

INVESTIGATING ENERGY CONSUMPTION IN ELECTRIC VEHICLES

Harry Vincent Smith, II

A dissertation submitted to the University of Dublin in partial fulfilment of
the requirements for the degree Doctor of Philosophy.

Department of Civil, Structural, & Environmental Engineering
Trinity College Dublin, the University of Dublin

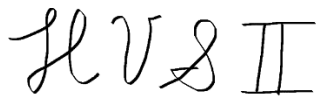
February 2026

Declaration

I declare that this thesis has not been submitted as an exercise for a degree at this or any other university and it is entirely my own work. I confirm that the work has been completed in full compliance with relevant university policies, including the Academic Integrity policy, the Trinity Policy on Good Research Practice, and the guidance on AI and GenAI in Teaching, Learning, Assessment and Research.

I agree to deposit this thesis in the University's open access institutional repository or allow the Library to do so on my behalf, subject to Irish Copyright Legislation and Trinity College Library conditions of use and acknowledgement.

I consent to the examiner retaining a copy of the thesis beyond the examining period, should they so wish (EU GDPR May 2018).

A handwritten signature in black ink, reading 'HVS II'. The letters are stylized and cursive, with the 'H' and 'V' being particularly prominent.

Harry Vincent Smith, II

30th September 2025

Abstract

As part of the fight to limit climate change, the European Union has set a target of zero-emissions from the transport sector in European cities by 2035. The main component of reducing emissions from transport is thought to be a complete transition from traditional internal combustion engine vehicles (ICEVs) to fully electric vehicles (EVs). However, EV sales are not on track to meet the ambitious goals set by government agencies like the European Union. Only 17% of new vehicles registered in Ireland in 2025 were EVs, but the goal is for that to reach 100% by 2030. One reason for the current scepticism around EVs is the phenomenon known as range anxiety, where drivers fear EVs lack the necessary range to complete desired trips.

An important piece of combatting range anxiety is developing better estimations of EV range. EV range is estimated both by the manufacturer to provide a selling point for the vehicle and by in-vehicle systems to advise drivers on how far the vehicle can travel before requiring a charge. However, these estimations can differ significantly from real-world driving scenarios due to a multitude of factors. Advances in the state-of-the-art literature, through developing models of energy consumption (EC) and understanding how different factors influence EC, can help to produce better estimations of EV range. To fill gaps identified in the literature, several objectives are set. The first objective is to build an EC model at a trip level. The accuracy of a macroscopic and a mesoscopic EC model are compared, and the factors influencing the model at each scale are also compared. The next objective is to find a method that relates driving behaviour to EC and determine how much driving behaviour can affect EC. The final objective is to model EC for realistic trips using a road network model and to analyse with these models how speed limits, route choice, and driving behaviour affect EC.

Data was collected from 36 EV drivers in Ireland to build an EC model. The EC model was built at the mesoscopic scale to model driving segments, which is shown to

decrease error by 53% compared to a macroscopic model when predicting EC of a trip. Additionally, the mesoscopic model is shown to be more comprehensive in modelling driving behaviour parameters, whereas the macroscopic model predicted EC mainly by distance travelled. At both scales, the most accurate models were artificial neural networks (ANNs) when compared to other machine learning (ML) methods. The model includes driving behaviour, environmental, and driver and vehicle parameters.

A sensitivity analysis (SA) is used to determine to what degree each modelled parameter affects EC. The results show that, among all driving segments, speed, gradient, time spent accelerating, and time spent decelerating have larger effects than any other parameter. It is shown that in both modelled and actual segments, the most energy efficient driving occurs at speeds between 20-60 km/h. Speeds above 100 km/h cause drastic increases in EC over the same distance travelled. As a result of the SA, a further investigation is conducted to determine how rates of acceleration, rates of deceleration, and speed affect EC at high and low speeds. The results show how moderate acceleration and deceleration rates can lead to improved energy efficiency at all speeds. The most energy efficient driving behaviour at high speeds can reduce EC by 71%. Additionally, it is demonstrated that common classification of driving behaviour into passive, moderate, and aggressive categories fails to sufficiently correlate with EC across different driving segments.

Finally, the EC model is applied to routes found through a model of the Ireland road network. Realistic drive cycles (DCs) are developed for routes found in the network model, and these DCs are used to estimate EC under different conditions such as reduced speed limits. Reduced speed limits have a significant effect on EC, especially on interurban trips. With the combined models, it is shown that drivers can save 10% of EC by slowing down, which results in an increase in travel time of 9%. In other words, increasing a journey time by about 15 minutes can save EV drivers about 5 kWh of battery power. The effect of eco-

driving behaviours and route choice on EC is also analysed by combining the EC and road network models.

Acknowledgements

I would like to thank first my family and friends who make me happy each day. My amazing mom and dad who love me unconditionally and who I always look forward to chatting with. Thank you, Emily, for being with me during this entire process and always encouraging me to do what I needed to. Thank you to my brothers, David and Charlie, for shaping who I am today. Thank you to all my friends for letting me enjoy my time during this long process. Even you, Mia. Without you all this dissertation is not possible.

Thank you to my supervisors, Margaret O'Mahony and Brian Caulfield, for believing in me and bringing me over here. I appreciated your guidance and attention to detail over the last three years that helped me grow confident as a researcher.

I would also like to thank European Union for funding my research as part of the POWERDRIVE project.

Publications

The following papers were published during the completion of this dissertation:

Smith, H., Akhtar, S., Caulfield, B., & O'Mahony, M. (2024). Validity of GPS data in driving cycles. *IET Intelligent Transport Systems*, 18, 3034-3040.

Akhtar, S., Smith, H., Caulfield, B., & O'Mahony, M. (2024, April). Markov Chain and K-means Clustering Analyses for Constructing Electric Vehicle Drive Cycles for Dublin, Ireland. In *Transport Research Arena Conference* (pp. 145-151). Cham: Springer Nature Switzerland.

Table of Contents

Chapter 1. Introduction	1
Chapter 2. Data Collection.....	7
2.1 Selection of Driving Behaviour Application.....	7
2.2 Selection of OBD Device.....	11
2.3 Data Collection Methodology	12
2.4 Participant Recruitment.....	14
2.5 Data Collection Process and Data Preprocessing	17
Chapter 3. Energy Consumption Model.....	19
3.1 Introduction.....	19
3.2. Literature Review.....	20
3.2.1 General Information - Energy Consumption of Electric Vehicles.....	20
3.2.2 Factors Influencing Energy Consumption.....	21
3.2.3 Energy Consumption Modelling	24
3.2.4 Data Collection Methodologies.....	28
3.2.5 Range Anxiety and Remaining Driving Range Prediction.....	29
3.2.6 Model Target – Energy Consumption and Energy Consumption Rate	31
3.2.7 Model Scale.....	33
3.2.8 Feature Importance.....	34
3.2.9 Summary and Gaps Found	35
3.3. Methodology	35
3.3.1 Data Processing: Streamlining the Database.....	35
3.3.2 Model Parameters.....	41
3.3.2.1 Driving Behaviour Parameters	42
3.3.2.2 GPS-Based Parameters.....	45
3.3.2.3 Driver and Vehicle Parameters	54
3.3.4 Machine Learning Models	57
3.3.5 Model Evaluation	61
3.3.6 Neural Network Hyperparameters	62
3.3.7 Feature Importance.....	66
3.3.8 Trip Segmentation	67
3.4 Results.....	69
3.4.1 Trip-Level Model Performance	69
3.4.2 Trip-Level Feature Importance.....	69
3.4.3 TEC Reliance on Distance and Model Alternatives	73
3.4.4 Model without Distance and ECR Models.....	77
3.4.5 Segment-Level Energy Consumption Model	80

3.4.6 Final Model Selection: Trip Energy Consumption as a Summation of Segments	84
3.5 Conclusion	88
Chapter 4. Driving Behaviour Effects on Energy Consumption and Sensitivity Analysis	91
4.1 Literature Review on Driving Behaviour.....	92
4.1.1 Driving Behaviour Definition and EC	93
4.1.2 Driving Behaviour Identification and Categorisation	95
4.2 Clustering Method.....	99
4.3 Limit-based Method	105
4.4 Empirical Method.....	112
4.5 Literature Review on Sensitivity Analysis.....	115
4.5.1 Sensitivity Analysis Introduction and Terms.....	115
4.5.2 Global Sensitivity Analysis Methods	116
4.5.3 Uncertainty Analysis	119
4.5.4 Sensitivity Analysis for Energy Consumption of Electric Vehicles	120
4.6 Sobol' Sensitivity Analysis	121
4.6.1 Sobol' Method Implementation.....	121
4.6.2 Sensitivity Analysis on Entire Range of Inputs.....	126
4.6.3 Sensitivity Analysis within Speed Intervals	130
4.7 Results of Driving Behaviour Classification on Sample Segments.....	140
4.7.1 Sample Segment Methodology	140
4.7.2 Sample Segment Results	143
4.7.3 Application of Driving Behaviour Classification Methods to Sample Segments	148
4.8 Discussion and Conclusion	151
Chapter 5. Estimating Energy Consumption and Journey Travel Time Under Varied Speed Limits Using Realistic Drive Cycles	155
5.1 Literature Review on Effects of Lowered Speed Limits.....	156
5.1.1 Speed Limit Reduction Studies.....	156
5.1.2 Effect of Speed on Energy Consumption.....	158
5.2 Literature Review on Drive Cycles.....	159
5.2.1 Route Selection	161
5.2.2 Data Collection	163
5.2.3 Driving Cycle Construction	165
5.2.3.1 Microtrip Method.....	166
5.2.3.2 Markov Chain Method	167
5.2.4 Characteristic Parameters.....	168
5.2.5 Driving Cycle Validation.....	171
5.2.6 Comparing Construction Methods	172

5.2.7 Localised and Representative Driving Cycles.....	175
5.2.8 EV Specific Driving Cycles	176
5.2.9 Limitations of Driving Cycles.....	178
5.2.10 Conclusion and Gaps in the Literature of Drive Cycles.....	181
5.2.11 Summary of Previous Drive Cycles	182
5.3.1 Visum Model Overview	183
5.3.2 Speed Limit Scenarios.....	189
5.3.3 Visum Route Selection	191
5.3.4 Visum Route Search Function.....	195
5.3.5 Drive Cycle Construction.....	198
5.3.6 Preparing Drive Cycles for Energy Consumption Modelling	202
5.3.7 Drive Cycles with Road Type TPMs.....	204
5.3.8 Drive Cycles with Combined Links	207
5.4 Results.....	208
5.4.1 Driving Behaviour and EC.....	208
5.4.2 Route Choice and EC	211
5.4.3 Effects of Speed Limit Reduction on Energy Consumption and Travel Time	218
5.5 Discussion and Conclusion	227
Chapter 6. Discussion	230
6.1 Contributions.....	230
Chapter 7. Conclusion	237
References.....	240
Appendix A. Summary of Previous DCs	273
Appendix B. Empirical Method Applied to Route DCs.....	279
Appendix C. Summary of EC Models	281
Appendix D. Summary of Driving Behaviour Studies	283
Appendix E. Summary of Sensitivity Analysis Studies	284

List of Figures

Figure 2.1 Car Scanner Application Dashboard	11
Figure 2.2. OBDLink MX+ Device (OBD Solutions, 2025).....	12
Figure 2.3. Social Media Post for Participant Recruitment	15
Figure 2.4. Data Collection Timeline	17
Figure 3.1. 5-7km Trip Distance and Duration by Driver	36
Figure 3.2. 10-15 km Trip Distance and Duration by Driver	36
Figure 3.3. Elbow Method Applied to Clustering of Trips with 10-15km Distance.....	38
Figure 3.4 K-Means Clustering of Distance and Duration Applied to All Trips with 10-15km Distance	39
Figure 3.5. Share of Database by Driver	40
Figure 3.6. TII Weather Stations	47
Figure 3.7. GPS Coordinates Rounded to Three Decimal Places.....	50
Figure 3.8. GPS Coordinates Rounded to Four Decimal Places	51
Figure 3.9. Ireland Legal Towns and Cities.....	52
Figure 3.10. Coordinate Locations of All Trips	54
Figure 3.11. Regression Tree Example	58
Figure 3.12. Artificial Neural Network Example	60
Figure 3.13. TEC vs. Average Speed of All Trips.....	61
Figure 3.14. Segment MAE by Number of Hidden Layers in ANN	63
Figure 3.15. Trip MAE by Number of Hidden Layers in ANN.....	64
Figure 3.16. Loss and Validation Loss vs. Epoch of ANN	65
Figure 3.17. Feature Importance of Trip-Level Energy Consumption within the ANN Model.....	70
Figure 3.18. Trip Energy Consumption vs. Trip Distance of All Trips.....	75
Figure 3.19. Predicted vs. Actual Energy Consumption Using a Macroscopic Model	76
Figure 3.20. Feature Importance of Trip-Level Average Battery Power Model.....	79
Figure 3.21. Average Moving Speed vs. Segment Energy Consumption of All Segments	81
Figure 3.22. Feature Importance of Segment-Level Energy Consumption Model	82
Figure 3.23. Feature Importance of Trip-Level EC Model with Segment Predictions.....	85
Figure 3.24. Predicted vs. Actual Energy Consumption for Final Model	87
Figure 3.25. Variation of ECR along a Trip.....	88
Figure 4.1. Clustering Method Results on Real Trip Segments	102
Figure 4.2. RPA Passive Limit Method Applied to Real Trips	107
Figure 4.3. VA ^{95th} Aggressive Limit Method Applied to Real Trips.....	107
Figure 4.4. ECR by Aggression of Real Trips: Limit-Based Method.....	110
Figure 4.5. Segment Average Battery Power by Empirical Method Group	112
Figure 4.6. ABP Sensitivity to Driving Behaviour Variables by Speed Interval	131
Figure 4.7. ABP Sensitivity to Weather and Slope Variables by Speed Interval	131
Figure 4.8. ABP Sensitivity to Drive Mode Variables by Speed Interval	132
Figure 4.9. ABP Sensitivity to Road Type Variables by Speed Interval	132
Figure 4.10. ABP Sensitivity to Vehicle and Driver Characteristics by Speed Interval	133
Figure 4.11. Linear Trendline between EC and Acceleration Time in the 20-30 km/h Interval	135
Figure 4.12. Linear Trendline between EC and Deceleration Time in the 20-30 km/h Interval	135
Figure 4.13. Sensitivity of ABP to Acceleration, Deceleration, and Gradient by Speed Interval	136

