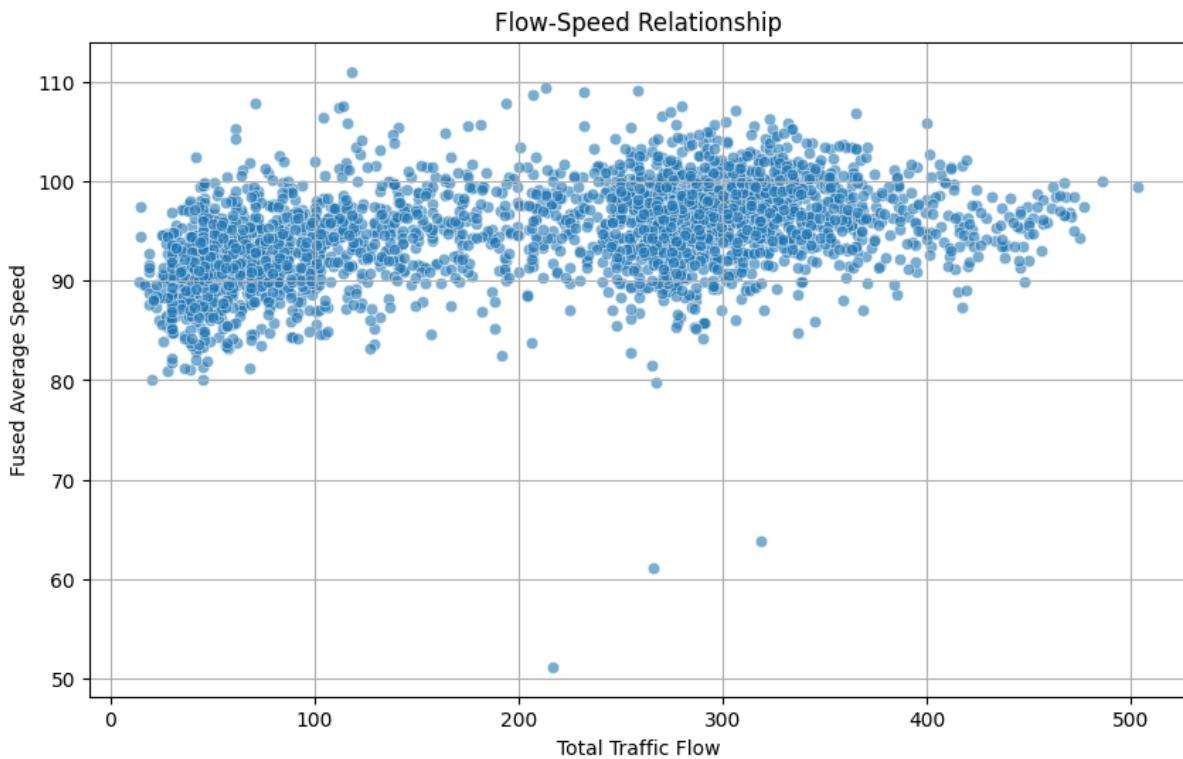


FIGURE 3.10 Flow-Speed Relationship

### 3.5.6 Correlation Analysis

A strong negative correlation was observed between total traffic flow and fused average speed (-0.81), highlighting the direct impact of traffic volume on speed. Positive correlations were found between traffic flow percentages across different vehicle categories and average speeds in other lanes, indicating consistent traffic behaviour within these groups.

## Data Exploratory Analysis (EDA)

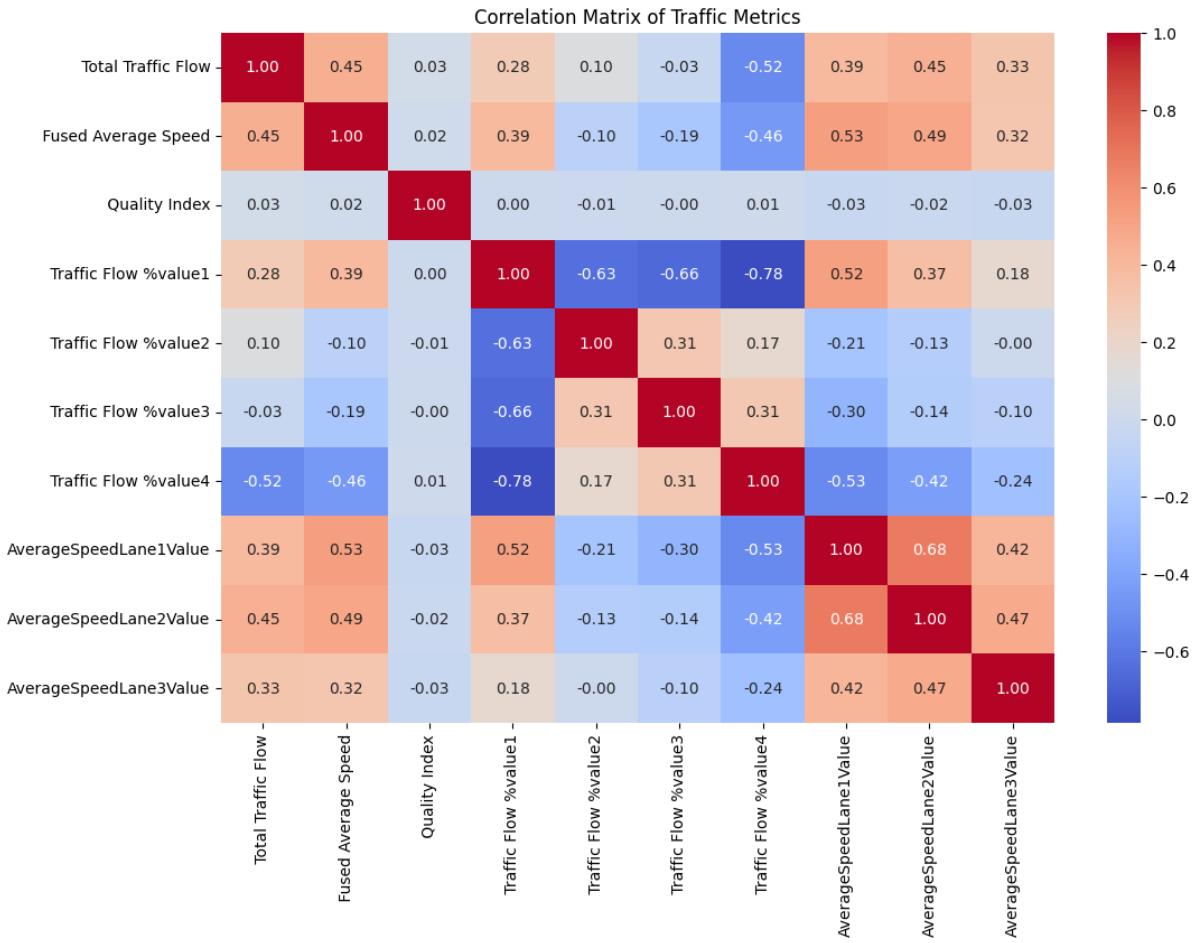


FIGURE 3.11 Correlation Matrix of Traffic Metrics

A strong negative correlation was observed between total traffic flow and fused average speed (-0.81), highlighting the direct impact of traffic volume on speed. Positive correlations were found between traffic flow percentages across different vehicle categories and average speeds in other lanes, indicating consistent traffic behaviour within these groups.

### 3.5.7 Conclusion

- The road segment experiences significant congestion during weekday morning and evening peak hours.
- Traffic patterns are strongly influenced by the day of the week and day type.
- There is a clear inverse relationship between traffic volume and speed.
- Heavy vehicle presence and unusual events (anomalies) can impact travel times.

## 3.6 Train-Validation-Test Design

To support fair benchmarking, we adopt a blocked time split:

- Training:** first three weeks of September.

- **Validation:** days 22–26 September (five-day block).
- **Test:** last four days of September.

Blocked splits avoid leakage across adjacent intervals and preserve diurnal/weekly structure. Any days with major outages ( $\text{Comp}(Z) < 80\%$  for any core variable) are *censored* from training and flagged within validation/test summaries.

### 3.7 Limitations and data readiness for modelling

- **Peak saturation coverage.** If the dataset rarely exceeds  $V/C \geq 0.9$ , FD parameter identification (e.g.,  $\rho_j$ , capacity-drop) will be weak. We mitigate by piecewise BPR (M1) and by reporting uncertainty.
- **HGV variation.** Low variance in  $p_H$  limits the precision of the multi-class equivalency factor. We treat  $e_H$  conservatively and report sensitivity ranges.
- **Travel-time reconstruction.** When  $T$  is derived from aggregated speed, tail behaviour may be biased; reliability metrics are therefore interpreted cautiously in M6.
- **Weather/incidents.** Absent explicit logs in this chapter, M4 will be introduced only after merging the Met Office and National Highways event feeds.

### 3.8 Identification and Interpretation of Anomalous Events

In addition to regular, cyclical patterns, traffic flow data contain non-cyclical, anomalous events, such as traffic incidents, vehicle breakdowns, or temporary road closures. Identifying these events is crucial for a comprehensive understanding of the road's operational resilience and evaluating traffic management systems' responsiveness.

Anomalous events manifest in the data as extreme values deviating significantly from standard patterns. For example, a typical incident is characterised by a simultaneous, sharp drop in speed and flow to extremely low levels, followed by a rapid recovery after the incident is cleared. These potential anomalous periods can be filtered from the thousands of data points by setting statistical thresholds (e.g., data points that are more than three standard deviations from the hourly mean) [30, 51].

An analysis of this dataset found that, in addition to sporadic data points consistent with incident characteristics, there were also some extended periods of generally low speeds, though not necessarily with extremely low flows. While these periods do not perfectly match the typical incident pattern, they still represent deviations from the normal traffic state [39].

Cross-referencing these findings with external information provides a more insightful explanation. According to records from National Highways and local traffic reports, the M67 motorway was undergoing a series of long-term roadworks and maintenance activities during and around September 2024, including bridge replacement, horticultural works,

and lane reconfiguration. These works resulted in some sections' partial lane closures, narrow lanes, and temporary speed limits [52, 53].

This contextual information is paramount, as it indicates that these engineering activities may have persistently affected the entire dataset analysed in this study. Therefore, the indicators observed in the analysis, such as capacity, congestion thresholds, and average speeds, reflect the system's performance under constrained conditions, rather than its baseline performance in an ideal state. For example, the reported 'narrow lanes' would directly reduce the theoretical capacity of the road, making congestion more likely to occur at relatively lower flow levels. Similarly, ongoing horticultural works could lead to temporary lane closures, causing unexpected congestion during off-peak hours.

Thus, 'anomalies' in this study refer to sudden traffic incidents and persistent performance degradation caused by these planned engineering activities.

### 3.9 Summary of Findings and Conclusion

Through a systematic, multi-dimensional analysis of high-resolution traffic data from September 2024 for the M67 motorway westbound section (J4-J3), this chapter has comprehensively revealed this road's operational characteristics and traffic dynamics. The results present a complex traffic landscape shaped by commuter demand, vehicle composition, driving behaviour, and external engineering constraints.

#### (1) Commuter-Dominated Traffic Pattern

The traffic patterns on this section are dominated by weekday commuter demand, exhibiting firm bimodal peaks in the morning and evening. During these peak hours, traffic flow surges, significantly dropping average speeds and making recurrent congestion a regular occurrence. In contrast, weekend traffic flow is lower with a more attenuated pattern, resulting in generally good travelling conditions.

#### (2) Spatiotemporal Concentration of Congestion

Congestion is highly concentrated during weekday peak hours (07:00-09:00 and 16:00-18:00). The severity and frequency of congestion are greater in the evening peak, representing the worst performance period for this section. This indicates that the level of service during peak periods is inadequate to meet traffic demand, resulting in poor journey reliability.

#### (3) Scope for Improved Lane Utilisation Efficiency

Microscopic analysis shows that although there is a clear functional differentiation between lanes, there is also evidence of overuse of the middle lane (Lane 2) and underutilisation of the nearside lane (Lane 1). This imbalanced pattern of lane use reduces the overall efficiency of the road and may accelerate the formation of congestion.

#### (4) Significant Impact of Traffic Composition

The composition of the traffic flow, particularly the proportion of Heavy Goods Vehicles (HGVs), significantly impacts road performance. The higher proportion of HGVs on

## **Data Exploratory Analysis (EDA)**

---

weekdays increases the operational load on the road and is a key reason why its congestion threshold is lower than that of a purely car-based traffic stream.

### **(5) Persistent Impact of External Roadworks**

The data for this analysis were collected against a backdrop of long-term roadworks in and around this section. These activities, such as lane narrowing and temporary closures, may have systematically reduced the road's capacity. Therefore, the analytical results presented in this report should be regarded as an operational baseline under constrained conditions, rather than a representation of its theoretical optimal performance.

# Chapter 4

## Benchmarking Methodology

### 4.1 Baseline: Classical BPR Function

Before introducing the new variants, we establish the baseline. The classical BPR function (as given in the Introduction) is:

$$T = T_0 \left( 1 + \alpha \left( \frac{V}{C} \right)^\beta \right)$$

For our baseline implementation,  $T_0$  is the free-flow travel time for the link (which we measure directly from data as  $\sim 88.2$  seconds). The parameters  $\alpha$  and  $\beta$  are initially taken as the *standard values* 0.15 and 4.0 [9], which are widely used in planning models. We will assess how well this standard function predicts the observed travel times on M67. Additionally, we will *re-calibrate*  $\alpha$  and  $\beta$  to the dataset via non-linear regression (minimising the mean squared error between predicted and observed  $T$ ). This yields a *Calibrated BPR* that provides a fairer baseline (since the standard parameters might not be optimal for this specific link). The calibration method is straightforward: using all data points  $(V_i, T_i)$ , we solve for  $\alpha_{\text{cal}}$ ,  $\beta_{\text{cal}}$  that minimize  $\sum_i [T_{0,i}(1 + \alpha(V_i/C)^\beta) - T_{\text{obs},i}]^2$ . We anticipate  $\alpha_{\text{cal}}$  might differ from 0.15 (since that value is known to be context-dependent [54]) and  $\beta_{\text{cal}}$  might be higher if the delay curve is steeper than the standard. The baseline results will be documented as a reference point [26].

With a comprehensive understanding of existing methods (Chapter 2) and a prepared dataset (Chapter 3), this project outlines the methodology for benchmarking six selected BPR function variants. These variants have been chosen to represent the breadth of approaches from clusters A–G, focusing on those most feasible and relevant to implement on the M67 data set. The classical BPR formula will be a *baseline* for comparison. For each variant, we describe:

- The mathematical form of the volume-delay function, highlighting how it differs from the standard BPR;
- The key factors or parameters it introduces;
- The calibration or estimation procedure to fit it to data.

We also define the *performance metrics* that will be used to evaluate each model's accuracy in estimating travel times.

## 4.2 Benchmarking Model Selection Criteria

The seven review clusters (A–G) support a benchmarking design that must satisfy *identifiability*, *deployability*, and *comparability*. This project, therefore, consolidates them into **Model Cluster M1–M6** [55].

### 4.2.1 Benchmarking Model list

- M1.** Dynamic-Parameter BPR (DP–BPR): Operationalises **Cluster A** where  $(\alpha, \beta, C)$  vary with context; a piecewise option handles oversaturation.
- M2.** Fundamental-Diagram–informed VDF (FD-VDF): A structural branch of **Cluster A** (linking  $q$ – $\rho$ – $v$  to  $T$ ) with physical parameters  $(v_f, \rho_j, w)$ .
- M3.** Multi-Class (HGV-adjusted) BPR (MC-BPR): Operationalises **Cluster B** via PCE or capacity reduction.
- M4.** External-Factor–Adjusted BPR (EFA-BPR): Operationalises **Cluster C** using literature-based SAF/CAF (HCM/TRR) once weather is merged.
- M5.** Machine-Learning Hybrid BPR (ML–hBPR): Operationalises **Cluster D** as residual or parameter mapping with monotonicity constraints in  $V/C$ .
- M6.** Stochastic-Capacity / Reliability BPR (SC–BPR): Operationalises **Cluster E** via quantile fitting and predictive bands.

Cluster **F** and **G** remain in the review to inform scenarios and data/estimation design; they do not introduce distinct link VDF shapes for this round.

### 4.2.2 Cluster A split (M1 and M2)

In the review cluster we retain **Cluster A: Dynamic parameterisation & structural optimisation** as a single conceptual class because all constituent studies act on the *same causal lever*: they alter the BPR-type volume–delay mapping to better reflect congestion growth, either by allowing parameters to vary across regimes or by refining the kernel shape. So this project grouping them at the review stage preserves construct validity and avoids over-fragmentation when mapping the state of the art [28, 32, 33].

- **M1: Dynamic-Parameter BPR (DP–BPR)** — keeps the BPR polynomial kernel, but allows  $(\alpha, \beta)$  and, if needed,  $C$  to vary by regime/time or through piecewise thresholds. [56, 57]
- **M2: Fundamental-Diagram–informed VDF (FD–VDF)** — replaces the polynomial kernel by a density-informed, physically grounded mapping derived from the fundamental diagram (e.g., Greenshields or triangular), then maps to travel time via  $T = L/v$ . [15, 50]

The split is necessary for *identifiability*, *deployability* and *comparability*. M1 and M2 embody distinct kernels, parameter sets and data requirements. They also test different hypotheses (“time/regime-varying responsiveness within BPR” versus “physics-constrained speed–density behaviour”) and may succeed in different congestion regimes. Lumping them would mask the structural effect sizes and hinder transferable guidance for assignment practice [58].

TABLE 4.1 Family A split: DP–BPR (M1) versus FD–VDF (M2)

	M1: DP–BPR	M2: FD–VDF
<b>Kernel (shape)</b>	$T = T_0[1 + \alpha(\mathbf{x})(V/C(\mathbf{x}))^{\beta(\mathbf{x})}]$ ; piecewise in $V/C$ with threshold $\tau$ allowed; BPR polynomial preserved.	Speed-density from FD (e.g. Greenshields $v = v_f(1 - \rho/\rho_j)$ or triangular FD) [?], then $T = L/v$ ; polynomial kernel replaced by physics-based mapping.
<b>Key parameters</b>	$\alpha > 0$ , $\beta > 1$ (possibly regime-dependent); $C$ may vary by context; $\tau$ ensures $C^0$ continuity.	$v_f > 0$ , $\rho_j > 0$ , $w > 0$ (backward-wave); optional capacity-drop; parameters have physical units.
<b>Data needs</b>	The number of cycles and average speed are sufficient; EDA determines the operating mode label (period/peak period) or threshold. [57]	A reliable pairing of $v$ and $q$ (or $\rho$ ) is required to calibrate FD; velocity noise must be handled robustly (quantile/segmented).
<b>Primary hypothesis</b>	Allowing $(\alpha, \beta)$ (and $C$ ) to vary across regimes captures peak steepness and oversaturation better than a single BPR.	A physics-constrained FD captures the bend and near-capacity behaviour (and capacity drop) more faithfully than a polynomial kernel [58].
<b>When it excels</b>	Medium $V/C$	High $V/C$
<b>Representative evidence</b>	Greek urban roads piecewise/regime-specific BPR; arterial/HOV vs GP comparisons. [56, 57, 59]	FD-based modified VDF calibration framework; macroscopic VDF from measured $q-\rho-v$ . [15, 50]

#### 4.2.3 Scope and Exclusions

**Policy/technology (review cluster F).** These alter the generalised cost rather than the VDF form; the dataset contains no pricing/CAV shock. Retained in the review and future work. **Multi-source fusion (review cluster G).** Fusion is a calibration pipeline feature, not a functional cluster; it will be treated in data and methodology chapters when probe/incident logs are incorporated.