Figure 4.14. Modelled Average Battery Power by Speed Interval	137
Figure 4.15. Modelled Average ECR by Speed Interval.....	138
Figure 4.16. Actual Average Battery Power by Speed Interval	139
Figure 4.17. Actual Average ECR by Speed Interval	139
Figure 4.18. Sample Driving Segment Example	141
Figure 4.19. EC vs. Average Speed for Urban Sample Segments	143
Figure 4.20. EC by Acceleration and Deceleration Rate for the Urban Sample Segments Starting at 40 km/h.....	144
Figure 4.21. EC vs. Average Speed for Motorway Sample Segments.....	146
Figure 4.22. EC by Acceleration and Deceleration Rate for Motorway Sample Segments Starting at 100 km/h.....	147
Figure 4.23. Application of Limit-Based Method to Sample Segments.....	149
Figure 4.24. Application of Clustering Method to Sample Segments	149
Figure 5.1. NTpM Roadway Network.....	184
Figure 5.2. Application of BPR Function (AECOM, 2019B)	187
Figure 5.3. Modelled Links of the NTpM in Dublin Area.....	192
Figure 5.4. Urban Routes.....	193
Figure 5.5. Interurban Routes	194
Figure 5.6. Sample Road Network	196
Figure 5.7. A* Algorithm Example Steps	197
Figure 5.8. Comparison of NTpM and OpenStreetMaps Roads	203
Figure 5.9. Speed Count Histogram on All Roads.....	204
Figure 5.10. Speed Count Histogram by Road Type	206
Figure 5.11. First Links of Least Distance Dub-Cork Route.....	212
Figure 5.12. First Links of Fastest Time Dub-Cork Route	212
Figure 5.13. EC by Route Choice on Urban Routes	213
Figure 5.14. EC by Route Choice on Interurban Routes	213
Figure 5.15. Travel Time by Route Choice on Urban Routes	215
Figure 5.16. Travel Time by Route Choice on Interurban Routes.....	216
Figure 5.17. Distance Travelled by Route Choice on Urban Routes.....	217
Figure 5.18. Distance Travelled by Route Choice on Urban Routes.....	217
Figure 5.19. Dub-Cork Base Case Drive Cycle.....	219
Figure 5.20. Dub-Cork Scenario C Drive Cycle.....	219
Figure 5.21. Real-World Drive Cycle Example 1	220
Figure 5.22. Real-World Drive Cycle Example 2.....	221
Figure 5.23. Average Link Speed of Dub-Cork Route: Base Case & Scenario C.....	222
Figure 5.24. Journey Time by Speed Limit Scenario on Urban Routes	223
Figure 5.25. Journey Time by Speed Limit Scenario on Interurban Routes.....	223
Figure 5.26. Energy Consumption by Speed Limit Scenario on Urban Routes	224
Figure 5.27. Energy Consumption by Speed Limit Scenario on Interurban Routes	225
Figure 5.28. Energy Consumption for Dub-Cork Route – Base Case and Scenario C	226

List of Tables

Table 2.1. Comparison of Smartphone Applications for Data Collection	9
Table 2.2. Comparison of Veepeak and OBDLink Devices.....	12
Table 3.1. Descriptive Statistics of Speed-Based Variables.....	43
Table 3.2. Descriptive Statistics of Acceleration-Based Variables	43
Table 3.3. Descriptive Statistics of Drive Mode Variables and Number of Stops	43
Table 3.4. Descriptive Statistics of Gradient and Weather Variables.....	48
Table 3.5. Descriptive Statistics of Road Environment	53
Table 3.6. Driver and Vehicle Information	56
Table 3.7. Performance of ML Models of TEC	69
Table 3.8. Distance, Duration, and Energy Consumption of Trips	70
Table 3.9. Correlation between Parameters and Trip-Level Energy Consumption	72
Table 3.10. MLR Results for TEC	73
Table 3.11. Error Comparison between Short and Long Trips	76
Table 3.12. Alternative Models for Trip-Level Energy Consumption	78
Table 3.13. Performance of ML Models of SEC	80
Table 3.14. Correlation between Parameters and Segment-Level EC	83
Table 3.15. Comparison of Final Model to Previous Models	86
Table 3.16. Error Comparison between Short and Long Trips of Final Model	86
Table 4.1. EC and ECR of Real Trip Segments Averaged by Cluster	103
Table 4.2. ANOVA Test Results on Clustering Method Applied to Real Trip Segments ..	104
Table 4.3. EC and ECR of Real Trips by Limit-Based Group.....	111
Table 4.4. ANOVA Test Results on Limit-Based Method Applied to Real Trips	111
Table 4.5. Average Statistics of Parameter by Group	113
Table 4.6. Distribution Params	125
Table 4.7. Parameter Distributions for Sobol' Sampling	125
Table 4.8. Sobol' Sensitivity Indices.....	127
Table 4.9. Average Statistics of Modelled Parameter by ABP Group.....	129
Table 4.10. Parameter Values for Sample Segments.....	142
Table 4.11. Average Statistics of Parameters by EC Group on Urban Segments	145
Table 4.12. Average Statistics of Parameters by EC Group on Motorway Segments	148
Table 4.13. ANOVA Test Results on Limit-Based and Clustering Methods Applied to Sample Segments	150
Table 5.1. Base Case and Reduced Speed Limits	190
Table 5.2. Urban Route Origins and Destinations	193
Table 5.3. Interurban Route Origins and Destinations.....	195
Table 5.4. Transition Probability Matrix Snippet	200
Table 5.5. Speed Counts by Road Type	205
Table 5.6. Energy Consumption with Standard and Combined Links.....	208
Table 5.7. Average Statistics by EC Group: Dub1 Route	209
Table 5.8. Average Statistics by EC Group: Dub-Cork Route.....	210
Table 5.9. ANOVA Test Results for EC by Route Choice	214
Table 5.10. ANOVA Test Results for EC by Speed Limit Scenario	226

Glossary of Terms

AADT	Annual Average Daily Traffic
ABP	Average Battery Power
ANN	Artificial Neural Network
BPR	Bureau of Public Roads
CAN	Controller Area Network
CP	Characteristic Parameter
DC	Drive Cycle or Driving Cycle
DEM	Digital Elevation Model
EC	Energy Consumption
ECR	Energy Consumption Rate
EV	Electric Vehicle
GIS	Geographic Information System
GPS	Global Positioning System
GSA	Global Sensitivity Analysis
HVAC	Heating, Ventilation, and Air Conditioning
ICEV	Internal Combustion Engine Vehicle
LHS	Latin Hypercube Sampling
LightGBM	Light Gradient Boosting Machine
MAD	Mean Absolute Deviation
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MC	Markov Chain
ML	Machine Learning
MLR	Multiple Linear Regression
MT	Microtrip
NEDC	New European Driving Cycle
NTA	National Transport Authority (Ireland)
NTpM	National Transport Model
OATSA	One-at-a-time Sensitivity Analysis
OBD	On-board Diagnostic
O-D	Origin-destination
OSM	OpenStreetMap
PCA	Principal Component Analysis
PID	Parameter Identifier
POWSCAR	Place of Work, School, or College Census of Anonymous Records
RDE	Real Driving Emissions
RELU	Rectified Linear Unit
RMSE	Root Mean Square Error
RPA	Relative Positive Acceleration
S1	Sobol' First-order index
SA	Sensitivity Analysis
SAFD	Speed-Acceleration Frequency Distribution
SEC	Segment-level Energy Consumption
SOC	State of Charge
SSE	Sum of Squared Errors
ST	Sobol' Total-effect index
Std Dev	Standard Deviation
SVM	Support Vector Machine

TEC	Trip-level Energy Consumption
TII	Transport Infrastructure Ireland
TPM	Transition Probability Matrix
WLTP	Worldwide Harmonised Light Vehicles Test Procedure
XGBoost	Extreme Gradient Boosting

Units of Measurement

Abbreviation	Full Name	Measurement Of
kWh	Kilowatt-hour	Energy
kW	Kilowatt	Power
MJ	Millijoule	Energy
kWh/km	Kilowatt-hour per kilometre	Energy Consumption Rate

Chapter 1. Introduction

EVs have been an emerging technology threatening to replace traditional ICEVs for over a decade. Yet EV adoption and overall market share still lag heavily behind ICEVs. A main attraction of EVs is that they provide consumers a choice to create less air pollution in their local environment while driving (Pamidimukkala et al., 2024). EVs also provide an enticing path to lower global greenhouse emissions to slow the pace of global warming (Braband et al., 2022; Chen and Yang, 2022). It should be noted that EVs are not a panacea to worries about air quality and emissions even though EVs remove all tailpipe emissions. Because EVs are generally heavier than comparable ICEVs, EVs have been found to emit a similar amount of particulate matter to ICEVs (Timmers and Achten, 2016). Particulate matter is generally small particles that are emitted through friction between the road, tyres, and brakes, and particulate matter emissions are a significant source of air quality reductions, which the European Union sets limits on due to the health and safety risks they pose (Patel and Aggarwal, 2022; Timmers and Achten, 2016).

As a result, the total environmental benefit of adopting EVs is directly tied to the energy sources used in the local region (Pamidimukkala et al., 2024). For an EV to maximise its environmental advantages over ICEVs, the energy used to fuel the EV must come from renewable energy sources with low carbon impact (Mekky and Collins, 2024; Qi et al., 2018). With the notable environmental benefits over ICEVs, it seems clear that EVs are the future of transport technology. So, the question remains, why are EVs not the clear choice of consumers today?

Among the policies to combat climate change and decarbonise Ireland, EVs have been singled out as an important sector to target. Specifically, the Climate Action Plan of 2025 outlines a desire for 30% of private vehicles to be fully electric and 100% of new vehicle registrations to be fully electric vehicles by 2030 (Government of Ireland, 2025). This bold proposal still has a long way to go before being a reality. As of August 2025, just

17% of new vehicles registered in Ireland in 2025 were fully electric, although that number is an increase on the 14% of new vehicles registered being electric from 2024 (Central Statistics Office, 2025). However, it is still far from the goal of 100% of new vehicle registrations. The global figures are similar. One study estimates 20% of all new vehicle sales worldwide were EVs in 2024, although this is mostly led by high EV sales in China (IEA, 2025).

Several policies have been enacted across the globe to improve EV sales. Ireland developed a National EV Charging Infrastructure Strategy to improve the number and availability of EV charging stations across the republic. (Government of Ireland, 2025). Financial incentives and rebates have also been popular globally, with Ireland giving grants of up to €3,500 for pure electric vehicles and discontinuing grants for hybrid vehicles (Citizens Information, 2025). Recently, Ireland has also introduced unique “green flash” on license plates of electric vehicles to improve the social perception of EVs (Citizens Information, 2025). An empirical review of policies in the U.S. showed that policies targeted at reducing carbon dioxide emissions and investing in green energy have the potential to increase EV adoption more than financial incentives or increased charging availability (Mekky and Collins, 2024).

The slow pace of EV adoption has been attributed to several factors. One significant factor is the relative lack of EV charging stations compared to petrol and diesel stations (Pamidimukkala et al., 2024). Ramachandaramurthy et al. (2023) cites several reasons why EVs have yet to reach the expected level of market acceptance, such as social pressures in addition to a limited charging network. Another reason that has been cited as a source of EV dissuasion is a phenomenon termed range anxiety.

Range anxiety has been defined as a psychological stress on a driver who fears the range of an EV will not be sufficient to manage either an ongoing or future trip (Rauh et al., 2015). Range anxiety can affect drivers on several levels, including a cognitive level, where

drivers may perceive EVs as insufficient, a behavioural level, where drivers may switch driving behaviour to extend an EV's range, and a physiological level, where drivers may exhibit physical symptoms of stress such as increased heart rate (Rauh et al., 2015). With modern improvements made to extend EV range, range anxiety in its current state is generally believed to be due to psychological barriers rather than technical ones (Noel et al., 2019).

When new car buyers consider EVs, they make comparisons to ICEVs in various factors, such as cost, range, and charging times (Kawgan-Kagan and Daubitz, 2014). This explains the origins of range anxiety, where drivers switching from ICEVs with large fuel tanks and near universal fuel stations feared the limited range and lack of available chargers with EVs. Range anxiety has been found to be the leading cause for avoiding EVs among common explanations (Noel et al., 2019). Range anxiety has been shown to negatively affect the perceived and actual utility of EVs, with added charging infrastructure only minorly improving drivers' perceptions of EVs (Neubauer and Wood, 2014).

Many studies validate the need for further research on EC in EVs to the phenomenon of range anxiety (Baek et al., 2019; Hariharan et al., 2022; Zhao et al., 2020). Still, in order to justify range anxiety as a significant barrier to EV adoption and not just a convenient antagonist to justify the slow sales, it must be shown that drivers actually are reluctant to use EVs for fear of a limited range. Noel et al. (2019) attempted to understand logical explanations for range anxiety despite recent improvements in EV range and charger availability. Noel et al. (2019) establishes the idea that although range anxiety is a phenomenon with some technological and psychological reasonings behind it, it is also often employed as a rhetorical argument for consumers who wish to reject EVs in favour of keeping the status quo.

Despite this, range anxiety exists as a real psychological issue for drivers who fear an EV may prohibit long-range trips (Rauh et al., 2015; Szumska and Jurecki, 2021).

Another concern comes from the fact that remaining range estimations provided by a vehicle may not match real conditions, and drivers may be left with less available range than expected (Rhode et al., 2020; Thorgeirsson et al., 2021; Varga et al., 2019). This variation in remaining range may be one reason why experience driving EVs does not appear to alleviate range anxiety concerns, as even experienced EV drivers still fear being stranded without charging options (Bühler et al., 2014; Franke et al., 2012; Jensen et al., 2013). Range anxiety persists despite EVs having generally been found to meet the daily driving needs of most people for most days (Wang et al., 2015).

Range anxiety serves as the primary motivation for the research presented in this thesis. The variation in remaining range in real-world conditions compared to the range predicted by an EV or stated by a vehicle manufacturer is an important topic that has gained significant attention in the recent literature (De Cauwer et al., 2020; Lee et al., 2024; Szumska and Jurecki, 2021). The range of an EV is determined by how much energy is consumed, so EC modelling is an important step to combat range anxiety. To better understand the factors that influence EC in EVs, better models are needed.

While reviewing the literature on EC models, it was discovered that several different scales are used when modelling EC. Microscopic models predict EC second-by-second, macroscopic models predict the EC of an entire trip, and mesoscopic models predict the EC of part of a trip (Straub et al., 2022). While all these models are used in the state-of-the-art literature, very little research has been done to compare the utility of these models. The first major step of this thesis is to build an EC model and determine which factors most influence EC. Then, a comparison of driving behaviour classification methods is undertaken to investigate the effect of driving behaviour on EC. Finally, a novel process is developed to study EC on trips made on a road network model using DCs. This process is used to determine the effects of a lowered speed limit on EC.

The arching idea of this thesis builds on previous work such as De Cauwer et al. (2020) and Straub et al. (2020). The structure of this thesis, from EC modelling to DC construction, provides the framework for a future advanced green vehicle routing application, where drivers could select routes with the lowest EC and even see how EC could change depending on driving behaviour. Each chapter provides novel insight into the study of EC in EVs. Several research questions are investigated through the thesis:

1. How does changing the scale of an EC model affect the degree to which different parameters influence the model?
2. Does a mesoscopic or macroscopic model provide better EC predictions?
3. Can driving behaviour be classified in a way that leads to different categories showing consistent differences in EC across short and long trips?
4. How does a speed limit reduction policy affect EC in EVs? Is the effect significant enough to warrant consideration in future econometric models?
5. How significant is the effect of changing driving behaviour on EC in realistic trip conditions?
6. Does choosing a route that minimises distance lead to more or less EC compared to a route that minimises journey time?

The objectives of this thesis are to answer each of the posed research questions in order to fill gaps identified in the literature. Following this introduction, Chapter 2 discusses the process of data collection used for this thesis. The thesis is then divided into three major sections of work: EC modelling in Chapter 3, analysing the effect of driving behaviour on EC in Chapter 4, and modelling how reducing speed limits changes EC and journey times in Chapter 5. Within each chapter, a review of the literature is presented which establishes state-of-the-art knowledge on that chapter's content and specific gaps in the literature where more research is needed. Though each chapter seeks to answer distinct research questions and produces new knowledge for the field, the chapters build upon each other. The analysis

in Chapter 5 would not be possible without building the EC model in Chapter 3 and analysing driving behaviour in Chapter 4. The result is one major study on EC in EVs broken into distinct ideas in each chapter. The thesis is finalised with a conclusion presented in Chapter 6, followed by a references section and appendices. The findings give quantitative results on how changes in driving behaviour can affect journey time and EC.