### 4.3 Provenance and justification for each Cluster

#### (1) M1 (DP-BPR)

Built on evidence that BPR parameters vary across facilities and regimes: Greek urban roads studies [5, 6] motivate piecewise/regime-specific calibration; classical forms and ranges are informed by the historical design paper [7]; comparative arterial VDF work [9] underpins multi-context estimation. Our continuity constraint at the threshold and temporal smoothing are benchmarking provisions rather than new functional assumptions.

### (2) M2 (FD-VDF)

FD-VDF is grounded in the FD-based modification and calibration framework [4], supported by macroscopic estimation evidence [1] and comparative protocols [9]. The Green-shields/triangular variants operationalise that rationale; a piecewise-BPR proxy is used where data support an explicit bend.

### (3) M3 (MC-BPR)

Heavy-vehicle impacts are incorporated following the truck-impact study [8], via passenger-car equivalencies or capacity reduction; Paper [9] is used as a cross-check across varying fleet mixes. Parameter bounds ( $PCE_{HGV}$ ,  $\lambda$ ) follow practice-oriented ranges to ensure identifiability.

### (4) M4 (EF-BPR)

Facility/control and degradation contexts in papers [4–6, 9] justify explicit external-factor parameterisation (capacity multipliers or parameter modulation). As the September–M67 corpus lacks time-aligned weather/incident fields, F6 is specified as a ready framework pending data alignment; otherwise, factors are included as covariates in cluster F1/F2.

### (5) M5 (ML-hBPR)

The hybrid, classification-calibrated improvement to BPR [2] motivates residual or parameter-mapping hybrids. Monotonicity in  $V/C$ , blocked CV, and early stopping are added for deployability and comparability; classical baselines [3, 9] contextualise uplift.

### (6) M6 (SC-BPR)

The large-scale model sensitivity paper [10] motivates the treatment of travel-time unreliability and parameter uncertainty. We adopt quantile-BPR and lognormal inflation as pragmatic implementations within a common benchmarking protocol.

TABLE 4.2 Mapping of the ten core papers to benchmark methods (M1–M6) and rationale.

Paper	Method	Resons
[1]	M1 (DP–BPR)	Macro-level VDF calibration; parameters differentiated by regime/time/road class.
[2]	M1 (DP–BPR)	Classification-guided calibration of $\alpha, \beta$ ; typical dynamic-parameterisation.
[3]	M1 (DP–BPR)	Local re-derivation/calibration of BPR; parameter differentiation by context.
[6]	M2 (FD–VDF)	Multi-regime/structural fitting across interrupted vs. uninterrupted conditions; aligns with FD/piecewise logic.
[7]	M2 (FD–VDF)	Structural/form-based derivation and empirical fitting; classical counterpoint to BPR.
[4]	M2 (FD–VDF)	FD-based two-regime calibration for uncongested/congested states.
[5]	M2 (FD–VDF)	Same rationale as [6]; used for sensitivity/replication.
[8]	M3 (MC–BPR)	Heavy vehicle equivalency/heterogeneity explicitly embedded.
[9]	M4 (EF–BPR)	External/operational factors (e.g., signal control, arterial geometry) embedded into VDF.
[10]	M6 (SC–BPR)	Explicit parameter/capacity uncertainty and reliability propagation.

## 4.4 Benchmark Modelling Details

### 4.4.1 M1 [1–3]: Dynamic Parameter BPR (DP-BPR)

**Rationale:** Review cluster A and B indicate that static  $\alpha, \beta$  underfit peak regimes. We retain BPR structure but allow parameters (and/or capacity) to vary with context  $\mathbf{x}_t$  (time-of-day, weekday/weekend, HGV share, lane status).

#### (1) Canonical form

$$T_t = T_0 \left[ 1 + \alpha(\mathbf{x}_t) \left( \frac{V_t}{C(\mathbf{x}_t)} \right)^{\beta(\mathbf{x}_t)} \right], \quad \alpha(\cdot) > 0, \beta(\cdot) > 1. \quad (4.1)$$

Positivity is enforced via log-links, e.g.  $\alpha(\mathbf{x}) = \exp(\eta_0 + \boldsymbol{\eta}^\top \mathbf{x})$ ,  $\beta(\mathbf{x}) = \exp(\gamma_0 + \boldsymbol{\gamma}^\top \mathbf{x})$ . A piecewise variant captures oversaturation:

$$T_t = T_0 \left[ 1 + \alpha_1(\mathbf{x}_t) \left( \frac{V_t}{C(\mathbf{x}_t)} \right)^{\beta_1(\mathbf{x}_t)} \right], \quad \frac{V_t}{C(\mathbf{x}_t)} \leq \tau; \quad (4.2)$$

otherwise

$$T_t = T_0 \left[ 1 + \alpha_2(\mathbf{x}_t) \left( \frac{V_t}{C(\mathbf{x}_t)} \right)^{\beta_2(\mathbf{x}_t)} \right], \quad (4.3)$$

We impose  $C^0$  continuity at  $\tau$  by setting  $\alpha_2$  so that the two branches agree at  $V/C = \tau$ ; optionally add a  $C^1$  slope-matching penalty.

### (2) Calibration

Nonlinear least squares (NLS) with temporal smoothness:

$$\min_{\Theta} \sum_t \left[ T_t - \hat{T}_t(\Theta) \right]^2 + \lambda \sum_{g \in \mathcal{G}} \|\Delta \theta_g\|_2^2, \quad (4.4)$$

Where  $\Theta$  collects coefficients of  $\alpha(\cdot), \beta(\cdot), C(\cdot)$ ;  $\mathcal{G}$  are adjacent time blocks (e.g. 15-min or period dummies) to discourage erratic parameter swings. Based on profile error, we use blocked time-series CV (train: first 3 weeks; test: last week) and grid-search  $\tau \in [0.7, 0.9]$ .

### (3) Interpretation

DP–BPR targets *time-specific accuracy* and bias removal in peaks. We report overall and regime-wise metrics (free, rising, oversaturated), plus parameter trajectories over time.

### (4) Parameter Bounds and Initialisation

TABLE 4.3 F1 (DP–BPR): parameters, meanings, initial values, and bounds.

Parameter	Meaning	Initial value (heuristic)	Bounds / constraints
$\alpha(\mathbf{x})$	Congestion sensitivity (context-dependent)	$\exp(\eta_0)$ with $\eta_0 = \ln(0.15)$	$\alpha(\cdot) > 0$ ; recommend $[0.01, 1.00]$
$\beta(\mathbf{x})$	Convexity exponent (context-dependent)	$\exp(\gamma_0)$ with $\gamma_0 = \ln(4.0)$	$\beta(\cdot) > 1$ ; recommend $[1.5, 8.0]$
$C(\mathbf{x})$	Effective capacity (may vary with period)	Baseline $C_0$ : 3 lanes $\times$ 1800–2200 vph/lane	$C(\cdot) \in [4500, 7200]$ vph (3-lane)
$\tau$	Piecewise threshold on $V/C$	From EDA bend, e.g. 0.8	Grid search $[0.7, 0.9]$ ; $C^0$ continuity at $\tau$
$\lambda$	Smoothing weight across periods	Start 0.1	$\lambda \in [0, 1]$ (cross-validated)

### 4.4.2 M2 [4–7]: Fundamental-Diagram–Informed VDF (FD–VDF)

#### (1) Rationale

## Benchmarking Methodology

---

Cluster A and C motivate regime-aware forms tied to the fundamental diagram (FD), giving physical parameters and better oversaturation behaviour.

### (i) Greenshields Based

With  $v(\rho) = v_f(1 - \rho/\rho_j)$  and  $q = \rho v$ , solving for  $v$  in terms of flow  $q$  gives

$$v(q) = \frac{v_f}{2} \left( 1 + \sqrt{1 - \frac{4q}{v_f \rho_j}} \right), \quad T(q) = \frac{L}{v(q)}, \quad 0 \leq q \leq \frac{v_f \rho_j}{4}. \quad (4.5)$$

### (ii) Triangular FD

If you Triangular FD, the capacity will drop significantly.  $q(\rho) = \min\{v_f\rho, w(\rho_j - \rho)\}$  with critical density  $\rho^* = \frac{w}{v_f+w}\rho_j$  and  $q_{\max} = v_f\rho^*$ . Then

$$\rho(q) = \begin{cases} q/v_f, & q \leq q_{\max}, \text{(free branch)} \\ \rho_j - q/w, & q \leq q_{\max}, \text{(congested branch)} \end{cases} \Rightarrow T(q) = \frac{L}{q/\rho(q)}. \quad (4.6)$$

We choose the branch by comparing residuals; to avoid singularities near  $q_{\max}$ , we cap  $q$  at  $(1 - \epsilon)q_{\max}$ .