Chapter 2. Data Collection

The data used in this thesis was collected as part of the POWERDRIVE project, funded by the Horizon Europe programme of the European Union (POWERDRIVE, 2025). Therefore, the data collection process was framed by the requirements of the POWERDRIVE project. The work done at Trinity College Dublin for the POWERDRIVE project included testing and calibrating a commercially available smartphone application that captures driving behaviour. This application would capture a vehicle's speed, battery information, and GPS location during trips. The data collection methodology is described in this chapter, including selection of the data collection application and device and recruiting of participants. All details describing data processing methods are presented in the subsequent chapters as relevant.

2.1 Selection of Driving Behaviour Application

The first step of the data collection process was to find a commercially available smartphone application capable of capturing the necessary data. An initial search was conducted into the DC literature to identify any applications mentioned. Smartphone applications were rarely mentioned directly in the methodology sections of the papers reviewed, even when a smartphone application was used. Out of the 38 studies reviewed that constructed at least one new DC, seen in Appendix A, only Romero et al. (2017) mentioned a specific application, Torque Pro, which was therefore included in the evaluation of applications (Torque, 2025). Due to the lack of information in the literature, another search was conducted through the internet.

Several characteristics were necessary for the selected application. First, the application must be able to capture a vehicle's speed and battery data during normal trips. Battery data here refers to the amount of energy expended by the vehicle, recorded by the vehicle as battery power in kW. Another parameter to capture is torque, which can aid powertrain optimisation in the POWERDRIVE project. Additionally, the application was

required to be widely compatible for use with different smartphone operating systems and vehicle makes and models. This would allow more potential participants to be included in the database. It was also preferred that the application was popular with positive reviews and was free and easy to use.

In the DC literature, speed data is generally collected through either GPS or on-board diagnostics (OBD). GPS data can be collected by a GPS-tracking device (Ashtari et al., 2014; Bishop et al., 2012) or a smartphone (Diaz Alvarez, 2014). OBD data collection requires a device to be plugged into a vehicle's OBD-II port (Ramai et al., 2022). The device in the OBD-II port is used to access parameter information from the vehicle's internal communications systems (Cura et al., 2021). Smith, II et al. (2024) showed that both GPS and OBD speed data produce similar resulting DCs, but GPS data is limited to speed and location data. GPS speed data also has more noise and requires filtering to lessen the effect of noise (Smith, II et al., 2024). OBD data, on the other hand, can capture the battery and torque parameters desired for the POWERDRIVE project (Ramai et al., 2022), so an application that connects to an OBD device was preferred. Modern EVs are equipped with many kinds of smart technologies. Some EVs collect an abundance of data, which has been used in studies as an equivalent to OBD (Kolarova et al., 2017). However, not every vehicle is able to store and send data, but every vehicle should be equipped with an OBD-II port. So, OBD data is again the preferred option.

Twelve potential smartphone applications were identified through the literature and internet searches. Research was conducted through the internet on how well each of the considered applications fit the needs of the project. These applications and the criteria used to compare are shown in Table 2.1 below.

Table 2.1. Comparison of Smartphone Applications for Data Collection

Application	Method	Vehicles	OS	Cost	Downloads	Citation
Piston	OBD	Unspecified	Android	Free	1M+	Stinieez, 2025
AutoPulse	OBD	Unspecified	Apple	Free	100+	SenSight Technologies, 2025
EVNotify	OBD	Unspecified	Android	Free	10K+	EVNotify, 2025
DRIVESUNG	GPS	Unspecified	Android	Free	Unknown	ZenSung, 2025
Car Drive	GPS	Unspecified	Apple	Free	Unknown	No longer active
HARMAN Spark	OBD	Unspecified	Both	Free	10K+	AT&T, 2025
B On Track	OBD	Unspecified	Both	Free	1K+	Apple, 2025
Car Scanner	Both	Unspecified	Both	Free	10M+	Car Scanner, 2025
EV Watchdog	OBD	Hyundia, Kia	Android	Free	1K+	EV Watchdog, 2025
Leaf Spy Pro	OBD	Nissan Leaf	Both	\$15	10K+	Google Play, 2025
BimmerLink	OBD	BMW, Mini	Both	\$35	100K+	BimmerLink, 2025
Torque Pro	OBD	Unspecified	Android	\$5	1M+	Torque, 2025

Based on the criteria shown in Table 2.1, there were few applications that satisfied all needs of the research team. Applications like DRIVESUNG and Car Drive were not selected because they only provided GPS data. Additionally, EV Watchdog, Leaf Spy Pro, and Bimmerlink were not selected because they are only compatible with a few EV models. Initially, EVNotify (EVNotify, 2025) was chosen for study despite its unavailability on Apple devices since it was specifically designed to capture the data of EVs. However, EVNotify was unable to connect to the test vehicle during the testing stage. The next top choices were three similar applications: Piston, Car Scanner, and Torque Pro. All three were popular, well-reviewed applications that provide vehicle diagnostics and live data to users that can be recorded. Car Scanner was preferred due to its availability on both Apple and Android smartphones and its popularity. So, Car Scanner (Car Scanner, 2025) was selected for further testing as the next best choice due to its compatibility with both major phone operating systems and many different vehicle types.

The Car Scanner application, like many commercial driving behaviour applications, relies on reverse engineering communication codes determined by the vehicle manufacturer into usable data (Sim and Sitohang, 2014). OBD-II ports are standard in all modern vehicles

for the purposes of diagnostic testing and data collection (Abukhalil et al., 2020). However, the parameters available to extract from a vehicle's OBD-II port differs between vehicle manufacturers, as does the numerical ID used to code each parameter (Sim and Sitohang, 2014). This means that the Car Scanner application relies on users to decode parameters and build different profiles for different vehicles (Car Scanner, 2025; Sim and Sitohand, 2014). Users of the application must select the profile that matches their vehicle, and only parameters available in that vehicle are able to be collected (Car Scanner, 2025).

Although the parameters available differ between vehicles, many parameters and parameter IDs are standard across most vehicles (Abukhalil et al., 2020). Vehicle speed, battery power, and battery state of charge were available in most vehicle profiles. While vehicle speed was able to be recorded in every vehicle during data collection, no other parameter was universally available through the OBD-II port. For instance, battery power was missing from the Nissan Leaf profile in Car Scanner. Torque was only available in a few vehicles, such as the Tesla Model Y. GPS coordinates and elevation were also recorded directly in the Car Scanner application for every vehicle using the smartphone's GPS signal. Participants were asked to track several parameters when available for the purposes of POWERDRIVE, but only speed, GPS, and battery power are relevant to this thesis. An example of the live dashboard used to view and capture data in the Car Scanner application can be seen in Figure 2.1.

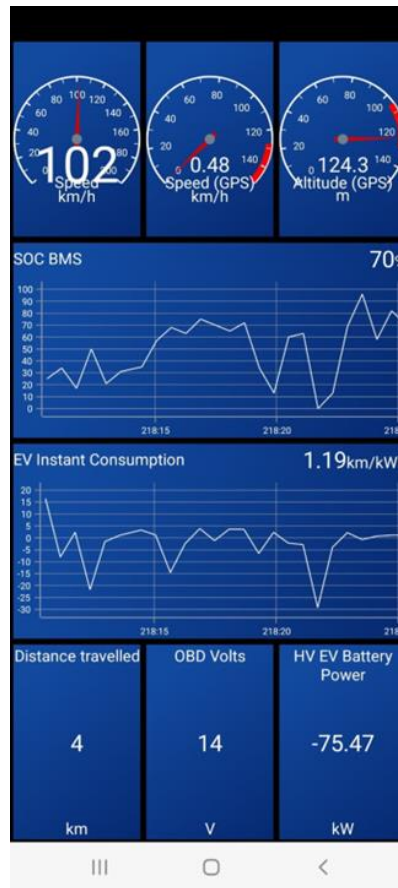


Figure 2.1 Car Scanner Application Dashboard

2.2 Selection of OBD Device

In addition to a smartphone application, a device to plug into the vehicles' OBD-II port was also required. The research team already had a device in their possession, the Veepeak OBDCheck BLE+ Plus (Veepeak Inc., 2025), so that device was used during initial driving tests in combination with the Car Scanner application. During these tests, it was found that the recorded data had frequent gaps. The literature on DCs establish a desired sampling rate, the frequency of recorded measurements, of 1 Hz (Corti et al., 2012; Zhao et al., 2020). Although the data collection was exceeding this standard on average, many gaps existed in the recorded trips where data was recorded at a lower sampling rate. The research team decided to test a higher quality device, the OBDLink MX+ (OBD Solutions, 2025), to reduce the number of gaps in the data. After several test trips with Car Scanner used in combination with the Veepeak and OBDLink devices, a comparison of the sampling rate and data gaps with each device was made, shown in Table 2.2.

Table 2.2. Comparison of Veepeak and OBDLink Devices

	OBDLink MX+	Veepeak
Average Sampling Rate (Hz)	1.98	1.09
Measurements with Gaps >1s (%)	1.24	36.70

Table 2.2 shows that both devices on average exceed the minimum sampling rate of 1 Hz (Corti et al., 2012; Zhao et al., 2020). The OBDLink MX+ has almost double the sampling rate of Veepeak, approaching almost two data readings per second on average. More importantly, the OBDLink MX+ significantly reduces the number of readings with gaps larger than one second from 36.70% to just 1.24% of all readings. Based on this improvement, the OBDLink MX+ was chosen as the device for data collection. A picture of the OBD Link MX+ is seen in Figure 2.2 below.



Figure 2.2. OBDLink MX+ Device (OBD Solutions, 2025)

2.3 Data Collection Methodology

During the testing phase, it was discovered that the Car Scanner application only consistently recorded all parameters when the smartphone remained on and opened to Car Scanner during the entire trip. This was seen as a significant inconvenience to participants, so a data collection option with zero participant interaction during trips was tested. Such devices had been used in previous DC studies such as Salihu et al. (2023). A CANedge1 device (CSS Electronics, 2025), used in Salihu et al. (2023), was acquired for testing.

However, it was determined that the CANedge1 and other zero-interface devices were unsuitable for this data collection due to the differences in coding OBD-II data discussed in Section 2.1.

As a result, the OBDLink MX+ device and Car Scanner application were selected as the data collection methodology for the POWERDRIVE project and this thesis. Because this data collection method required active participation by the driver before and after each trip, a set of thorough and concise instructions was developed to be shared with participant. The instructions were listed in a document, and an accompanying video was made in case participants preferred that form. The instructions covered all steps of data collection, starting with plugging in and turning on the OBDLink MX+ device. The user was then required to turn on the Car Scanner application, connect to the device via Bluetooth, and start recording data on Car Scanner. The application must then be left on the dashboard shown in Figure 2.1 for the entire duration of the trip. At the end of a trip, the driver would stop recording and send the data file to the research team via email. It was emphasised through the instructions and direct communication to leave the application open while driving as an essential step of data collection. Participants were requested to share their first trip data as soon as possible to ensure all necessary data was being collected. Once data was approved, participants were asked to continue sending trip data files at their convenience, and the research team reached out periodically to any participants who were not sending data to reestablish communication.

Based on the testing phase, a protocol was developed to ensure trips sent by drivers were suitable for the database. As mentioned previously, it was found during testing that the data on several trips cut out when Car Scanner was not left open on the smartphone. Therefore, each file sent by participants were manually inspected to ensure the data was consistent along a trip. A vehicle's speed, battery power, and GPS coordinates were graphed versus time of the trip, so that the research team could see any periods with missing data. In addition, the percentage of seconds during a trip without a datapoint for each of these

parameters was calculated, and any trip missing data for more than 5% of a trip's seconds was not included in the database.

Many trips sent by participants, especially early in the data collection period, were found to have missing data for sections of the trip. These trips were either not included in the database or trimmed to exclude missing data. A trimmed section of a trip was included in the database if data was consistent between stops, and this trimmed section was treated as a normal trip. Other trips lacked GPS data for an unknown reason, and these were not included in the final database. If a trip file captured all parameters continuously throughout the trip, it was added to the database. Any trip file added to the database was recorded in a trip log, which logged the date, time, distance, duration, and participant number of that trip. Although the number of trips sent by participants and not included in the database was not logged, it is estimated that roughly 10% of files did not meet the standards of the data collection protocol.

2.4 Participant Recruitment

With the data collection methodology established, the next step was to search for participants. Before continuing with the study, an ethics application was submitted to the Research Ethics Committee at Trinity College Dublin. The study was approved by the Research Ethics Committee on 12 April 2023 under reference number 1965. The research team simultaneously went through an additional process with the Research Ethics Committee to obtain permission to email the staff of Trinity College Dublin to recruit participants from the school. Approval to email the staff list was granted to the research team on 14 March 2024.

EVs account for just 5% of vehicles on the road in Ireland, which made the recruitment process difficult (Mac Uidhir and Rogan, 2024). Additionally, participants were not compensated financially for participating, so the search was reliant on goodwill and aid from the public. An initial search was made by the POWERDRIVE research team through

social media (Twitter and LinkedIn) requesting EV drivers in Ireland. The posts contained the advertisement seen in Figure 2.3 and included a link for interested drivers to share their emails for further contact.



Figure 2.3. Social Media Post for Participant Recruitment

Any drivers who registered interest were sent a participant information leaflet with information on the study, and they were required to fill out an informed consent form before participating. This process is mandated by the Research Ethics Committee. The participants were then sent or given a device along with a document and video containing instructions for data collection. Forty-five EV drivers responded to the initial form, of which twelve provided data for this thesis. The remainder of the interested EV drivers did not fill out the required informed consent form and did not respond to follow-up communications. The social media posts were made in July 2023, and the first data was received in August 2023. Participants were asked to record data for an initial period of two weeks, as the research team determined two weeks to be a compromise between obtaining as much trip data as possible without losing participant enthusiasm. After this, participants were requested to continue recording trips if desired or return the device. Some participants recorded trips

several months after the initial two-week period because the device was not required for additional participants due to the number of devices available to the research team.

A second phase of participant recruitment began in April 2024. For this phase, participants were found through email communication to staff at Trinity College Dublin. Following approval by the Research Ethics Committee, an email was sent to all staff members with the participant interest form. Again, an information leaflet and consent form were sent to all interested EV drivers. This phase received an even higher response, with 95 responses to the interest form leading to 24 additional participants. Again, the majority of participants did not return the required informed consent form despite multiple contacts by the research team. Data collection continued until the end of December 2024. However, only trips recorded by the end of August 2024 are included in this thesis.

An additional thirteen participants were sent devices but were not included in the database. Five of these participants lost interest or were unable to set up the data collection device and application. For six participants, the Car Scanner app did not contain the profile of their EV to enable data collection. Three of these vehicles were made by Tesla, along with one each by Renault, Mercedes, and BMW. One participant's vehicle, the MG 4, was unable to provide data consistently even when following the correct protocol. The last participant's smartphone would not record GPS data. A timeline of the year-long data collection process can be seen in Figure 2.4 below.



Figure 2.4. Data Collection Timeline

2.5 Data Collection Process and Data Preprocessing

The data collection process required constant communication with participants to troubleshoot and request trip data. Because the data collection process involved several steps from participants, there were several chances for error on each trip. This led to most participants sending several trips unsuitable for the database before sending trip files with the necessary data. While many issues presented themselves over the data collection period, the most common reoccurring issue was data not being recorded for part or all of a trip because the Car Scanner application was not kept open. For this reason, it is recommended that future studies attempt further testing with zero-interface devices, when suitable options become available commercially. Also, a standardisation of OBD-II data and parameter ID codes across vehicle manufacturers could aid future studies.

Once trips were added to the database, they went through one step of preprocessing. In order to fit the standard sampling rate of 1 Hz used in DCs (Corti et al., 2012; Zhao et al., 2020), trip data was standardised to a 1 Hz data frame. This means each second of a trip was given exactly one datapoint. Because the sampling rate was higher than 1 Hz, this generally meant reducing the amount of datapoints for each trip. In any second of data where more than one datapoint was measured, all measured datapoints were averaged into one resulting datapoint. In the small number of cases where a second had no datapoint, linear interpolation from the nearest measurements was used to estimate the value.

Any EV driver included in the database was also sent a survey to gather basic information on the driver and vehicle. In this survey, drivers provided the year, make, model, and battery size of their EV. Participants also shared their age, gender, and years of experience driving EVs. One participant did not return this survey, so their data is not included in this thesis. With that, the database was finalised for use in this thesis. Some additional data processing measures are described in the following chapters in the appropriate sections. In total, 2,373 trips, captured from August 2023 to August 2024 were included in the database for this thesis. This includes over 1,046 hours of driving data covering 62,770 kilometres.

Chapter 3. Energy Consumption Model

3.1 Introduction

EVs are considered to be an important step toward an emission-free future. For the past decade, governments and vehicle manufacturers have been pushing to increase the popularity and usage of EVs through monetary incentives and higher quality vehicles. However, there still exists a reluctance among ICEV owners that prevents the widespread adoption of EVs. A leading factor of this reluctance has been defined in the literature through the term range anxiety. Range anxiety describes the fear of an EV driver or potential EV driver that, without the quick and widespread fuelling stations and extensive range which are available to ICEVs, the driver will be stranded or forced to wait during a lengthy recharge of the battery. This is especially present on trips which near or exceed a vehicle's rated range, for which EVs have yet to match the distances common to ICEVs.

A vehicle's rated range is a parameter provided by a manufacturer as a key selling point for new EVs. However, the rated range tends to be idealistic and not representative of actual range due to variation in EC in real world driving scenarios. A legal disclaimer on the Kia website states that, despite EVs being labelled as having specific ranges, "individual driving style and other factors, such as speed, outside temperature, topography, and the use of electricity-consuming devices ... have an influence on real life range and can possibly reduce it" (Kia, 2024). Similarly, Hyundai (2024) warns "driving range may vary depending on road conditions, your driving style and the temperature" as well as the tyres. Discovering the impacts of these listed and other conditions present in real-world driving on EC and range has been a major topic of EV research in the literature, and it also provides the motivation for this research.

Specifically, the goal of this chapter is to create an EC model at the trip level. Because this model is intended for use as part of a larger thesis, the model must be made usable for the purposes of the later chapters. The later chapters include traffic assignment,

macrosimulation, and the building of synthetic drive cycles (DCs) to represent the actual speed and acceleration patterns of a trip. For this reason, the model is restricted to parameters that can be obtained from the combination of a macrosimulation trip with a known route, distance, and average speed as well as any speed- and acceleration-based parameters that can be obtained from a DC.

There exists a wide variety of methods to model EC found in the literature. One major difference between methods is the scale of the model. EC can be modelled at the microscopic level, which is generally accomplished using physical equations to estimate instantaneous EC second by second (Fiori et al., 2016; Mediouni et al., 2022; Wu et al., 2015; Appendix C). Alternatively, models can be made at the macroscopic level, which generally predict EC based on empirical data aggregated at the trip level (Ahmed et al., 2022; Ji et al., 2022; Jiang et al., 2023). Mesoscopic models have been introduced, where EC is modelled for smaller sections of a trip so that the data is aggregated over less time (Pan et al., 2023; Qi et al., 2018; Straub et al., 2022). Despite these different scales, little to no comparisons on how the scale of the model affects the model output have been made in the literature. Following a review of the literature presented here, this chapter aims to answer a few different research questions to fill gaps in current knowledge:

1. Do mesoscopic or macroscopic models give equivalent predictions of EC? If not, which scale leads to more accurate predictions.
2. Does the magnitude of parameters influencing EC change depending on the scale of the model used to predict EC?
3. Which ML model provides the most accurate EC predictions?

3.2. Literature Review

3.2.1 General Information - Energy Consumption of Electric Vehicles

EVs provide a more sustainable form of transport than traditional internal combustion engine vehicles (ICEVs). Alanzi (2023) summarised the current state of EVs

including the benefits and challenges of the emerging technology. EVs are fueled by the power grid as opposed to gasoline or diesel in ICEVs. Similar to fuel economy, minimising EC of EVs provides benefits by extending the driving range and lessening environmental externalities. This literature provides an overview on state-of-the-art research surrounding factors influencing EC, categorising driving behaviour, regenerative braking, range anxiety, modelling EC, and predicting an EV's remaining range and state of charge (SOC). Information on the EC models included in this literature review is presented in Appendix C.

3.2.2 Factors Influencing Energy Consumption

There are four major categories of factors that influence the amount of energy consumed by an EV: driving behaviour, weather, road environment, and heating, ventilation, and air conditioning (HVAC). Because driving behaviour can be controlled and is an important part of this thesis, it will be considered in the subsequent sections. This section outlines the knowledge surrounding the impact of other factors on EC. It is important to note that increasing EC is synonymous with decreasing range and SOC.

Varga et al. (2019) grouped the factors influencing EV range into vehicle design, driver, and environmental parameters, including both constant and variable parameters. Constant parameters are those such as vehicle weight, while variable parameters are those like SOC and environment. Sivak and Schoettle (2012) categorised driver choices that influence fuel consumption into the following types: strategic (vehicle type), tactical (route choice), and operational decisions (behaviour such as speed, HVAC, and idling).

EV drivers experience seasonal differences in the EC and range of their vehicle. Besides the changes in battery mechanics due to ambient temperature, additional impacts on battery usage are likely to be imposed by auxiliary systems in the car such as HVAC (Kambly and Bradley, 2014). Both with and without the use of HVAC, EC is higher at temperatures less than 20°C due to increased energy loss and higher auxiliary power consumption. (Liu et al., 2018). Lee et al. (2024) found 20-30°C to be the most energy efficient temperature range.

Yi et al. (2018) estimated differences of up to 20% in simulated EC and 60% in charging demand due to seasonal temperature differences in New York City.

A case study in Lille, France measured an EC increase of 33% in winter and only up to 15% in the summer, mostly due to HVAC auxiliary power consumption (Desreveaux et al., 2020). Yang et al. (2023) corroborated the findings, adding that summer and autumn have similar EC characteristics in Tianjin, China. Shibata and Nakagawa (2015) found that the air conditioning can make up 25-50% of a vehicle's EC in a hot climate. Hariharan et al. (2022) showed that the use of battery for air conditioning leads to an increase of EC as well as higher variability in SOC during a trip. In terms of HVAC, the car's heating systems generally require more EC than the air conditioning (Liu et al., 2018; Lee et al., 2024). Extreme cold temperatures are known to have extreme limiting effects on range. For example, EV range is less than 60% of the rated range at -15°C mainly due to heating (Abediasl et al., 2023).

Modi and Bhattacharya (2022) include wind speed in EC prediction but assume wind to be a headwind to the vehicle. Although there are major difficulties in aligning wind speed to direction of travel and the model still showed some correlation between wind speed and EC, the assumption of a headwind can negatively impact the accuracy of EC prediction for trips. Vepsäläinen et al. (2019) included direction in their wind measurements but noted how quickly wind direction changed. The study included uncertainty in the input parameters such as wind speed and direction in their EC model (Vepsäläinen et al., 2019).

Road environment such as road type and traffic conditions can have a significant effect on EC, making route choice an important part of EV driving. Yang et al. (2023) utilised principal component analysis (PCA) and K-means clustering to divide trips into urban, suburban, and high-speed microtrips (MTs) and found EC to be highest for urban driving and lowest for high-speed. Xie et al. (2020) found that heavy traffic conditions can negatively affect EC, but this could be minorly aided by regenerative braking. Felipe et al.

(2015) found that route choice had a larger effect on EC than driving behaviour, but behaviour still had a significant effect. Hariharan et al. (2022) modelled trip-long energy consumption and theorised its use in map applications to choose the lowest-energy route. Lee and Wu (2015) had previously proposed a similar idea using big data. Road grade (also called gradient) is known to have a non-negligible effect on SOC. Liu et al. (2018) found that road grade only significantly affects EC at slopes higher than 3% uphill or downhill.

Many studies have already been performed to measure the impact of several factors on EC, and this paragraph outlines the findings. Skuza and Jurecki (2022) summarized the extent to which a variety of parameters can affect energy consumption. Neubauer and Wood (2014) observed that high ambient temperatures add more wear to the battery and are likely to increase SOC depletion through HVAC usage. The same study found that aggressive driving can reduce EV efficiency. Ahmed et al. (2022) found that battery capacity, top speed, and kerb weight have high positive correlations with driving range. On the other side, the time of acceleration from 0 to 100km had a strong negative correlation with range. Through a sensitivity analysis, it was found that auxiliary power has a major effect on EC especially speeds lower than 80km/h when the EC to propel the vehicle is less. The same analysis revealed that vehicle weight only has an important effect during trips with varied gradients, and regenerative braking impacts EC most on downhill trips (Asamer et al., 2016). Generally, weight is considered to be an important factor in energy consumption in both passenger EVs and electric buses, including load and number of passengers (Liu et al., 2020; Jaworski et al., 2023; Settey et al., 2021).

Al-Wreikat et al. (2021) found ambient temperature to have the largest effect on EC. The same study found that EC is higher on average for trips with aggressive driving, longer trips, trips with longer idle times (due to auxiliary power demands), and uphill trips. Szumska and Jurecki (2021) confirmed ambient temperature and vehicle load to have large impacts on EC. On the other hand, Janpoom et al. (2023) found load to have a negligible

impact. Morlock et al. (2020) developed an EC model using current traffic conditions to estimate speeds and number of stops. Xie et al. (2020) added that frequency distribution of deceleration while braking can be an EC predictor. Abukhalil et al. (2020) used support vector machines (SVMs) to predict fuel consumption based on engine RPM, vehicle speed, and throttle position. Wei et al. (2023) experimented with using different variables to predict and reduce future energy consumption with mixed results. Loganathan et al. (2021) shows that EVs with higher range require more time to accelerate due to the weight of the extra battery. Madziel and Campisi (2023) present the correlation between various parameters and energy consumption for their database. This study also compares energy consumption by vehicle manufacturer and type. Galvin (2017) observes that higher speeds may increase energy consumption, but they do so for a shorter period of time. Therefore, it is important to look at the energy consumption of the entire trip. Dynamometer tests in this study suggest that high accelerations cause great losses of energy, challenging the idea that EVs are most efficient in urban areas if the number of acceleration periods surpasses a threshold. Mavlonov et al. (2023) found that battery size and efficiency map, which converts the speed and torque requirements to electrical energy, can affect SOC drain by 10-21%. Some studies estimate EC based on battery mechanics rather than driving behaviour. Zhao et al. (2020) shows the strong linear but not one-to-one relationship between battery power and used SOC over an entire trip.

3.2.3 Energy Consumption Modelling

Modelling and predicting the energy consumption of EVs is a major emerging topic. As discussed in the previous section, many studies have identified parameters that affect EC to various degrees. New models are being constantly developed to predict EC from these parameters. Romli and Iskandar (2023) identified three methods in the literature for SOC prediction: simulation, data-driven, and model-based. Data-driven models relate known input parameters to known EC outputs while simulation and models use knowledge of

physical properties and relationships to predict EC. Difficulties and complexities in modelling have led to a shift toward data-driven approaches in the literature (Das and Kumar, 2023). Model type can be inferred from the methods column of Appendix C.

EC models range from simple physics calculations to deep machine learning (ML) methods. Wu et al. (2015) showed that even a simple physics-based model can predict EC relatively well for an EV with sample data. Physics- or kinematics-based models calculate forces on the vehicle to estimate how much power the vehicle needs. Agrawal et al. (2023) present detailed mathematical models for predicting energy requirements of EVs based on dynamic forces. Miri et al. (2021) developed a detailed model of internal vehicle dynamics to predict EC. Vatanparvar et al. (2019) focused on predicting future speed and provided a number of equations that model SOC changes over time from both propelling the vehicle and HVAC systems. Alhanouti and Gauterin (2024) developed a model that combines accuracy and computational efficiency by back-calculating the mechanics of powertrain components from speed. Mediouni et al. (2022) created a model that relates mechanical power to electrical power when predicting energy consumption.

ML techniques are used in many modern studies, especially those that aim to predict EC as part of an energy management system (Adnane et al, 2023). A few examples are listed in this paragraph. Rhode et al. (2020) tested various ML algorithms to predict EC from speed and acceleration and found that kernel adaptive filters could predict energy well, with a normalised root mean squared error of around 90% at the cost of great computational power. Sasank et al. (2023) developed a MATLAB model to estimate range and SOC given a drive cycle and other input parameters. This was done to provide a standard method of EC modelling for various driving simulators in MATLAB (Sasank et al., 2023). Janpoom et al. (2023) showed that ML algorithms can have success in predicting EC, but the models had access to battery parameters such as current and voltage. All of the tested ML algorithms (support vector regression, kernel ridge regression, random forest, and extreme gradient

boosting) provided lower errors of EC, and extreme gradient boosting made its predictions with the fastest computational time (Janpoom et al., 2023). Madziel (2024) created a blueprint method of predicting energy consumption from microsimulation using ML techniques. Gradient boosting and neural networks predicted EC with less error than a random forest model (Madziel, 2024). These examples show different use cases for ML in energy consumption modelling, and the prevalence of ML in state-of-the-art studies demonstrates how important ML is to the field.

The success of regression models compared to other ML models such as artificial neural networks (ANN) and decision trees has been discussed in detail in the literature. De Cauwer et al. (2017) utilised a multiple linear regression model to estimate energy consumption based on kinematic parameters, gradient, and ambient temperature. Liu et al. (2018) highlighted the importance of interaction terms when estimating EC through linear regression. Sun et al. (2019) showed that machine learning estimations of range using gradient boosting can remove over half of the error of regression-based estimations. Santhosh et al. (2023) found that ML techniques like SVMs could outperform linear regression when predicting energy consumption. Conversely, Ekici et al. (2023) found that multi-variable linear regression outperformed many ML techniques when predicting average consumption. This could be because the big data that the study used showed very high correlations for parameters like average speed, average acceleration, and average slope to EC. Both linear and logarithmic regression models are able to accurately predict EC differences based on road gradient (Liu et al. 2017). Marzbnai et al. (2023) suggested a hybrid modelling approach for EC as linear models may not adequately consider non-linear relationships of parameters and ML or stochastic non-linear models may overfit the model to the data.