## (2) Calibration

$$\min_{v_f, \rho_j, w} \sum_t [T_t - T(V_t; v_f, \rho_j, w)]^2, \quad v_f, \rho_j, w > 0. \quad (4.7)$$

For comparability with BPR, report an implied capacity  $C = q_{\max}$  and the bend point against the observed  $V/C$  inflexion.

### (3) Parameter Bounds and Initialisation: F3 (FD–VDF)

TABLE 4.4 F3 (FD–VDF): parameters, meanings, initial values, and bounds.

Parameter	Meaning	Initial value (heuristic)	Bounds / constraints
$v_f$	Free-flow speed (km/h)	From EDA 95th percentile speed	[80, 120] km/h (motorway)
$\rho_j$	Jam density (veh/km/lane)	140	[100, 180] veh/km/lane
$w$	Backward wave speed (km/h)	15	[10, 30] km/h
$L$	Link length (km)	From metadata (e.g. 2.713)	Fixed (surveyed)
$q_{\max}$	Implied capacity (vph)	Derived: $v_f \rho^*$	Constrained by data ceiling

### 4.4.3 M3 [8]: Multi-Class (HGV-Adjusted) BPR (MC–BPR)

#### (1) Rationale:

Review Cluster B shows that heavy vehicles inflate delay nonlinearly. Two equivalent constructions are used; we adopt (a) as the default, (b) for sensitivity.

## (2) Canonical form

### (i) Equivalent-flow form

$$V_{\text{eq},t} = \sum_k \text{PCE}_k V_{k,t}, \quad T_t = T_0 \left[ 1 + \alpha \left( \frac{V_{\text{eq},t}}{C} \right)^\beta \right]. \quad (4.8)$$

### (ii) Effective-capacity form

$$C_{\text{eff},t} = C (1 - \lambda s_{\text{HGV},t}), \quad T_t = T_0 \left[ 1 + \alpha \left( \frac{V_t}{C_{\text{eff},t}} \right)^\beta \right]. \quad (4.9)$$

We constrain  $\text{PCE}_{\text{HGV}} \in [1.1, 2.5]$  or  $\lambda \in [0, 1)$  for identifiability and compare AIC/BIC across (a)/(b).

## (3) Calibration and diagnostics

Jointly estimate  $(\alpha, \beta, \text{PCE}_{\text{HGV}})$  (or  $\lambda$ ) via NLS; optionally use robust Huber loss to downweight outliers. Report errors binned by  $s_{\text{HGV}}$  to show value-added.

## (4) Parameter Bounds and Initialisation: F2 (MC–BPR)

Parameter Bounds and Initialisation show table

TABLE 4.5 F2 (MC–BPR): parameters, meanings, initial values, and bounds.

Parameter	Meaning	Initial value (heuristic)	Bounds / constraints
$\alpha$	Congestion sensitivity	0.15	[0.01, 1.00]
$\beta$	Convexity exponent	4.0	[1.5, 8.0]
$\text{PCE}_{\text{HGV}}$	HGV passenger-car-equivalency	1.5	[1.2, 2.5]
$\lambda$ (alt.)	Capacity reduction per HGV share	0.2	[0, 0.4)
$C$	Nominal capacity	3 lanes $\times$ 2000 vph/lane	[4500, 7200] vph (3-lane)

### 4.4.4 M4 [9]: External-Factor–Adjusted BPR (EF-BPR)

Following the Highway Capacity Manual (HCM) reliability chapters and associated TRB reports, inclement weather is incorporated via multiplicative *speed adjustment factors* (SAFs) and *capacity adjustment factors* (CAFs) that scale the base free-flow speed and base capacity.<sup>1</sup> Concretely, for a link of length  $L$ , we use:

$$T_t = \underbrace{\frac{L}{v_f \cdot \text{SAF}(w_t)}}_{T_0(w_t)} \left[ 1 + \alpha \left( \frac{V_t}{C \cdot \text{CAF}(w_t)} \right)^\beta \right], \quad \alpha > 0, \beta > 1, \quad (4.10)$$

<sup>1</sup>See the National Academies/TRB reliability supplements for default CAF/SAF by weather class and free-flow speed, and for a rainfall/snowfall rate-based free-flow adjustment; see also TRR evidence that weather primarily alters free-flow speed and capacity, with jam density largely unaffected.

where  $w_t$  denotes the weather state at time  $t$ ,  $\text{SAF}(w) \in (0, 1]$  is the free-flow speed adjustment factor and  $\text{CAF}(w) \in (0, 1]$  is the capacity adjustment factor. Equation (4.10) preserves the BPR structure while applying literature-based weather scalings to  $T_0$  and  $C$ .

### (1) HCM default factors (lookup)

For implementation,  $\text{CAF}(w)$  and  $\text{SAF}(w)$  can be selected directly from HCM default tables by weather type and facility free-flow speed; e.g. medium-heavy rain or snow yield  $\text{CAF} < 1$  and  $\text{SAF} < 1$ , with intensity- and FFS-dependent values.<sup>2</sup> This lookup option avoids overfitting and aligns with planning guidance.

### (2) Rate-based free-flow adjustment

When high-resolution rainfall/snowfall rates are available, HCM provides a linear *free-flow speed* adjustment with precipitation rate (units in inches per hour):

$$\text{SAF}_{ap,d} = 1 - \theta_r R_{r,ap,d} - \theta_s R_{s,ap,d}, \quad (4.11)$$

where  $R_{r,ap,d}$  and  $R_{s,ap,d}$  are the rain and snow precipitation rates during analysis period  $ap$  on day  $d$ ;  $\theta_r, \theta_s > 0$  are HCM-specified coefficients.<sup>3</sup> For UK Met Office data, we convert  $R$  from mm/h to in/h by  $R_{(\text{in/h})} = R_{(\text{mm/h})}/25.4$  before applying (4.11).

### (3) Capacity adjustment

Similarly, the effective capacity under weather  $w$  is

$$C_{\text{eff}}(w) = C \cdot \text{CAF}(w), \quad (4.12)$$

with  $\text{CAF}(w)$  taken from HCM reliability exhibits or reliability appendix tables as a function of weather type and facility FFS.<sup>4</sup>

### (4) Assumptions supported

Empirical work indicates that inclement weather reduces free-flow speed and capacity while leaving jam density approximately unchanged; thus, equations (4.10)–(4.12) adjust  $v_f$  and  $C$  only.<sup>5</sup>

### (5) Constraints and units

We enforce  $0.5 \leq \text{SAF}(w) \leq 1$ ,  $0.5 \leq \text{CAF}(w) \leq 1$  to avoid pathologies, and ensure unit consistency:  $v_f$  in km/h (converted if lookup is in mph),  $R$  in in/h for (4.11), and  $L$  in km.

<sup>2</sup>Default CAF/SAF by weather category are tabulated in the proposed HCM reliability chapters (Exhibit 36–26 and related tables).

<sup>3</sup>HCM provides explicit coefficients and fallback constants for “wet pavement, no rain” (e.g. fixed SAF) and “snow/ice on pavement, no snowfall”. We adopt those defaults when zero rates, but surfaces remain wet/icy.

<sup>4</sup>The reliability appendix gives default capacity adjustment factors by weather type and intensity; these are applied multiplicatively to the base (clear, dry) capacity per lane.

<sup>5</sup>See Transportation Research Record results quantifying free-flow speed and capacity reductions by precipitation/visibility class, with negligible change in jam density.

## (6) Calibration protocol

We follow a two-stage procedure: (i) fix SAF, CAF by HCM lookup (or by (4.11) plus HCM CAF), then (ii) estimate  $(\alpha, \beta)$  by non-linear least squares under clear-dry conditions and *hold* them fixed when weather scalings are applied; alternatively, allow mild re-optimisation with a minor ridge penalty to preserve comparability. Performance is reported overall and stratified by weather regime (explicit vs. rain vs. wet-no-rain), using RMSE, MAE, MAPE,  $R^2$ , and reliability coverage where applicable.

TABLE 4.6 F6 (EF-BPR): parameters, meanings, initial values, and bounds.

Parameter	Meaning	Initial value	Bounds / constraints
$\alpha$	Congestion sensitivity	0.15	[0.01, 1.00]
$\beta$	Convexity exponent	4.0	[1.5, 8.0]
$C_0$	Nominal capacity (no-external) vph/lane	3 lanes $\times$ 2000 vph	[4500, 7200]
$\boldsymbol{\delta}$	Capacity multipliers for external factors ( $\phi = \exp[-\boldsymbol{\delta}^\top \mathbf{z}]$ )	zeros	$\delta_i \geq 0; \phi \in [0.5, 1]$ (truncation)
$\eta$ (alt.)	External-factor effects on $\alpha$	zeros	implicit positivity via $\exp(\cdot)$
$\gamma$ (alt.)	External-factor effects on $\beta$	zeros	implicit positivity via $\exp(\cdot)$
Regularisation	Penalise collinearity/time confounding	con- $\lambda = 0.1$	$\lambda \in [0, 1]$ (CV tuned)

## 4.4.5 M5 [2, 9]: Machine Learning Hybrid BPR (ML-hBPR)

**Rationale:** Cluster D suggests data-driven improvements. We use a *residual* or *parameter-mapping* hybrid rather than a pure black box to preserve interpretability.

### (1) Residual form

$$T_t = T_0 \left[ 1 + \alpha \left( \frac{V_t}{C} \right)^\beta \right] \times (1 + r(\mathbf{x}_t)), \quad r(\cdot) \in [-0.5, 0.5], \quad (4.13)$$

where  $r(\cdot)$  is learned by a shallow gradient-boosted regressor with *monotonic constraint* in  $V/C$  to prevent non-physical shapes.

### (2) Parameter-mapping form (alternative)

$$\alpha(\mathbf{x}) = \exp\{g_\alpha(\mathbf{x})\}, \quad \beta(\mathbf{x}) = \exp\{g_\beta(\mathbf{x})\}, \quad (4.14)$$

where  $g_\alpha, g_\beta$  are small trees/NNet trained on training folds only. We use blocked CV and early stopping to avoid leakage/overfitting.

### (3) Parameter Bounds and Initialisation

TABLE 4.7 F5 (ML-hBPR): residual/parameter-mapping settings and constraints.

Component	Meaning	Initial value	Bounds / constraints
$r(\mathbf{x})$	Multiplicative residual corrector	Zero-mean init.	Monotone in $V/C$ ; $r \in [-0.5, 0.5]$
Model depth	GBM tree depth / NN hidden width	Depth 3 / width 16	Depth [2, 5]; early stopping
Learning rate	GBM learning rate / NN step size	0.05	[0.01, 0.2]
Features	Inputs to $r(\cdot)$ or $g_\alpha, g_\beta$	$V/C$ , time dummies, HGV share	No future leakage; time-block CV
Regularisation	$L_2$ or shrinkage	0.001	Tune via CV; avoid overfit

#### 4.4.6 M6 [10]: Stochastic-Capacity / Reliability BPR (SC-BPR)

**Rationale:** Cluster E highlights travel-time unreliability. We evaluate two complementary implementations.

##### (1) Canonical form

###### (i) Quantile-BPR (percentile fit)

Fit BPR to the  $p$ -th conditional percentile  $T^{(p)}$  via pinball loss:

$$\min_{\alpha^{(p)}, \beta^{(p)}} \sum_t \rho_p \left( T_t - T_0 \left[ 1 + \alpha^{(p)} \left( \frac{V_t}{C} \right)^{\beta^{(p)}} \right] \right), \quad \rho_p(u) = u(p - \mathbb{I}\{u < 0\}). \quad (4.15)$$

We use  $p \in \{0.5, 0.9\}$  to obtain median and 90th-percentile curves.

###### (ii) Lognormal inflation (predictive band).

Assume  $\log T \sim \mathcal{N}(\log T_{\text{det}}, \sigma^2)$  with  $T_{\text{det}} = T_0[1 + \alpha(V/C)^\beta]$ . Then the  $p$ -th percentile is

$$T^{(p)} = T_{\text{det}} \exp(z_p \sigma), \quad z_p = \Phi^{-1}(p). \quad (4.16)$$

Estimate  $\sigma$  from residuals of the deterministic fit. We report *coverage* (share of observations below  $T^{(p)}$ ) as a reliability metric.

##### (2) Parameter Bounds and Initialisation: F4 (SC-BPR)

TABLE 4.8 F4 (SC–BPR): parameters, meanings, initial values, and bounds.

Parameter	Meaning	Initial value (heuristic)	Bounds / constraints
$\alpha^{(p)}, \beta^{(p)}$	Quantile-BPR parameters (e.g. $p = 0.9$ )	From median fit as warm start	$\alpha^{(p)} \in [0.01, 1.0], \beta^{(p)} \in [1.5, 8.0]$
$\sigma$	Lognormal dispersion for predictive band	Residual sd of log-error	[0.05, 0.60]
$p$	Target percentile	0.90	[0.50, 0.95] (study design)
$C_{\text{eff}}$	Reliability-adjusted capacity (alt.)	$0.9 \times C$	$[0.8, 1.0] \times C$

#### 4.4.7 Common Calibration and Evaluation Protocol

##### (1) Data Splits

Train in the first three weeks of September 2024, and test in the last week.

##### (2) Objective and optimisers

Deterministic fits use NLS (Levenberg–Marquardt) with multiple random initialisations; piecewise thresholds via grid search; ML via gradient boosting with monotone constraint in  $V/C$ .

##### (3) Constraints

$\alpha > 0, \beta > 1$ ; continuity at  $\tau$  for piecewise; bounds  $\text{PCE}_{\text{HGV}} \in [1.1, 2.5], \lambda \in [0, 1]$ ; FD parameters  $v_f, \rho_j, w > 0$ .

##### (4) Metrics

RMSE, MAE, MAPE,  $R^2$ , 90th-percentile absolute error, runtime per 1,000 predictions; report by regime (free/rising/oversaturated), by weekday/weekend, and by  $s_{\text{HGV}}$  bins.

##### (5) Robustness

Bootstrap CIs for  $(\alpha, \beta)$ ; sensitivity to  $\tau$  and  $\text{PCE}_{\text{HGV}}$ ; coverage of  $T^{(0.9)}$  (F4).

### 4.5 Performance Metrics Definition

Throughout the benchmarking, we will use several standard metrics to compare model performance: A single error metric is insufficient for a comprehensive comparison of VDFs. A model might excel in one area (e.g., low average error) while failing in another (e.g., poor performance under congestion). Therefore, a multi-criteria evaluation framework is essential for a nuanced and insightful benchmarking analysis. This framework, inspired by the comprehensive approach in Mohammadian et al. [58], groups performance

## Benchmarking Methodology

---

metrics into three distinct categories: accuracy, replication of phenomena, and implementation considerations.[5] This structure ensures that the models are judged based on their statistical accuracy, theoretical validity, and practical utility. The selected metrics are detailed in Table 4.2.

TABLE 4.9 Performance Metrics for Benchmarking Analysis

Category	Metric	Purpose
<b>1. Accuracy of Operational Measures</b>	Root Mean Square Error (RMSE) $\sqrt{\frac{1}{N} \sum (T_{pred} - T_{obs})^2}$	Measures the average magnitude of prediction error in seconds, penalising larger errors more heavily.
	Mean Absolute Percentage Error (MAPE) $\frac{100\%}{N} \sum \left  \frac{T_{pred} - T_{obs}}{T_{obs}} \right $	Measures the average prediction error as a percentage, providing a scale-independent measure of accuracy.
<b>2. Replication of Traffic Phenomena</b>	Goodness-of-Fit ( $R^2$ ) - Uncongested ( $V/C < 0.7$ )	Assesses model performance specifically in free-flow and stable flow conditions.
	Goodness-of-Fit ( $R^2$ ) - Congested ( $V/C \geq 0.7$ )	Assesses model performance specifically during the transition to and within congested states.
	Bias $\frac{1}{N} \sum (T_{pred} - T_{obs})$	Determines if a model systematically over-predicts (positive bias) or under-predicts (negative bias) travel times.
<b>3. Implementation Considerations</b>	Number of Parameters	A measure of model complexity and parsimony. Fewer parameters are generally preferred if accuracy is comparable.
	Parameter Interpretability	A qualitative assessment of whether the calibrated parameters have a clear and plausible physical meaning.

- **Root Mean Squared Error (RMSE):** RMSE =  $\sqrt{\frac{1}{N} \sum_{i=1}^N (\hat{T}_i - T_i)^2}$ , where  $\hat{T}_i$  is the model-predicted travel time and  $T_i$  is observed, and  $N$  is number of data points. This gives an absolute error measure in seconds (or whichever unit  $T$  is in). RMSE penalises large errors more due to squaring, which is useful since we particularly care about big underestimations of delay.

- **Mean Absolute Percentage Error (MAPE):**  $\text{MAPE} = \frac{100\%}{N} \sum_i \left| \frac{\hat{T}_i - T_i}{T_i} \right|$ . This measures relative error as a percentage of actual travel time. It's scale-free, allowing error comparison during free-flow (when  $T$  is small) vs congested (when  $T$  is large). For instance, a 10s error is huge if  $T = 90$  (11%), but minor if  $T = 180$  (5%), and MAPE reflects that. We will use MAPE to see if models are consistently within, say, 5–10% of actual times.
- **Coefficient of Determination ( $R^2$ ):** This indicates the fraction of variance in  $T$  explained by the model.  $R^2 = 1 - \frac{\sum(\hat{T}-T)^2}{\sum(T_{\text{mean}}-T)^2}$ . An  $R^2$  closer to 1 means the model captures most variability in travel times. We expect the classical BPR (standard params) to have a moderate  $R^2$ , perhaps (maybe 0.5–0.7 based on literature for a single link). Improved models hopefully raise it closer to 0.8 or above by explaining more variance (especially temporal variance). We will report  $R^2$  for each variant on calibration and validation sets to gauge fit vs generalisation [36].
- **Bias measures:** We will check for any systematic bias (e.g. mean error, or whether the model tends to underpredict or overpredict). While BPR forms are unbiased in a symmetric error sense if calibrated by least squares, specific variants like piecewise could introduce slight bias in different regions. We will mention if a model always overshoots at low or undershoots at high volumes, indicating a form mismatch [60].
- **Computational efficiency:** Though not a numerical metric per se, we will comment on the relative complexity (e.g. ML model training time, or iteration needed for dynamic model). For instance, the ANN might take a few seconds to train, which is trivial here, but conceptually, if we scaled up to network-wide, a parametric model would be faster in assignment iterations.