Model-based methods model the EV's battery to examine its physical properties during driving. Adding battery models can improve SOC estimation over conventional

methods where emphasis is not placed on physical imperfections that occur when drawing power from an EV (Baek et al., 2019). A review on ML techniques for EV battery modelling found that gradient boosting, SVMs, and ANNs were all able to accurately predict SOC to a high degree (Manoharan et al., 2022). For battery modelling, voltage, current, and temperature are the most common parameters for determining SOC (Manoharan et al., 2022). Sun and Fu (2024) developed a state-of-the-art model for EV batteries using decision trees to capture the performance of EV batteries under various ambient temperatures. The results of Ali et al. (2022) imply that battery models should be adaptive in order to more accurately capture real world performance and SOC loss. Zhou et al. (2023) provided recent literature on the difficulties and successes of modelling battery states. Many battery models only predict SOC in the 20-80% range due to the relatively stable characteristics of the battery in that range (Manoharan et al., 2022).

Di Martino et al. (2022) determined that data-driven models of EC can better consider driving styles and driving behaviour than vehicle model-driven estimations. Gurusamy et al. (2023) added that data-driven models generally are more accurate at the cost of computational effort. The latter also states that data-driven approaches do not consider auxiliary power consumption, but several data-driven models have accounted for the additional power consumption by auxiliary sources (De Cauwer et al., 2017; Yang et al., 2023). Jiang et al. (2023) used MLR to estimate auxiliary power consumption after noting that auxiliary power consumption showed a relationship with ambient temperature and time of day. The model of auxiliary power consumption contained high errors (RMSE of 1.89kW) due to the high variability of the target variable. Still, the model was unbiased and provided some insight to the whole EC model.

Prehofer and Mehmood (2020) demonstrated the potential of cloud computing live vehicle data, with the example case being EC prediction. The physics-based model shows

good levels of accuracy in a matter of seconds, but the model fails at high acceleration and deceleration peaks.

A noise filter, such as an unscented Kalman filter, can improve the accuracy of battery SOC estimators using least squares (Shebab El Din et al., 2018). Pan et al. (2023) utilised MTs as the duration for which to predict EC as a compromise between continuous prediction and prediction for the whole trip. Yang et al. (2023) suggested that MTs are too short for precise EC calculations but allow for analysis of driving behaviour. Straub et al. (2022) improved route-based SOC prediction by separating EC between propulsion and recuperation based on the flow of energy determined by analysis of the powertrain as opposed to the common method of separation by acceleration.

A review on EC prediction modelling (Marzbnai et al., 2023) provides five challenges to research on the topic: lack of broad, real-world data including limited geographic locations studied, a lack of sensitivity analysis for the input parameters, prioritisation of accuracy without regard for computation time, no justification for chosen prediction methods, and lack of concern for under- or over-fitting ML models. The model presented in this chapter has taken steps to address all of these issues.

3.2.4 Data Collection Methodologies

Data collection is an important step in every study on EC. Because EC is affected by many internal parameters, parameters other than speed are generally needed. Carmona et al. (2015) noted the importance of OBD parameters in addition to GPS and inertial data when describing driving behaviour. In this thesis, parameters from the vehicle's internal communication such as in the control area network (CAN) bus are referred to as OBD data as that is the port where such information is retrieved.

A combination of OBD and smartphone data is the state-of-the-art method for driving behaviour data collection, but these methods still lack environmental data (Fafoutellis et al., 2021). Diaz Alvarez et al. (2014) validated data from a smartphone's accelerometer through

OBD data, finding that jerk was a good predictor of EC. De Cauwer et al. (2017) collected GPS and CAN data from EVs in Belgium with a minimum frequency of 1Hz for both data types. The frequency was variable across different instrumented vehicles. Weather data was taken from a national meteorological institute. Hong et al. (2021) collected trip data using an OBD dongle connected via Bluetooth to a phone app. Cheap data loggers can be built by a research team (Ramai et al., 2022), but that method is not suitable for compatibility across many vehicle types. There is no standard for parameter IDs (PIDs) in EVs (Ramai et al., 2022). Therefore, most OBD data has been reversed engineered and is available through apps such as Car Scanner ELM OBD2 (Car Scanner, 2025) used in this thesis.

Ayman et al. (2022) outlined an algorithm for mapping GPS data points to actual roads in order to better incorporate weather, traffic, and elevation data in EC estimation. Schäfers et al. (2023) improved the accuracy of resistance modelling (e.g., slope and wind resistance) by combining an online map with empirical driving data. Road type (urban, rural, or motorway) can be approximated by maximum speed (Al-Wreikat et al., 2021).

3.2.5 Range Anxiety and Remaining Driving Range Prediction

Major differences between EV and ICEV drivers come from a particular phenomenon for EV drivers known as range anxiety (Pan et al., 2017, Settey et al., 2017). This phenomenon describes how EV drivers may choose their routes and driving styles based on a fear their EV will run out of battery and leave them stranded. Researchers have used the past decade to study the effects, learn how to ease range anxiety, and improve remaining range prediction.

Kim et al. (2017) determined that range should not have a major effect on EV adoption, as the daily mileage of most households do not approach maximum range of most EVs. Still, their regression models supported the idea that increasing EV range is more cost-effective than adding charging stations for increasing EV adoption. According to Lu et al. (2020), average daily usage for EV drivers is only 34% of an EV's range.

Range anxiety and reliability of travel time and charge are critical to the success of EVs. Of vehicle-related factors, performance and reliability are the most important for adoption of EVs. Pevec et al. (2019) used survey responses to determine the effects of various parameters on range anxiety based on how far they would be willing to drive before charging. One finding of this study is that participants were more likely to base their range anxiety on the vehicle's estimation of remaining range than on SOC, but drivers experienced range anxiety more at levels of SOC below 30%. (Palit et al., 2022). Through data analysis and communication with an EV driver, Wu et al. (2015) found that EV drivers choose routes based on travel time and EC. Additionally, the same driver may choose different routes when driving an ICEV compared to an EV. De Cauwer et al. (2020) showed that the most energy efficient routes tend to favour lower distances when not using much auxiliary power consumption, but they favour lower durations when auxiliary consumption is high. This finding makes sense as auxiliary consumption is active for the total duration of a trip, independent of whether the vehicle is moving.

Jung et al. (2015) compared driving behaviour when EV drivers were told an exact range remained versus when the displayed range was more ambiguous. They observed that with an ambiguous range, drivers were more aggressive with a higher SOC and more energy-saving with a lower SOC compared to those with an exact range. Varga et al. (2019) suggested that separating driving factors from auxiliary power use during range estimation could reduce range anxiety as these factors are independent of each other.

One emerging area of the research is real-time EC prediction to improve remaining range accuracy and therefore decrease range anxiety (Modi et al., 2020). Fukushima et al. (2018) noted the importance of modelling remaining range as it relates to finding suitable charging points before the end of the journey if necessary. Gurusamy et al. (2023) provided a thorough literature review of topics related to EV modelling and remaining range prediction. This study highlights the importance of realistic DCs and modelling the

uncertainty of internal and external parameters such as battery states, weather, and road environment. The importance can be seen through the fact that EV range can be 30% lower in real driving conditions compare to rated estimates, and which can drop by another 30% in cold conditions (Al-Wreikat et al., 2022).

Several models have been developed to predict remaining range, each with important insights. Foiadelli et al. (2018) developed a vehicle-specific plot of speed to remaining range that fit the data to an R-square value of 0.999. Li et al. (2023) provided a good template for integrating EV EC modelling with O-D route modelling to better understand the needs of charging infrastructure. O-D matrices have been used in the context of EVs mostly for the planning of charging infrastructure (Xiang et al., 2016). Mei et al. (2023) emphasised the need to incorporate conditions of future travel in addition to historical data when predicting remaining range if the data is not complete. Thorgeirsson et al. (2021) introduced probabilistic range predictions to their model with a factor of safety added in case of unforeseen problems. This study notes the factor of safety should be as small as possible to avoid lengthening travel time by excess charging during the trip. Qi et al. (2018) modelled link-level EC to allow for energy efficient route choice and network simulation of EC.

3.2.6 Model Target – Energy Consumption and Energy Consumption Rate

It was identified in the literature of EC modelling that there are different target variables. In particular, there were two main types of models discovered: EC models and energy consumption rate (ECR) models. EC models predict the amount of energy consumed, generally in kilowatts. On the other hand, ECR models predict the rate of energy consumption, generally in some form of kilowatts per distance travelled. This section identifies some research of both types of models as well as the model accuracies. The model target is one factor investigated in Appendix C.

Miri et al. (2021) combines several models, such as an electric motor model, auxiliary power model, and battery model to predict the energy consumption rate in Watt-

hours (Wh) per km. This study includes a simplified driver model which exists to accelerate or decelerate the model vehicle to match a desired speed. The model achieved errors between -15.6% to 5.9% on energy consumption rates between 131 Wh to 176 Wh.

De Cauwer et al. (2020) predicts energy in kWh based on distance, speed, acceleration, altitude, duration, and a temperature-dependent variable to represent auxiliary power usage. The MLR model achieved an R^2 of 0.89 based on one driver and vehicle. In this model, a term which consisted of squared speed multiplied by distance had the greatest effect on EC with a coefficient value of 3.78, but the impact of acceleration and deceleration were of only slightly less importance with coefficient values of 1.79 and 1.29, respectively. This means a change in the speed-distance term relative to the mean value will have double the effect on increasing EC compared to the same change in acceleration or deceleration. All other parameters had an impact on the model at least one order of magnitude lower. Still, energy-efficient routes were found to be similar to distance-efficient routes especially when auxiliary power consumption is low.

Pan et al. (2023) predicts EC in kWh at the MT level. The target value of EC, however, is a calculated value by finding the difference in SOC and multiplying the percent change by the rated energy capacity of the vehicle. Target values are the values of the parameter that ML models are trained to predict. An ANN model was developed based on speed, acceleration, distance, battery voltage and current, and temperature. The individual effects of these parameters on the model were not noted. The mean EC of the dataset was 0.425 kWh, and the model resulted in a MAE of 0.065 kWh and an RMSE of 0.122 kWh.

Modi and Bhattacharya (2022) estimate EC in millijoules in ten-second intervals based on speed, acceleration, gradient, auxiliary power, initial SOC, wind speed, and temperature. The model gives excellent predictions, with MAE and RMSE as low as 0.20 MJ and 0.27 MJ, respectively; however, it should be noted that the actual EC is obtained from simulation.

Galvin (2017) achieved an R^2 higher than 0.96 for 7 out of 8 tested EVs on a simple regression of speed and acceleration terms to predict power demand in W. This demonstrates the ability of even simple physical models to fit certain datasets well. Qi et al. (2018) developed an MLR and an ANN model which both gave an R^2 of about 0.80 to model EC rate (kWh/km) based on cumulative positive and negative kinetic energy terms which be simplistically explained as cumulative positive and negative accelerations. Zhang et al. (2020A) implemented XGBoost to achieve an RMSE of 0.159 kWh and an MAPE of 12.86 while predicting trip-level EC between 0.5-3.5 kWh using parameters for distance, duration, speed, acceleration, and temperature. These results show that many types of models have been developed to predict EC accurately, but it is unclear for which datasets certain models might perform better.

Yang et al. (2023) predicts ECR in kWh/100km after first obtaining kWh by multiplying the change in SOC by the vehicle's battery capacity, setting a minimum threshold of 10% in SOC to make sure the output variable is consistent. This study lacks definitive information on the makeup of characteristic parameters after using the PCA algorithm; however, some information on correlation between ECR and various parameters is included through a correlation matrix. It can be seen through the correlation matrix that average speed and auxiliary power usage (in this case the parameter is air conditioner state) have the highest Spearman rank correlation values with ECR at around 0.5. Other variables such as drive mode percentages and temperature are slightly less correlated with ECR, while average acceleration and deceleration show only slight correlations with ECR.

3.2.7 Model Scale

Another factor of any EC model is the scale of the model. As mentioned in Section 3.1, three levels of EC models have been identified by Basso (2021) and Straub (2022): microscopic, mesoscopic, and macroscopic. Microscopic models predict instantaneous energy consumption second-by-second usually with a physics-based model (Galvin et al.,

2017; Jiang et al., 2023; Xie et al., 2020). Macroscopic models predict the energy consumption of an entire trip (Ullah et al., 2022; Yang et al., 2023). Mesoscopic models provide a middle ground where fractions of a trip are modelled (Pan et al., 2023; Wei et al., 2023). It has been suggested that microscopic models are unable to capture driving behaviour and changes in road conditions by looking at only a moment of the trip (Li et al., 2021; Pan et al., 2023). Model scale is one factor detailed in Appendix C, where it can be seen that the majority of models are microscopic.

3.2.8 Feature Importance

While there have been many studies made to determine the features that influence EC, only a few quantitatively examine the effects of model parameters. Blades et al. (2024) tests various methods of feature selection: recursive feature elimination, sequential forward selection, and sequential backwards selection. Average road gradient and average speed were consistently found to be the best predictors for EC of electric buses, even outperforming total distance. The models reached an R^2 value of 0.85. Still, when out-of-bag variable importance was tested, trip distance was the most influential variable, all other variables were less than half as important. Maity and Sarkar (2024) implemented permutation feature importance to determine that gradient changes have a higher effect on ECR than acceleration and speed parameters. The features included by each study can be found in Appendix C.

Ji et al. (2022) categorises the importance of a variable in prediction as follows, with the threshold values being absolute values to account for both positive and negative correlations: a correlation between 0.1 and 0.3 with the dependent variable is a weak predictor, between 0.3 and 0.5 is a medium predictor, and higher than 0.5 is a strong predictor variable. According to this standard, trip distance and temperature were the only strong variables at predicting trip-level EC of buses among the variables used in the model, with distance being the best predictor. Vehicle mass, trip duration, and average speed were in the

medium range of correlation, and auxiliary power consumption (in this model, time with air conditioning turned on) had a low correlation with EC.

3.2.9 Summary and Gaps Found

EC modelling is an important subject for EVs and therefore many papers are being published each year on the subject. This literature has covered the wide range of EC models and related topics such as range anxiety. Various types of EC models were identified, most prominently physics-based models, data-driven models, and simulation models. These models can be applied at different scales including at the trip-level EC, second-by-second EC, and EC of part of a trip. Many features have been presented as important to EC, although each model uses different inputs. Methods to identify and classify different levels of aggressiveness in driving behaviour were shown. Finally, a use case for EC models was identified for estimating remaining range of a battery. Remaining range estimation can make a practical difference on the adoption of EVs by reducing the phenomenon of range anxiety.

Several areas for further research were identified through the literature review. First, a direct comparison of accuracy across model scales is lacking in the literature. In addition, the importance of features on EC has been widely represented in the literature, but quantitative analyses of feature importance and comparisons across EC and ECR models are scarce. Data-driven models also lack representation in the literature compared to other models. The methodology presented in this chapter attempts to add important information to fill these gaps in the literature, and the research outcomes to inform the literature are highlighted in the conclusion of this chapter.

3.3. Methodology

3.3.1 Data Processing: Streamlining the Database

Data collection was conducted over a full year from August 2023 to August 2024. As drivers were not all recruited at the same time, there are inconsistencies in the frequency and total number of recorded trips by drivers. A major concern of the data processing stage

was the potentially heavy contribution of just a few drivers within the database. This resulted from the fact that recruited drivers were allowed to send as much trip data as they wished. As such, some initial analysis of the data was required to determine if the excessive contributions of some drivers would cause overrepresentation in the modelling of one type of trip, vehicle, or driving style. Figures 3.1 and 3.2 were developed to observe the trip database at different distance intervals, 5-7km and 10-15km respectively. In Figures 3.1 and 3.2, each point represents an individual trip, and each colour represents an individual driver.

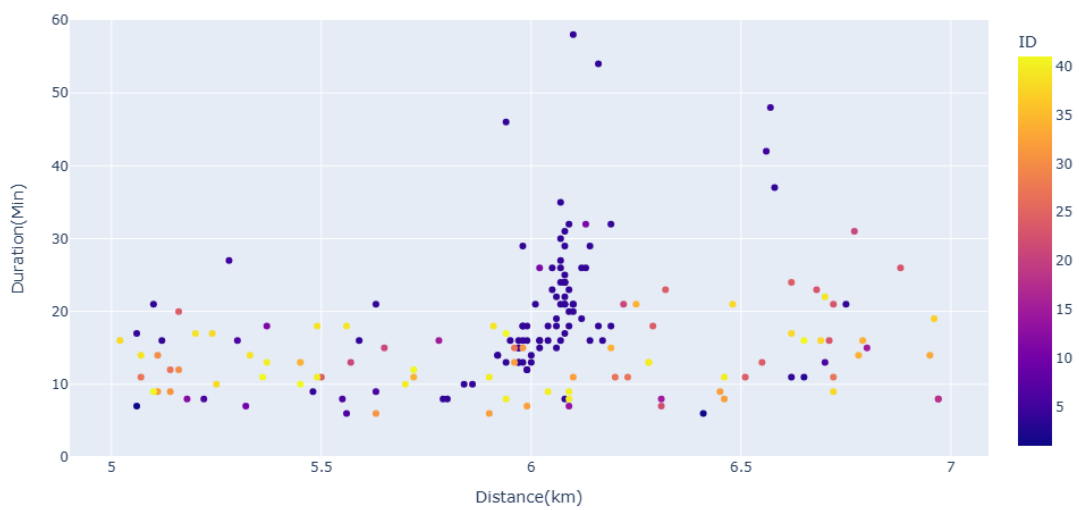


Figure 3.1. 5-7km Trip Distance and Duration by Driver

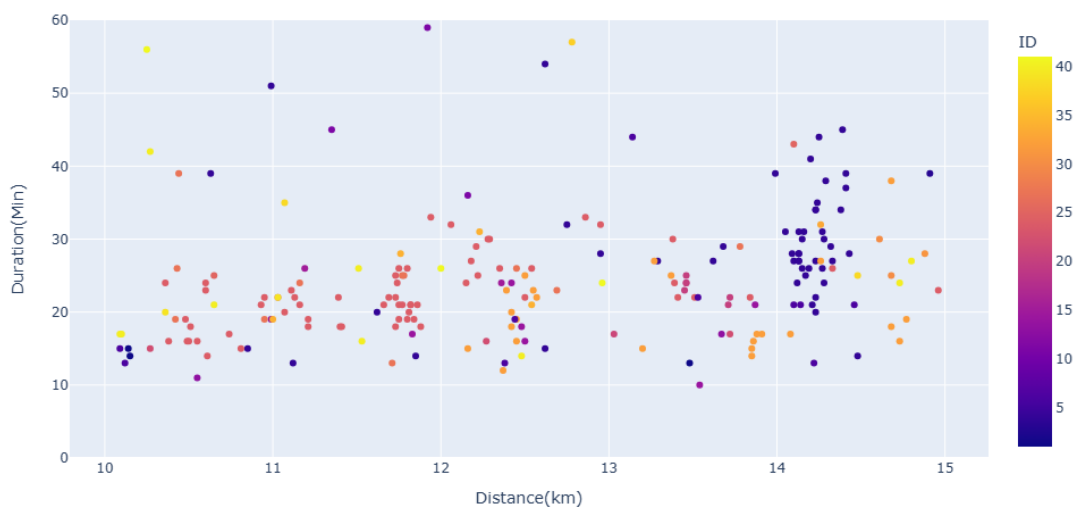


Figure 3.2. 10-15 km Trip Distance and Duration by Driver

Figures 3.1 and 3.2 show that there is a wide spread of trips across all distances; however, there exist a few distinct groups of points where one driver contributes many instances of the same or a similar trip to the database. For example, there are clear clusters of indigo around six and fourteen kilometres in addition to a cluster of tangerine around 12 kilometres. These clusters are most likely the most regular commutes of these drivers, which should be more highly represented in the database. The aforementioned indigo clusters show a large range of trip durations even on these commuting trips with similar distances. Unfortunately, only information about the participant's speed is known, but it can be theorised that these differences were due to different traffic conditions. On the other hand, the tangerine cluster around 12km experienced a much narrower range in trip duration, meaning the commute duration does not experience much volatility through different traffic conditions.

In order to better understand if these repeated trips stand out from the data of all drivers, a cluster analysis was conducted. K-means is a popular ML algorithm for grouping data into clusters. A more detailed explanation of the algorithm is given in Section 4.2, but for now it is important to know that the algorithm forms k number of clusters that minimises the distance error between points and the centroid of the cluster each point belongs to (Kodinariya and Makwana, 2013). In this instance, the error calculated is the sum of squared errors (SSE), which is a sum of the squared distance between each point and its cluster's centroid. As the error term is squared, the k-means algorithm is sensitive to outliers (Abdel-Wahed et al., 2023).

K-means clustering was implemented to cluster all trips by distance and duration in the 10-15km range. To determine the optimal number of clusters to suit this data, the elbow method was used. The elbow method plots how SSE decreases as the number of clusters increases, and an inflection point can be seen where the SSE does not decrease substantially by adding another cluster (Syakur et al., 2018). The k-means algorithm used here is from the

scikit-learn library in Python (Pedregosa et al., 2011). The results of the elbow method on this data are shown in Figure 3.3 below.

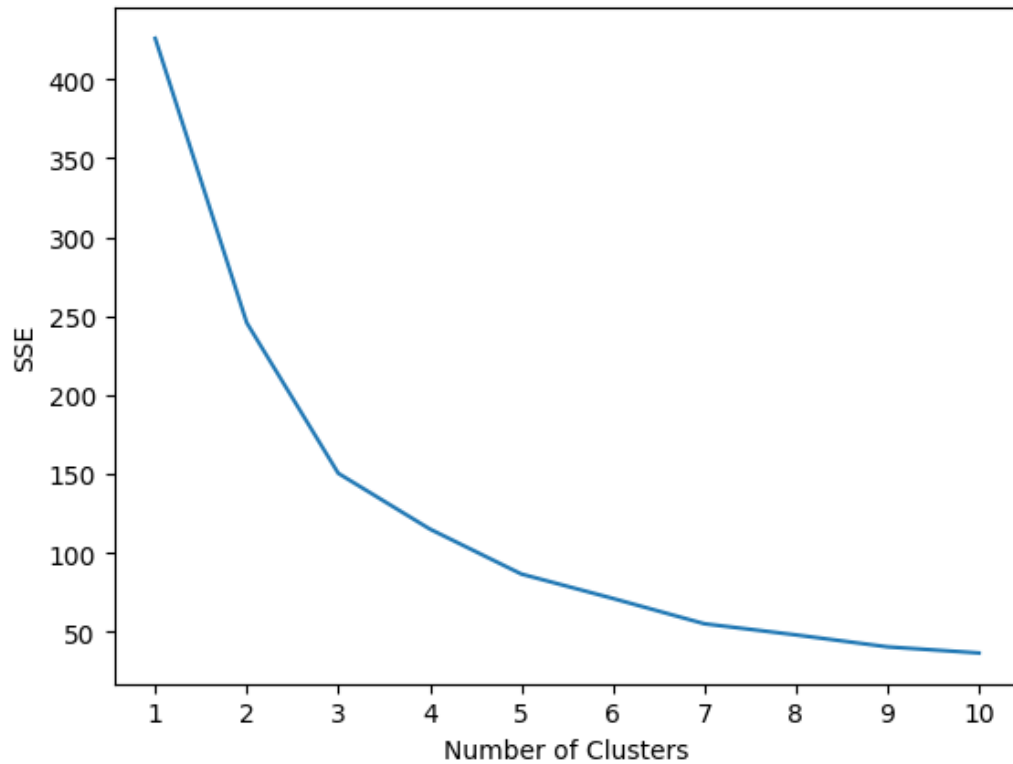


Figure 3.3. Elbow Method Applied to Clustering of Trips with 10-15km Distance

Figure 3.3 shows that the optimal number of clusters for this data is 5, because the decrease in SSE after 5 clusters is not as large as the decrease in adding up to 5 clusters. As a result, the k-means algorithm was implemented on this data with $k=5$ clusters. The clusters are shown in Figure 3.4.

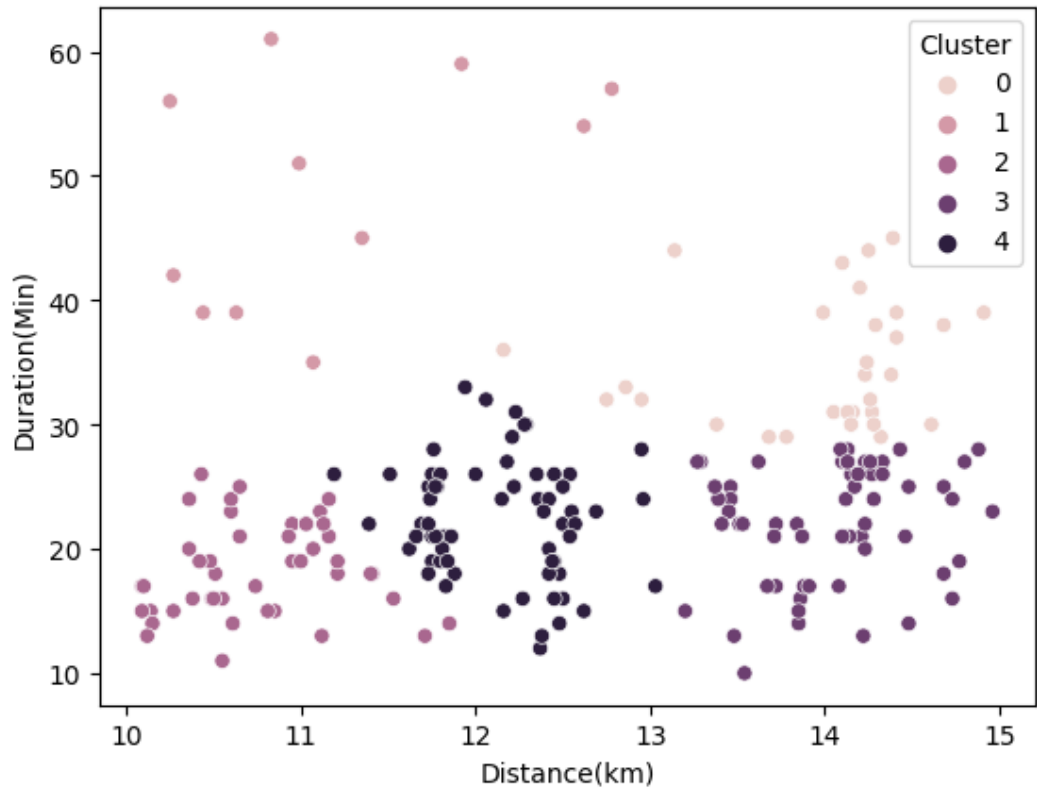


Figure 3.4 K-Means Clustering of Distance and Duration Applied to All Trips with 10-15km Distance

Figure 3.4 shows similar trends to the clusters of trips submitted by individual participants noted in Figure 3.2. For example, the tangerine-coloured cluster around 12km from Figure 3.2 was grouped into Cluster 4 along with other trips of similar distance and duration in Figure 3.4. The indigo cluster around 14km from Figure 3.2 is split between Clusters 0 and 3 in Figure 3.4. This shows that even regular commutes, which appear to stand out in Figure 3.2, do not differ statistically from trips recorded by other participants. As a result, it does not appear that different driving behaviour is a justifiable reason to limit the database. However, the total number of trips submitted by just a few drivers could undesirably influence the model by weighting the driving and vehicle factors of the highest-contributing participants more in the EC model.

Figure 3.5 shows the contribution of each driver in the original database, the database limited to two weeks per driver, and the database of two weeks broken into 500m distance segments. The importance of segmenting the database as well as the segmentation

methodology will be explored in the following sections. In the final model, both the trip database and segmented database will be utilised.

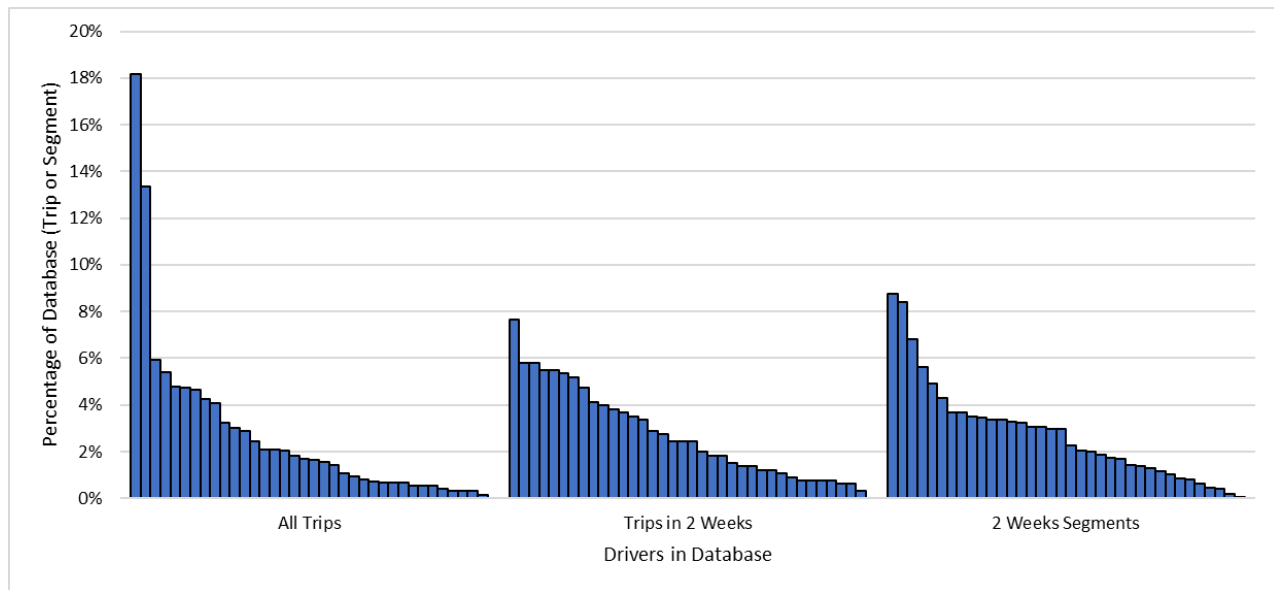


Figure 3.5. Share of Database by Driver

From Figure 3.5 it can be seen that the reduction of trips by the highest contributors, with each column representing an individual driver, was necessary to create a more generalised dataset. Before this change, one driver out of 36 was contributing over 18% of the trips to be modelled. The standard deviation of trip contribution percentage by driver dropped from 3.64% in the original database to just 1.94% in the two-weeks database, marginally increasing to 2.14% in the segmented database. The total number of trips decreased from 1520 to 653 when limited to two weeks of data per driver. The 653 trips consist of 22,916 segments.