By defining and using these consistent metrics, we ensure an objective comparison. Each variant has different strengths – some minimise maximum error, others overall RMSE. Thus, multiple metrics are needed to evaluate them thoroughly. Ultimately, selecting a “best” model may depend on which performance aspect one prioritises (average accuracy vs. reliability vs. interpretability), a point we will discuss in the conclusion.

# Chapter 5

## Results and Comparative Evaluation

This chapter outlines the structure of results that will be obtained once the models described in Chapter 4 are calibrated and run. Since this is a Year 2 progress report, the actual numerical results and figures are placeholders, indicating how the comparisons will be presented. This chapter reports the planned structure of results and the templates for quantitative comparison of the six benchmark methods introduced in Chapter 4:

- M1 Dynamic–Parameter BPR (DP–BPR),
- M2 Fundamental–Diagram–informed VDF (FD–VDF),
- M3 Multi–Class BPR (MC–BPR),
- M4 External–Factor–Adjusted BPR (EF–BPR),
- M5 Machine–Learning Hybrid BPR (ML–hBPR)
- M6 Stochastic–Capacity / Reliability BPR (SC–BPR).

At the time of writing, complete model fitting is ongoing. Therefore, this chapter provides a reproducible reporting framework with clearly marked placeholders (*TBC*) for numerical results.

### 5.1 Evaluation Design (fixed before analysis)

#### 5.1.1 Datasets and splits

All models will be trained and evaluated on the National Highways M67 dataset (September 2024), as curated in Chapter 3. To avoid temporal leakage, we adopt a \*\*blocked split\*\*:

- **Training/validation:** Weeks 1–3 (*TBCdates*) with rolling validation windows.
- **Hold-out test:** Week 4 (*TBCdates*). No refitting on the test set.

Where daily calibration is required (e.g. M1), parameters are learned on Week 1–3 and applied to Week 4.

### 5.1.2 Outcome and metrics

The primary outcome is the link travel time  $T$  (seconds) per 15-minute interval. We report, for each method and data subset, the following metrics:

$$\begin{aligned} \text{RMSE} &= \sqrt{\frac{1}{n} \sum_i (\hat{T}_i - T_i)^2}, & \text{MAE} &= \frac{1}{n} \sum_i |\hat{T}_i - T_i|, \\ \text{MAPE} &= \frac{100}{n} \sum_i \left| \frac{\hat{T}_i - T_i}{T_i} \right|, & R^2 &= 1 - \frac{\sum_i (\hat{T}_i - T_i)^2}{\sum_i (T_i - \bar{T})^2}, \\ \text{Bias} &= \frac{1}{n} \sum_i (\hat{T}_i - T_i), & \text{P95Err} &= 95\text{th percentile of } |\hat{T}_i - T_i|. \end{aligned}$$

For the reliability model (M6), we additionally report the *prediction interval coverage probability* (PICP) for nominal 90% bands and the mean interval width (MIW).

All metrics are reported overall and by strata: AM peak (07:00–09:00), PM peak (16:00–18:00), inter-peak (09:00–16:00), weekend, and high-HGV periods (top decile of HGV share).

## 5.2 Headline Results

TABLE 5.1 Overall test-set accuracy (Week 4).

Method	RMSE (s)	MAE (s)	MAPE (%)	$R^2$	Bias (s)	P95Err (s)
M1 DP-BPR	TBC	TBC	TBC	TBC	TBC	TBC
M2 FD-VDF	TBC	TBC	TBC	TBC	TBC	TBC
M3 MC-BPR	TBC	TBC	TBC	TBC	TBC	TBC
M4 EF-BPR	TBC	TBC	TBC	TBC	TBC	TBC
M5 ML-hBPR	TBC	TBC	TBC	TBC	TBC	TBC
M6 SC-BPR (mean)	TBC	TBC	TBC	TBC	TBC	TBC

Table 5.1 is the planned interpretation template. We will first compare RMSE/MAE across methods. We *expect* ML-hBPR (M5) to minimise overall RMSE, while FD-VDF (M2) is anticipated to reduce tail errors (P95Err) around oversaturated conditions. DP-BPR (M1) should reduce peak-period bias relative to a static BPR. MC-BPR (M3) may show modest gains overall, given low HGV share, but improved fit in high-HGV strata. SC-BPR (M6) will be evaluated on reliability criteria (PICP, MIW) and mean accuracy.

### 5.3 Stratified Performance

TABLE 5.2 Peak vs. inter-peak and weekend accuracy (test set).

Method	AM Peak	PM Peak	Inter-peak	Weekend
M1 DP–BPR	TBC	TBC	TBC	TBC
M2 FD–VDF	TBC	TBC	TBC	TBC
M3 MC–BPR	TBC	TBC	TBC	TBC
M4 EF–BPR	TBC	TBC	TBC	TBC
M5 ML–hBPR	TBC	TBC	TBC	TBC
M6 SC–BPR (mean)	TBC	TBC	TBC	TBC

### 5.4 Calibration Diagnostics

#### 5.4.1 Parameter estimates

TABLE 5.3 Key parameter estimates per method (training on Weeks 1–3).

Method	Parameter	Estimate	SE / 95% CI
M1 DP–BPR	$\alpha_{AM}, \beta_{AM}$	TBC	TBC
	$\alpha_{PM}, \beta_{PM}$	TBC	TBC
M2 FD–VDF	threshold $\theta$ (or $\rho_c$ )	TBC	TBC
	$(\alpha_1, \beta_1) / (\alpha_2, \beta_2)$	TBC	TBC
M3 MC–BPR	HGV equivalency $e_H$	TBC	TBC
M4 EF–BPR.	weather coefficient(s)	TBC	TBC
M6 SC–BPR	reliability terms ( $g, d$ or $\lambda$ )	TBC	TBC

#### 5.4.2 Goodness-of-fit plots

Placeholder for: (a)  $V/C$  vs.  $T$  scatter with fitted curves for M1–M6;  
 (b) residuals vs.  $V/C$ ; (c) QQ-plot of residuals; (d) calibration curve for PIs  
 (M6).

FIGURE 5.1 Diagnostic plots to be inserted after calibration.

## 5.5 Reliability Analysis for M6 (placeholders)

TABLE 5.4 Prediction-interval performance for SC–BPR (M6) on the test set.

	Nominal 90%	Observed PICP (%)	Mean Interval Width (s)
All periods	90	TBC	TBC
AM peak	90	TBC	TBC
PM peak	90	TBC	TBC
Weekend	90	TBC	TBC
High-HGV decile	90	TBC	TBC

## 5.6 Ablation and Robustness

### 5.6.1 Ablations

- **M1 smoothing:** compare piecewise-constant vs. rolling updates of  $(\alpha, \beta)$ .
- **M2 threshold:** fixed  $\theta = 0.8$  vs. data-driven  $\theta$  (minimising validation RMSE).
- **M3 HGV:** fit with/without  $e_H$  when HGV share < 5%.
- **M4 weather:** dry-only vs. all days (once weather is merged).
- **M5 complexity:** 1-hidden-layer vs. gradient boosting (same features).

### 5.6.2 Robustness checks

- **Temporal robustness:** train on Weeks 2–4, test on Week 1.
- **Distribution shift:** weekday-only training, weekend test.
- **Bootstrap CIs:** day-level block bootstrap for RMSE differences (1,000 resamples).

# Chapter 6

## Conclusion and Forward Plan (Year 2)

### 6.1 Synthesis of Year-2 Contributions

#### (1) Evidence base and taxonomy

Completed a protocol-guided systematic review (2010–2025) and consolidated a corpus of 48 studies. Produced a transparent screening flow, an A–G taxonomy (dynamic parameterisation/structural optimisation; multi-class; external-factor integration; machine-learning; uncertainty; policy; data fusion), and a crosswalk to six benchmarkable method families (M1–M6).

Delivered bibliometric mappings (yearly counts and topical proportions), an evidence map, and a reproducible labelling protocol to ensure consistent cluster assignment.

#### (2) Data, pre-processing, and EDA

Ingested and audited September 2024 National Highways data for M67 (15-minute resolution). Established a reproducible pipeline for time alignment, de-duplication, gap handling, and outlier control; reconstructed travel time where needed and derived density and  $V/C$ .

Estimated a defensible static capacity  $C$  (upper-envelope and peak-percentile methods) and defined regime bands for subsequent calibration and evaluation.

#### (3) Benchmark design and model readiness

Fixed six benchmark families with explicit equations, constraints, and calibration routines: M1 Dynamic-Parameter BPR; M2 Fundamental-Diagram-informed VDF / piecewise BPR; M3 Multi-Class (HGV-adjusted) BPR; M4 External-Factor-adjusted BPR (weather-integrated); M5 Machine-Learning hybrid; M6 Stochastic-Capacity / reliability BPR.

Specified a common metric suite (RMSE, MAE, MAPE,  $R^2$ ) with reliability measures (quantile errors and coverage), and a blocked train/validation/test split to avoid temporal leakage.

### (4) Key empirical insights from EDA (decision-relevant)

Clear weekday bimodality and sharper degradation of speed at high flow support *time-regime* and *two-regime* formulations (M1, M2). And HGV share is low but time-varying; multi-class adjustment (M3) is expected to yield marginal but targeted benefits in specific windows. Meanwhile the Tail behaviour (wider travel-time dispersion at high  $V/C$ ) motivates reliability-oriented modelling (M6) and argues for reporting percentile performance, not just mean fit.