After limiting the database to two weeks, further cleaning of the database took place. First, nine trips were removed for not having weather data including temperature and wind speeds. Twenty-nine trips did not have complete GPS data, which limited the information on gradient and other important variables. One trip was removed from the model because it showed an EC of -33.05 kWh. This issue was limited to the final segment of the trip, meaning the trip probably included data when the driver was charging their vehicle. Additionally, three outlier trips were removed from the database for having unrealistic speed or battery

power values. One trip reported speeds of 255 km/h, which is known to be an error code recorded by the OBD device (Pan et al., 2023). This error was confined to one period of idle time, so only the segment with the error value was removed from the segment-level model. The two other trips were removed for having negative battery power values, which was found to be an issue of missing values and excessive interpolation due to errors in data recording.

In total, 652 trips remained in the database for use in modelling. This amounted to 1,057,201 seconds of driving data and 25623 driving segments, for an average of 41.26 seconds per 500m segment and 24.25 minutes per trip.

3.3.2 Model Parameters

As explained in the introduction to this chapter, the EC model is designed for use later in the thesis to address other objectives to be discussed later. This does not limit the potential applications of this or a similar model. The later applications will only use variables from synthetic DCs and routes on a map, which can provide a wealth of information on factors that influence EC. Synthetic DCs vary from real DCs because they are not directly recorded speeds of actual vehicles but are developed through algorithms to satisfy the specific needs of the research. Synthetic DCs provide speed and acceleration data. The route can give information on the road environment such as gradient, distance, and duration. The only excluded variables then are battery related parameters and auxiliary power consumption, both discussed in Section 3.2. Battery parameters such as voltage and current are sometimes used to model EC but are not needed when other variables are used to predict battery power and EC. Auxiliary power consumption, which has been shown to have an effect on EC (Desrevaux et al., 2020; Hariharan et al., 2022; Liu et al., 2018), is not included in this model as it was not able to be captured by all drivers. If needed, it may be supplemented to the model using average values in a similar manner to Miri et al. (2021).

Thus, the model includes speed- and acceleration-based parameters, drive mode percentages, trip characteristics such as distance and duration, and characteristics from the route including gradient, road type, and if the location is urban. In addition, environmental parameters such as temperature and wind speed are examined to see if they have a major effect on EC. This set the framework for the variables used in the model which are explained in more detail below.

3.3.2.1 Driving Behaviour Parameters

The exact parameters which describe driving behaviour are broadly similar for both validating DCs and modelling EC. The most important of these parameters are the mean and maximum values of speed, acceleration, and deceleration (Kamble et al., 2009; Komorska et al., 2021; Wu et al., 2023; Xie et al., 2020). Maximum speed was not included in this analysis because of its high correlation with average speed at the segment level, and average speed was kept because of its higher correlation with the target variable. The standard deviation of these parameters can also be included in the modelling to account for the variability of driving style throughout a trip or segment (Komorska et al., 2021; Li et al., 2021; Maity and Sarkar, 2024; Wei et al., 2023; Xing et al., 2020). Drive mode percentages, which describe the amount of time the driver spends cruising, idling, accelerating, and decelerating, are more likely to be included as measurements in data-driven models such as in Yang et al. (2023). Other models, such as physical or simulated models, would not be able to include these variables as the models account for the continuous instantaneous movements of the vehicle (Genikomsakis and Mitrentsis, 2017; Modi and Bhattacharya, 2022; Schäfers et al., 2023). For the data-driven model in Lee et al. (2023), drive mode percentages are a key predictor especially as acceleration and deceleration phases can cause large differences in EC, in part due to regenerative braking during decelerations. Setting a threshold of acceleration is common practice in the development in DCs to allow for small variations in speed during cruising, so it was adopted here (Pan et al., 2015; Gong et al.,

2018). In this thesis, cruise time is defined as a period of speed above zero while acceleration and deceleration are below a minimum threshold of 0.14 m/s^2 , equivalent to a change of 0.5 km/h from second to second. Average moving speed is taken as the average speed when idle times (times when speed is equal to zero) are removed from the trip or segment.

After considering the literature, it is necessary to return to the data collected for this thesis. The descriptive statistics for the included parameters of speed-based, acceleration-based, and drive mode percentages can be found in Tables 3.1-3.3. The descriptive statistics displayed here and in the following sections contain the statistics of the segmented database and not the trip-level data.

Table 3.1. Descriptive Statistics of Speed-Based Variables

	Avg. Speed (km/h)	Avg. Moving Speed (km/h)	Std Dev Speed (km/h)
Min	0.33	1.10	0.00
Max	146.82	146.82	43.05
Median	69.14	69.14	3.69
MAD	28.06	27.18	5.20
Mean	67.32	68.19	6.25
Std Dev	32.49	31.26	6.42

Table 3.2. Descriptive Statistics of Acceleration-Based Variables

	Avg. Accel. (m/s^2)	Max. Accel. (m/s^2)	Std Dev Accel. (m/s^2)	Avg. Decel. (m/s^2)	Max. Decel. (m/s^2)	Std Dev Decel. (m/s^2)
Min	0.00	0.00	0.00	-2.41	-6.67	0.00
Max	2.62	6.11	1.68	0.00	0.00	1.89
Median	0.28	0.56	0.13	-0.28	-0.56	0.14
MAD	0.22	0.65	0.18	0.23	0.68	0.19
Mean	0.34	0.81	0.21	-0.35	-0.85	0.22
Std Dev	0.28	0.80	0.22	0.29	0.84	0.23

Table 3.3. Descriptive Statistics of Drive Mode Variables and Number of Stops

	Idle Time (%)	Cruise Time (%)	Accel. Time (%)	Decel. Time (%)	Number of Stops
Min	0.00	0.00	0.00	0.00	0.00
Max	98.00	100.00	100.00	100.00	25.00
Median	0.00	45.00	22.00	22.00	0.00
MAD	6.14	27.02	15.72	15.28	0.49
Mean	3.51	50.72	23.28	22.50	0.28
Std Dev	11.52	30.98	18.92	18.55	0.99

The minimum, maximum, mean, median, and standard deviations are common statistical properties, which are included in descriptive statistics tables to give readers an

idea about a parameter's distribution. The relationship between a parameter's mean and median can provide information about the parameter's distribution. Generally, if a median and mean are equal, then the parameter can be assumed as a random variable with a normal distribution. If the median is not equal to the mean, then the distribution is skewed (Lewis, 1993). Additionally, the mean absolute deviation (MAD) is provided here to give the reader an additional measurement of spread. The MAD is helpful in determining spread because it is a calculation of the average distance of points from the mean (van Eekelen et al., 2022).

From Table 3.1, it can be determined that average speed and average moving speed follow closely with normal distributions. For the standard deviation of speed parameter, the mean is much higher than the median. As such, this parameter has many low values, but the distribution is skewed, with a long tail on the right-hand side due to several higher values. Each parameter in Table 3.2 follows a similarly skewed distribution, with a higher number of points closer to zero. The drive time parameters in Table 3.3 do not follow a normal distribution as the most common values in a segment are either 0% or 100%, indicating one drive mode was constant throughout the segment. Additionally, most segments did not include any stops.

Tables 3.1-3.3 show an expected range of parameters for a diverse database. The range of velocities in Table 3.1 exemplifies the range of trips from urban to motorway with high standard deviations of speed. Table 3.2 also provides examples of a complete, well-constructed database since the acceleration parameters closely mirror the deceleration parameters. Finally, Table 3.3 displays mean drive model percentages that could be expected in a segmented database where idle time is generally low on average for all segments and again acceleration behaviour closely mirrors deceleration. The average idle times are low because most 500m segments only include the vehicle moving. Acceleration and deceleration should complement each other because they are similar actions. A final parameter based on the driving behaviour measured the number of stops during a segment,

when the speed and acceleration values are at zero. This parameter captures information on the traffic level of the trip (Ji et al., 2022).

3.3.2.2 GPS-Based Parameters

The GPS coordinates of trip routes were obtained during data collection, but by themselves no information can be drawn. The raw coordinate data serves as a springboard for gathering information on the trip. Other sources of data are required to add information that increase the usefulness of GPS to model EC. This section describes how other sources of data were incorporated and attached to GPS coordinates to add important parameters such as gradient and temperature.

The first GPS-based parameter, gradient, is also the only parameter added without the use of a separate data source. Along with GPS latitude and longitude, elevation was also recorded by each driver through the Car Scanner ELM OBD2 app on their smartphone. Gradient has been shown to have a major influence on EC including heavily influencing regenerative braking capabilities (Genikomsakis and Mitrensis, 2017; Liu et al., 2017; Liu et al., 2018; Settey et al., 2021), but it is not a simple process to gather high quality elevation data. Liu et al. (2021) among others note that GPS location on smartphones can be inaccurate, especially in the z-direction which provides elevation data. Qi et al. (2018) collected GPS location data which was then mapped on a high-fidelity map to find elevation data. However, due to the unavailability of a high-fidelity elevation map covering Ireland, this elevation data was used with minor adjustments to see if meaningful gradient information could be added to the model. A method similar to that presented in Duran and Earleywine (2012) was used to filter the raw GPS elevation data. First, excessively large jumps in elevation were removed. Although the Design Manual for Urban Roads and Streets recommends a maximum gradient of 8.3% (Government of Ireland, 2019), rural and old roads around Ireland are likely to have greater elevation changes than newly designed urban roads. So, gradient limits were determined after observation of the data to be a 12% increase

or decrease in slope. Anything higher than these values were outliers in the data and replaced by a rolling average with a window of 5 adjacent values.

Next, weather data was required to provide insight on temperature (Al-Wreikat et al., 2022; Liu et al., 2018; Yang et al., 2023). and wind (Agrawal et al., 2023; Modi and Bhattacharya, 2022; Schäfers et al., 2023) which are heavily present as features in the EC modelling literature. Transport Infrastructure Ireland (TII) maintains a set of roadside weather stations around Ireland which measure current information such as temperature, wind speed, and wind direction. Figure 3.6 shows the locations of each weather station which are spread throughout the country. Figures 3.6, 3.9, and 3.10 are displayed on a standard projection EPSG:4326.

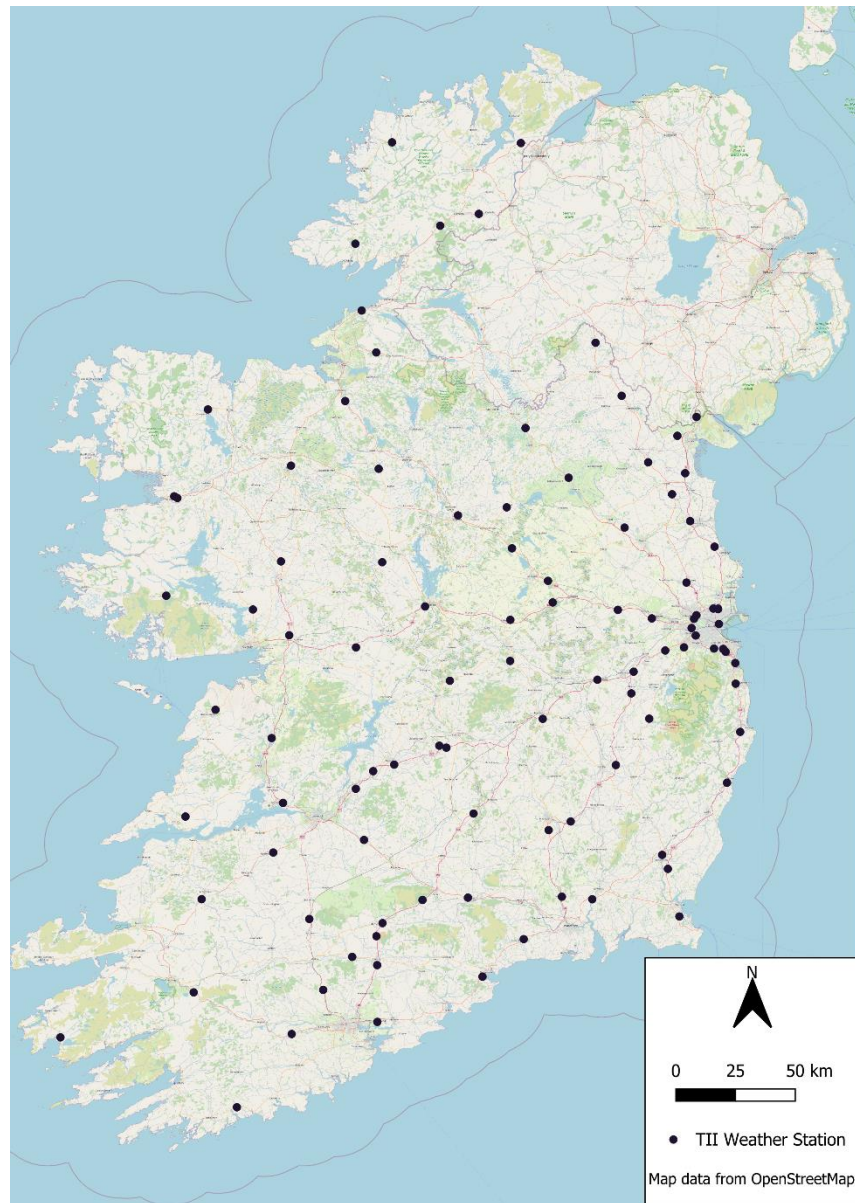


Figure 3.6. TII Weather Stations (Transport Infrastructure Ireland, 2024A)

TII provided historical information for temperature, precipitation, wind speed, and wind direction between August 2023 and August 2024 (Transport Infrastructure Ireland, 2024A). To make use of these data, an algorithm was developed to detect the nearest weather station to the coordinates of each trip at each second. The algorithm determined the closest weather station using the haversine formula to calculate the distance from the trip to each weather station. Then, by matching the date and nearest time, the aforementioned parameters from the closest weather station at the trip's date and time were attached to trips. The average values of temperature, wind speed, and wind direction were recorded for the trip or segment.

Rain was added as a dummy variable to show whether precipitation was present during each trip or segment. It was raining for only 8% of the recorded segments.

An additional variable called wind factor was developed to observe the differing impacts of headwinds and tailwinds on the trip. In addition to wind speed, wind direction in comparison to trip direction has a significant effect on EC (Vepsäläinen et al., 2019). A headwind acting against the vehicle will have a negative effect on EC by increasing the wind resistance (Shibata and Nakagawa, 2015), while a tailwind behind the vehicle could lessen EC. To examine this effect, a parameter called a wind factor was developed. The wind factor was calculated by first determining the angle between wind direction and trip direction. This angle was converted to a number between -1 and 1, where 1 indicated a direct headwind acting on the trip and -1 representing a direct tailwind. Wind acting perpendicular to the trip direction was given a value of 0. This number was then multiplied by wind speed to get the wind factor. To still capture the effects of wind speed perpendicular to the vehicle, measured wind speed was also included as a predictor variable. Descriptive statistics of gradient and weather parameters are included below in Table 3.4.

Table 3.4. Descriptive Statistics of Gradient and Weather Variables

	Gradient (%)	Temperature (°C)	Wind Speed (km/h)	Wind Factor (km/h)
Min	-11.47	1.90	0.00	-10.27
Max	11.08	25.70	10.30	8.98
Median	0.00	14.20	1.70	0.02
MAD	1.06	3.10	1.01	1.05
Mean	0.00	14.05	2.04	0.02
Std Dev	1.65	4.00	1.34	1.44

The descriptive statistics show means near zero for gradient and wind factor, which is exemplary of an unbiased database, and the ranges of all parameters are as expected. The temperature range in Ireland is quite mild, but most of the yearly variation is captured in the database. Limiting the database to two weeks of data from each driver also reduced the range of variables captured. For reference, a six-month study performed in China captured temperatures between -2°C and 47°C (Pan et al., 2023). All parameters in Table 3.4 can be

assumed to be normally distributed random variables because the mean and median are exactly or nearly equal.