## 6.2 Conclusions

### (1) Dynamic responsiveness is essential

A single set of  $(\alpha, \beta)$  cannot capture off-peak and peak conditions without bias. Allowing parameters to vary across interpretable time regimes (M1) is a pragmatic, assignment-compatible step that remains transparent for planning use.

### (2) Structure matters near capacity

Two-regime BPR (or FD-informed VDF) provides the necessary curvature to reduce underestimation in near-saturation. The EDA-inferred bend offers a data-driven threshold for the piecewise specification (M2).

### (3) Factor expansion should be parsimonious

Given the limited variance in  $p_H$  on M67, multi-class adjustment (M3) should be conservative and validated with sensitivity bands. External factors (M4) are promising, but require careful alignment and missing-data handling to avoid spurious gains.

### (4) Uncertainty must be explicit

Reliability metrics (quantile errors, coverage) will complement mean-error metrics so that the proposed function is useful for operations, not only for average-case planning (M6).

## 6.3 Forward Plan (next 12–18 months)

### 6.3.1 Objective O1: Complete the benchmark on M67 (all M1–M6)

- *Calibration protocol:* (i) blocked time split; (ii) identical inputs across families; (iii) monotonicity and positivity constraints; (iv) hyperparameter search bounded by literature-informed priors.
- *Reports:* overall metrics; regime-wise metrics by  $V/C$  band and peak/off-peak; reliability table (90th/95th percentile errors); ablation notes for each additional factor.
- *Acceptance thresholds:* each candidate must improve at least one reliability metric by  $\geq 10\%$  versus classical BPR without degrading RMSE by  $> 5\%$ .

### 6.3.2 Objective O2: Integrate external factors and optimise parameters

- *Data integration (M4):* merge Met Office weather (precipitation intensity, visibility, wind) and any available incident logs at 15-minute granularity; define robust imputation and indicator design (e.g. rain/heavy-rain dummies; visibility tiers).
- *Parameter optimisation:* Bayesian optimisation for  $(\alpha, \beta)$  and factor multipliers with cross-validation and stability checks; identifiability diagnostics (profile likelihoods; ridge-trace style plots) to prevent over-fitting.

### 6.3.3 Objective O3: Reliability-oriented modelling

- *Stochastic capacity (M6):* estimate effective capacity distributions and calibrate percentile-targeted variants (e.g. 90th percentile travel time curves). Report coverage of observed tails and confidence bands for deployment.

### 6.3.4 Objective O4: Development of a New BPR Function

#### (1) Aim

To design a demand-responsive BPR (DR-BPR) that synthesises the strongest elements observed across M1–M6, whilst remaining parsimonious, identifiable and network-assignment-ready.

#### (2) Success criteria

- Out-of-sample RMSE/MAPE improvement  $\geq 10\text{--}20\%$  over calibrated classical BPR;
- Reduced peak-period bias;
- Stable parameters across weeks;
- Differentiability for assignment/optimisation.

#### (3) Deployment

Embed DR-BPR in static User Equipment(UE) and Dynamic Traffic Assignment DTA Problem.

## 6.4 Milestones, deliverables and risks

### 6.4.1 Milestones

TABLE 6.1 Planned milestones and deliverables ( $R2 \rightarrow$  early  $R3$ ).

Quarter	Milestone	Deliverables
Q1 (now– +1.5 mo)	Finalise M1–M3 calibration on M67	Reproducible code; metric tables; residual diagnostics; sensitivity to $p_H$ .
Q2 (+1.5– +2.5 mo)	Weather integration and M4 calibration	Matched weather dataset; factor design doc; ablation study; improved reliability metrics.
Q3 (+2.5– +3 mo)	M6 (stochastic) and percentile curves	Coverage analysis; uncertainty bands; deployment note.
Q4 (+3– +12 mo)	Selection and network embedding	Chosen function embedded in UE; convergence and realism checks; transfer test on extra month/link.

### 6.4.2 Risks and mitigations

- *Sparse high-saturation samples*: extend data window; prioritise M2/M6 with explicit uncertainty reporting.
- *Noisy external factors*: use blocked matching, robust indicators, and pre-registered ablations to avoid over-fitting.
- *Identifiability of additive factors*: constrain multipliers with literature priors and apply cross-regime stability tests.

## 6.5 Publications plan

### (1) Paper A: Benchmarking next-generation BPR families on a UK motorway link

- Scope: M1–M6 on M67; unified metrics; reliability emphasis; transparent reproducibility pack.
- Target venues: *Transportation Research Part C/A, IEEE T-ITS*.

### (2) Paper B: Weather-integrated, reliability-aware link performance functions

- Scope: factor design, identifiability, percentile-fit performance; transferability across months/links.
- Target venues: *Transportation Research Part B/C, CEUS*.

### **(3) Paper C: Embedding the selected function in assignment and network optimisation**

- Scope: UE/DNUE embedding, convergence, and network performance improvements under demand variability [6, 54].
- Target venues: *TR-B, Networks and Spatial Economics.*

# Bibliography

- [1] Rafał Kucharski and Arkadiusz Drabicki. Estimating macroscopic volume delay functions with the traffic density derived from measured speeds and flows. *Journal of Advanced Transportation*, 2017(1):4629792, 2017.
- [2] Hong Chen, Han Zeng, Ji-Biao Zhou, and Rong Zeng. *Improvement on the BPR Link Performance Function and Classification Calibration in Parameter*, pages 373–382.
- [3] Aulia Rahman, Muhammad Zudhy Irawan, and Mukhammad Rizka Fahmi Amrozi. A new approach to estimating link performance on indonesian urban roads: deriving the bpr 1964 function. In *IOP Conference Series: Materials Science and Engineering*, volume 1289, page 012043. IOP Publishing, 2023.
- [4] Yuyan “Annie” Pan, Han Zheng, Jifu Guo, and Yanyan Chen. Modified volume-delay function based on traffic fundamental diagram: A practical calibration framework for estimating congested and uncongested conditions. *Journal of transportation engineering, Part A: Systems*, 149(11):04023112, 2023.
- [5] Robert Neuhold and Martin Fellendorf. Volume delay functions based on stochastic capacity. *Transportation research record*, 2421(1):93–102, 2014.
- [6] Roza E Barka and Ioannis Politis. Fitting volume delay functions under interrupted and uninterrupted flow conditions at greek urban roads. *European Transport-Trasporti Europei*, 127(1):13–20, 2001.
- [7] N Hossam, U Gazder, M Al Hashimi, N Salah, and S Abdulameer. Estimation of time delay functions for design of traffic systems. In *IET Conference Proceedings CP859*, volume 2023, pages 17–22. IET, 2023.
- [8] Seongsoon Yun, Wade W White, Daniel R Lamb, and Yongqiang Wu. Accounting for the impact of heavy truck traffic in volume–delay functions in transportation planning models. *Transportation research record*, 1931(1):8–17, 2005.
- [9] Lee Klieman, Wang Zhang, Vincent L Bernardin Jr, and Vladimir Livshits. Estimation and comparison of volume delay functions for arterials and freeway hov and general purpose lanes. Technical report, 2011.
- [10] Stefano Manzo, Otto Anker Nielsen, and Carlo Giacomo Prato. Effects of uncertainty in speed–flow curve parameters on a large-scale model: Case study of the danish national model. *Transportation Research Record*, 2429(1):30–37, 2014.
- [11] Aathira K Das and Bhargava Rama Chilukuri. Link cost function and link capacity for mixed traffic networks. *Transportation Research Record*, 2674(9):38–50, 2020.

## BIBLIOGRAPHY

---

- [12] Benton J Underwood. Ten years of massed practice on distributed practice. *Psychological Review*, 68(4):229, 1961.
- [13] Highway Capacity Manual. Hcm2d10. *Transportation Research Board, National Research Council, Washington, DC*. <https://www.sciencedirect.com/topics/engineering/highway-capacity-manual>, 2010.
- [14] Heinz Spiess. Conical volume-delay functions. *Transportation Science*, 24(2):153–158, 1990.
- [15] Yuyan Pan, Han Zheng, Jifu Guo, and Yanyan Chen. Modified volume–delay function based on traffic fundamental diagram: A practical calibration framework for estimating congested and uncongested conditions. *Journal of Transportation Engineering, Part A: Systems*, 149(6):04023023, 2023.
- [16] Yasir Ali, Fida Hussain, and Md. Mizanul Haque. Advances, challenges, and future research needs in machine learning-based crash prediction models: A systematic review. *Accident Analysis & Prevention*, 194:107378, 2024.
- [17] Saeed Mohammadian, Zuduo Zheng, Md. Mazharul Haque, and Ashish Bhaskar. Performance of continuum models for realworld traffic flows: Comprehensive benchmarking. *Transportation Research Part B: Methodological*, 147:132–167, 2021.
- [18] Yasir Ali, Anshuman Sharma, and Zuduo Zheng. Empirical research on car-following and lane-changing: Recent developments, emerging vehicle technologies’ impact, and future research needs. *Transportation Research Interdisciplinary Perspectives*, 31:101368, 2025.
- [19] Konstantin Dragomiretskiy and Dominique Zosso. Variational mode decomposition. *IEEE transactions on signal processing*, 62(3):531–544, 2013.
- [20] Proposed hcm chapter 37: Travel time reliability — supplemental. TRB/National Academies Press, 2013. Eq. 37–66: free-flow speed adjustment factor as a linear function of rainfall/snowfall rates.
- [21] Joseph S Drake. A statistical analysis of speed density hypothesis. *HRR*, 154:53–87, 1967.
- [22] Han Ru, Jinliang Xu, Zhihao Duan, Xingliang Liu, and Chao Gao. Modelling travel time after incidents on freeway segments in china. *IEEE Access*, 7:162465–162475, 2019.
- [23] Yu-Qing Wang, Chao-Fan Zhou, Bin Jia, and Hua-Bing Zhu. Reliability analysis of degradable networks with modified bpr. *Modern Physics Letters B*, 31(36):1750353, 2017.
- [24] Qixiu Cheng, Zhiyuan Liu, Yuqian Lin, and Xuesong Simon Zhou. An s-shaped three-parameter (s3) traffic stream model with consistent car following relationship. *Transportation Research Part B: Methodological*, 153:246–271, 2021.
- [25] KB Davidson. A flow travel time relationship for use in transportation planning. In *Australian Road Research Board (ARRB) Conference, 3rd, 1966, Sydney*, volume 3, 1966.

## BIBLIOGRAPHY

---

- [26] Xingliang Zhang, Dandan Song, Zhuo Liu, Yukun Jing, Jie Pu, and Haitao Yu. A new method for calculating traffic delay based on modified bpr function model. In *International Conference on Cryptography, Network Security, and Communication Technology (CNSCT 2023)*, volume 12641, pages 292–297. SPIE, 2023.
- [27] Zhaoyang Lu, Qiang Meng, and Gabriel Gomes. Estimating link travel time functions for heterogeneous traffic flows on freeways. *Journal of Advanced Transportation*, 50(8):1683–1698, 2016.
- [28] Nan He and Sheng-chuan Zhao. Urban road traffic impedance function—dalian city case study. *Journal of Highway and Transportation Research and Development (English Edition)*, 8(3):90–95, 2014.
- [29] Ehsan Mazloumi, Sara Moridpour, and Hassan Mohsenian. Delay function for signalized intersections in traffic assignment models. *Journal of Urban Planning and Development*, 136(1):67–74, 2010.
- [30] Yu ZHANG and Shuanglin JIANG. Function of journey time on dynamic highway section compare with bpr function. *Communications Science and Technology Heilongjiang*, (7).
- [31] Gerard Casey, Bingyu Zhao, Krishna Kumar, and Kenichi Soga. Context-specific volume–delay curves by combining crowd-sourced traffic data with automated traffic counters: A case study for london. *Data-Centric Engineering*, 1:e18, 2020.
- [32] Rahmi Akçelik. Travel time functions for transport planning purposes: Davidson’s function, its time dependent form and alternative travel time function. *Australian Road Research*, 21(3), 1991.
- [33] Rui Feng She and Yanfeng Ouyang. Generalized link cost function and network design for dedicated truck platoon lanes to improve energy, pavement sustainability and traffic efficiency. *Transportation Research Part C: Emerging Technologies*, 140:103667, 2022.
- [34] Jinbiao Huo, Xinhua Wu, Cheng Lyu, Wenbo Zhang, and Zhiyuan Liu. Quantify the road link performance and capacity using deep learning models. *IEEE Transactions on Intelligent Transportation Systems*, 23(10):18581–18591, 2022.
- [35] Fei Zhao, Liping Fu, Ming Zhong, Shaobo Liu, Xudong Wang, Junda Huang, and Xiaofeng Ma. Development and validation of improved impedance functions for roads with mixed traffic using taxi gps trajectory data and simulation. *Journal of Advanced Transportation*, 2020(1):7523423, 2020.
- [36] Salvatore Manzo, Otto Anker Nielsen, and Carlo Giacomo Prato. Effects of uncertainty in speed–flow curve parameters on a large-scale model: Case study of the danish national model. *Transportation Research Record*, 2429:32–41, 2014.
- [37] Pushkin Kachroo and Shankar Sastry. Traffic assignment using a density-based travel-time function for intelligent transportation systems. *IEEE Transactions on Intelligent Transportation Systems*, 17(5):1438–1447, 2016.
- [38] Masayuki ANDOU, Jun-ichi TAKAYAMA, and Sho-ichiro NAKAYAMA. Presumption of travel time by a bpr function that considers the vertical inclination on the road. In *Proceedings of the Eastern Asia Society for Transportation Studies Vol. 8*

## BIBLIOGRAPHY

---

(*The 9th International Conference of Eastern Asia Society for Transportation Studies*, 2011), pages 332–332. Eastern Asia Society for Transportation Studies, 2011.

- [39] Richard Dowling, Alexander Skabardonis, Vassili Alexiadis, et al. Traffic analysis toolbox, volume iii: Guidelines for applying traffic microsimulation modeling software. Technical report, United States. Federal Highway Administration. Office of Operations, 2004.
- [40] Sang-Jin Oh, Sang-Hyuk Park, and Byung-Ho Park. Development of travel time functions considering intersection delay. *Journal of Korean Society of Transportation*, 26(4):63–76, 2008.
- [41] Johannes Asamer and Martin Reinthaler. Estimation of road capacity and free flow speed for urban roads under adverse weather conditions. In *13th International IEEE Conference on Intelligent Transportation Systems*, pages 812–818. IEEE, 2010.
- [42] Feng Jiang, Zhen-ni Lu, Min Gao, and Da-ming Luo. Dp-bpr: Destination prediction based on bayesian personalized ranking. *Journal of Central South University*, 28(2):494–506, 2021.
- [43] Sang Nguyen. An algorithm for the traffic assignment problem. *Transportation Science*, 8(3):203–216, 1974.
- [44] Ning Wang, Jian Gu, Wei Jing, Fu-An Huang, and Hang Zhu. Study of an impedance function for mixed traffic flows considering the travel time–cost characteristics of long-distance electric vehicle trips. *Applied Sciences*, 14(16):7138, 2024.
- [45] Hesham Rakha, Mohamadreza Farzaneh, Mazen Arafah, and Emily Sterzin. Inclement weather impacts on freeway traffic stream behavior. *Transportation Research Record*, 2071:8–18, 2008. Provides empirical weather adjustment factors to free-flow speed and capacity; jam density approximately unaffected.
- [46] N Liu, SC Zhao, and N He. Further study of impedance function based on bpr function. *Journal of Wuhan University of Technology (Transportation Science & Engineering)*, 37:545–548, 2013.
- [47] Ninad Gore, Shriniwas Arkatkar, Gaurang Joshi, and Constantinos Antoniou. Modified bureau of public roads link function. *Transportation Research Record*, 2677(5):966–990, 2023.
- [48] Leslie C Edie. Car-following and steady-state theory for noncongested traffic. *Operations research*, 9(1):66–76, 1961.
- [49] Bruce D. Greenshields et al. A study of traffic capacity. *Highway Research Board Proceedings*, 14:448–477, 1935.
- [50] Rafał Kucharski and Arkadiusz Drabicki. Estimating macroscopic volume delay functions with the traffic density derived from measured speeds and flows. *Journal of Advanced Transportation*, 2017:Article ID 4629792, 2017. File: 14-Estimating Macroscopic Volume Delay Functions with the Traffic...
- [51] Seongsoon Yun, Wade W. White, Daniel R. Lamb, and Yongqiang Wu. Accounting for the impact of heavy truck traffic in volume–delay functions in transportation planning models. *Transportation Research Record*, 1931:8–17, 2005.

## BIBLIOGRAPHY

---

- [52] Mohsen Babaei, Mojtaba Rajabi-Bahaabadi, and Afshin Shariat-Mohaymany. Estimation of travel time reliability in large-scale networks. *Transportation Letters*, 8(4):229–240, 2016.
- [53] Junjie Zhang, Miaomiao Liu, and Bin Zhou. Analytical model for travel time-based bpr function with demand fluctuation and capacity degradation. *Mathematical Problems in Engineering*, 2019(1):5916479, 2019.
- [54] Denos C Gazis, Robert Herman, and Richard W Rothery. Nonlinear follow-the-leader models of traffic flow. *Operations research*, 9(4):545–567, 1961.
- [55] Lee Vien Leong. Delay functions in trip assignment for transport planning process. In *AIP Conference Proceedings*, volume 1892, page 060006. AIP Publishing LLC, 2017.
- [56] Roza E. Barka and Ioannis Politis. Fitting volume delay functions under interrupted and uninterrupted flow conditions at greek urban roads. *European Transport / Trasporti Europei*, (83):Paper 7, 2021. File: 4-Fitting Volume Delay Functions... (Greek urban roads).
- [57] Lee Klieman, Wang Zhang, Vincent L. Bernardin Jr., and Vladimir Livshits. Estimation and comparison of volume delay functions for arterials and freeway HOV and general purpose lanes. In *TRB 90th Annual Meeting Compendium of Papers*, pages 1–19, Washington, D.C., 2011. Paper #11-0242.
- [58] Adolf Darlington May. *Traffic flow fundamentals*. 1990.
- [59] N. Hossam and U. Gazder. Estimation of time delay functions for design of traffic systems. *Examples and Counterexamples*, 6:100151, 2024.
- [60] Yong Chen, Weishao Wang, Jian Wang, Kui Yu, and Xuesong Wang. Improvement on the BPR link performance function and classification calibration in parameter. In *CICTP 2012: Multimodal Transportation Systems, Proceedings of the 12th COTA International Conference of Transportation Professionals*, pages 250–262, Beijing, China, 2012. ASCE.