A final set of GPS-based parameters was added to incorporate information on the road environment, including road type and area of travel. Straub et al. (2022) demonstrates the importance of trip environment on EC calculations, especially when creating synthetic driving cycles. This study divides driving into three categories: urban, motorway, and rural. Morlock et al. (2020) considers five different road types such as freeways and primary roads, with the percentage of each type being used as a predicting variable. This study takes a similar approach based on how roads were classified in OpenStreetMap (OSM) (OpenStreetMap, 2024A). OSM is an open-source online map with details on real world roads and other features contributed by the public. Road information is commonly gathered in OSM for use in EC modelling and other research (Ayman et al., 2022; Guo et al., 2020; Wang et al., 2018). Six major road categories were found in the OSM data for Ireland and Northern Ireland: motorway, primary, secondary, tertiary, residential, and unclassified. Roadways in Ireland are classified as motorway, national primary, and national secondary with distinct speed limits and other characteristics (Transport Infrastructure Ireland, 2024B); however, it is still up to individuals to classify roads on OSM to the standards of that program instead of official government classification (OpenStreetMap, 2024B). That means many roads remain unclassified, and the national primary roads of Ireland, for example, may not always be identified correctly as a primary road on OSM.

Along with road type, the type of area in which the trips took place was considered. In addition to the previously mentioned Straub et al. (2022) including the percentage of each trip in urban areas as a factor in EC modelling, other studies have noted the importance of regional context on driving style and EC (Chu et al., 2023; De Cauwer et al., 2017; Rauh et al., 2015; Varga et al., 2019; Wu et al., 2023). It is natural to expect slower speeds, traffic, and frequent stopping in urban areas that would not occur on interurban trips on motorways

(Shibata and Nakagawa, 2015). Because the data was collected across the entire country of Ireland with no obvious discernment between types of trips, it was imperative to correctly classify trips based on their GPS coordinates. Since coordinates were recorded for every second of the data, a total of 1,004,278 coordinate pairs were included in the database. The computational time to classify each of these pairs would be extraordinary, and the GIS program would crash during the attempt. So, to limit the computational difficulty, the coordinates needed to first be rounded to a certain degree for only the remaining set of unique coordinates to be classified. The classification task was accomplished with the free geographic information system software QGIS (QGIS.org, 2024) as used in Zhang et al. (2022). First, the coordinates were rounded to three decimal points, which lowered the unique coordinate pair count to a much more manageable 44,297. Upon visual inspection, however, the coordinates were too rounded and could not reliably be connected to the nearest road as seen in Figure 3.7.

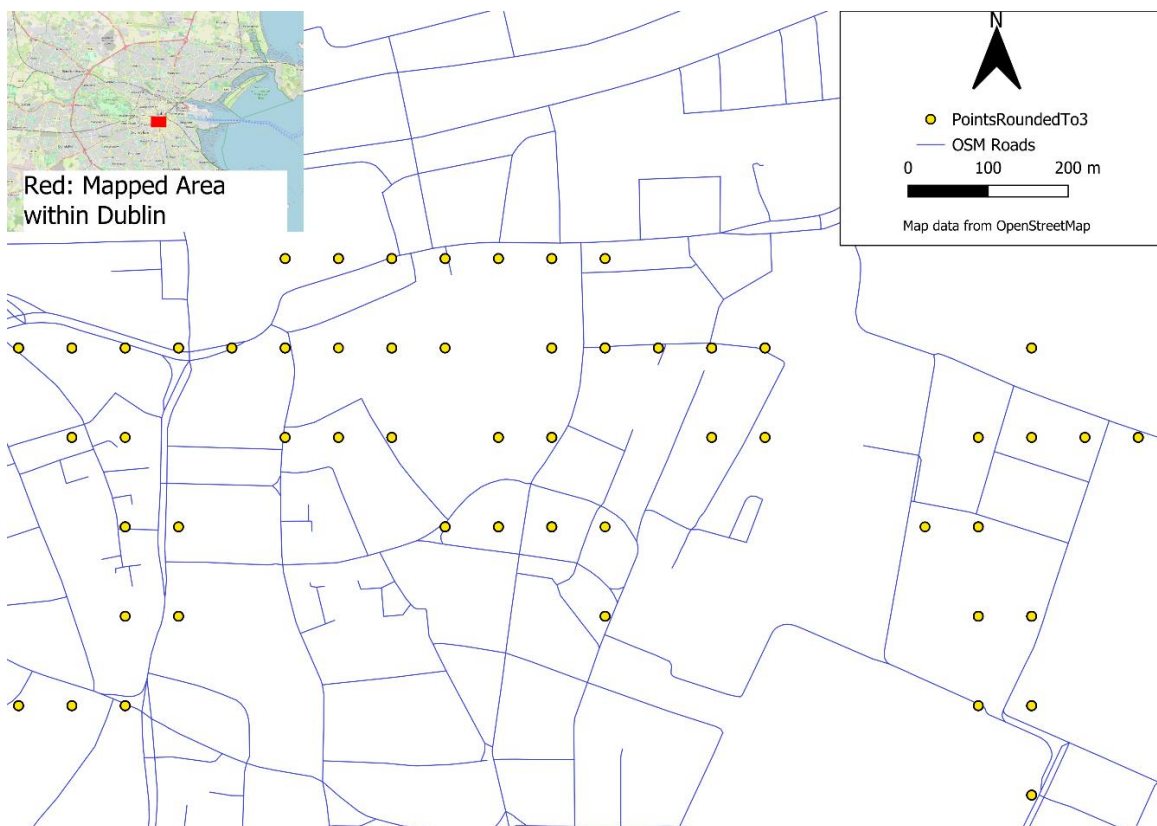


Figure 3.7. GPS Coordinates Rounded to Three Decimal Places

The rounded GPS coordinate pairs, each yellow dot in Figure 3.7, did not seem to reliably follow a route and connect to a discernible road. So, the coordinates were instead rounded to four decimal places, with the improved result seen in Figure 3.8.

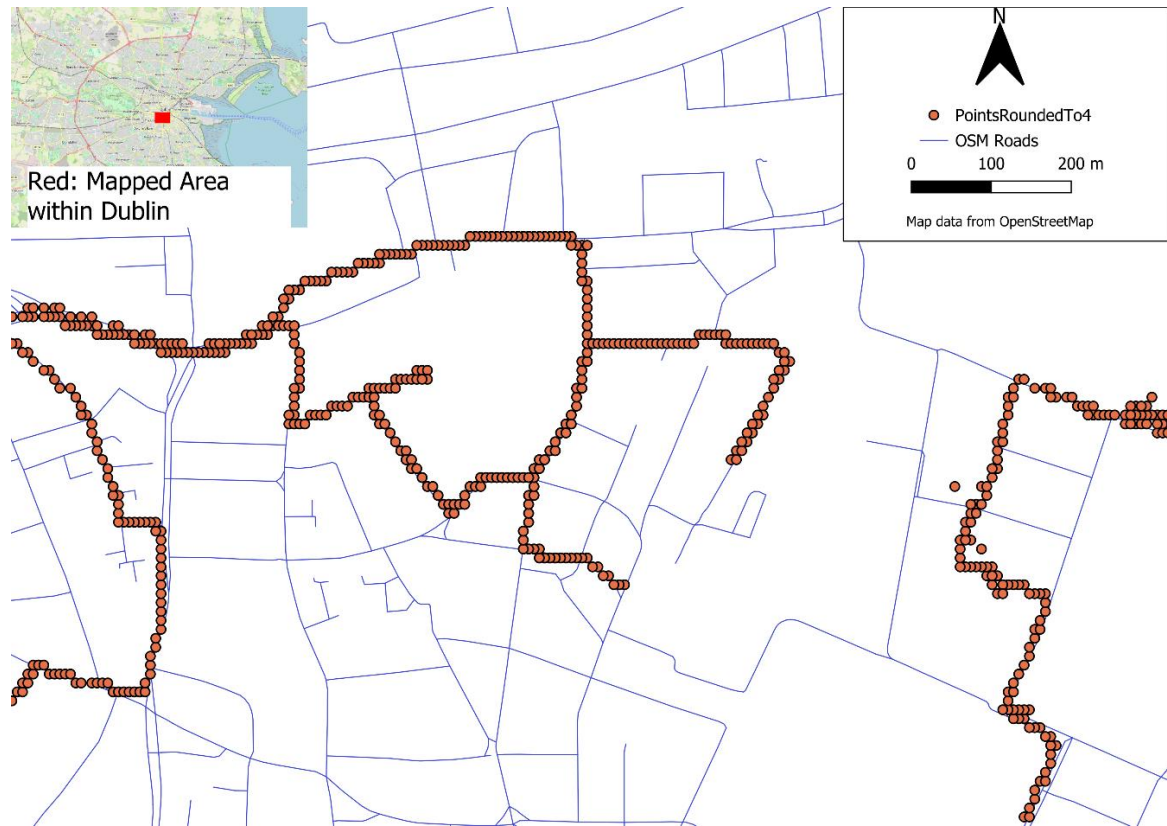


Figure 3.8. GPS Coordinates Rounded to Four Decimal Places

Although increasing the detail of unique coordinates to four decimal places raised the number of pairs to 305,964, the improvement to the data was clear. Although still not directly aligned with the roads, the coordinates could still be reliably attached to the nearest road in an automated process for a suitable understanding of road type during a trip.

To understand the percentage of trips in urban areas, a map of urban areas across Ireland was needed. The University College Dublin Library provides a map of all legal towns and cities across Ireland (UCD Library, 2024). This map was easily incorporated into the QGIS landscape, and the urban area shapes can be seen in Figure 3.9.

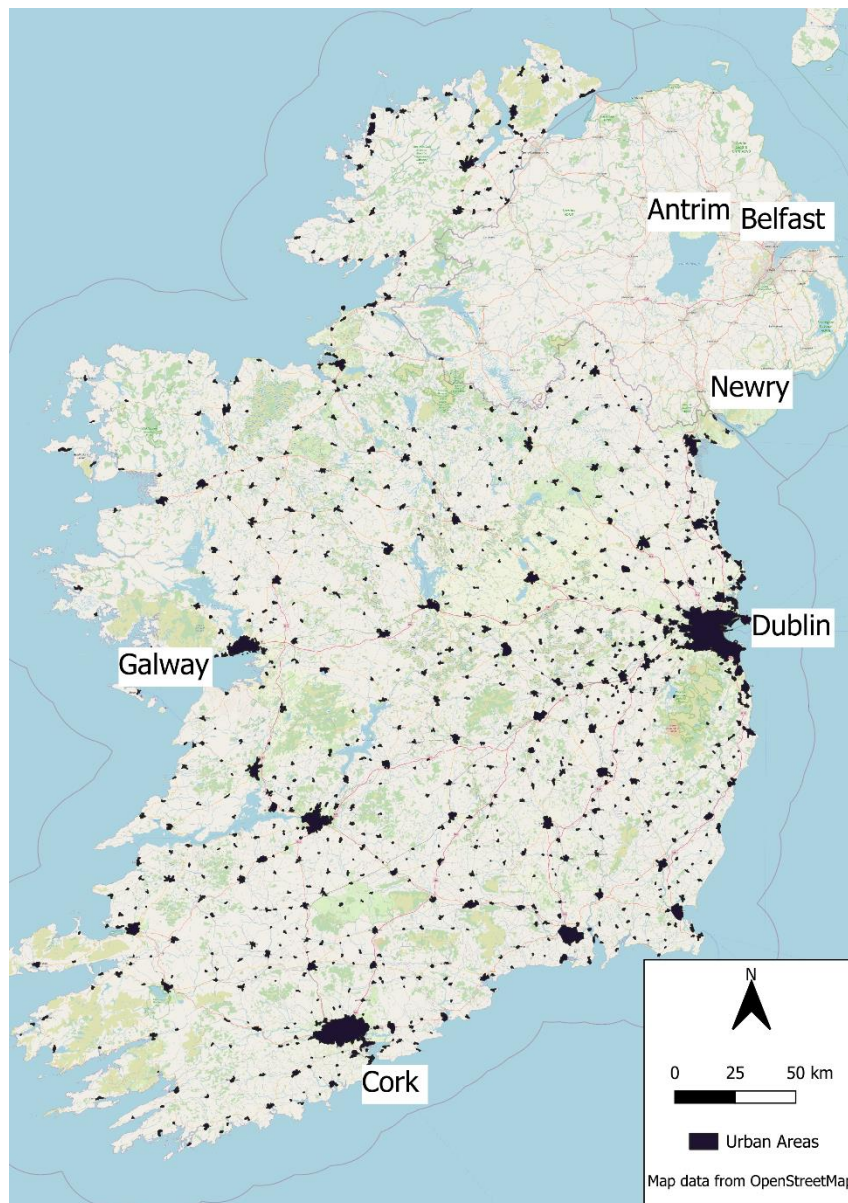


Figure 3.9. Ireland Legal Towns and Cities

The towns and cities cover all types of urban areas in Ireland, but Northern Ireland remained devoid of shapes in Figure 3.9 due to the different legislative authority over the area. So, additional shapes were drawn over the cities of Antrim, Belfast, and Newry in Northern Ireland where some trips were made. After this preparation, the coordinate pairs were labelled as urban if they were contained in any of these urban shapes through an automated process in QGIS. Each coordinate was also labelled with a road type through a separate automated process in QGIS which found the nearest road to the coordinates. A final step was made to classify urban trips by removing the urban tag on any second of data where the

speed was recorded to be over 50 km/h. This was done because the speed limit in Irish towns and cities is 50 km/h (Citizens Information, 2024), and some trips may have been classified as urban even though they were motorway trips passing through a town. So, high speed datapoints were not classified as urban.

Road surface was also considered as an important characteristic in the EC calculations (Lee et al., 2023). Road surface can alter the coefficient of friction which changes the amount of energy needed to propel the vehicle forward (Mediouni et al., 2022; Schäfers et al., 2023). However, information from OSM indicated that 140,032 roads in Ireland were paved with asphalt and only 5765 were made of concrete, the next most used material. So, little information could be gathered on road surface types and the parameter was not included.

The descriptive statistics for road environment can be seen in Table 3.5.

Table 3.5. Descriptive Statistics of Road Environment

	Urban (%)	Primary (%)	Secondary (%)	Tertiary (%)	Unclassified (%)	Motorway (%)	Residential (%)
Min	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Max	100.00	100.00	100.00	100.00	100.00	100.00	100.00
Median	0.00	0.00	0.00	0.00	0.00	0.00	0.00
MAD	32.89	10.33	39.08	21.48	22.33	31.85	18.69
Mean	22.58	5.51	29.65	13.21	13.54	20.18	11.38
Std Dev	39.13	22.39	43.26	31.02	32.28	39.59	28.62

As with the drive mode percentages of Table 3.3, the road environment parameters in Table 3.5 have an unequal distribution, with most segments having either 0% or 100% for these variables. The mean values in Table 3.5 can help to inform that motorway and secondary roads were the most common road types. An interesting piece of the data is that only about 23% of the recorded trips occurred in urban areas. Since most participants come from Dublin, it was supposed that most trips would be around that area. However, there appears to be a wide range of trips across urban and rural areas, and different road types. Figure 3.10 displays the coordinates for all recorded data to show the locations of all trips in the database.

While urban areas such as Dublin and Galway are hubs in the trip map, many trips are spread across rural areas.

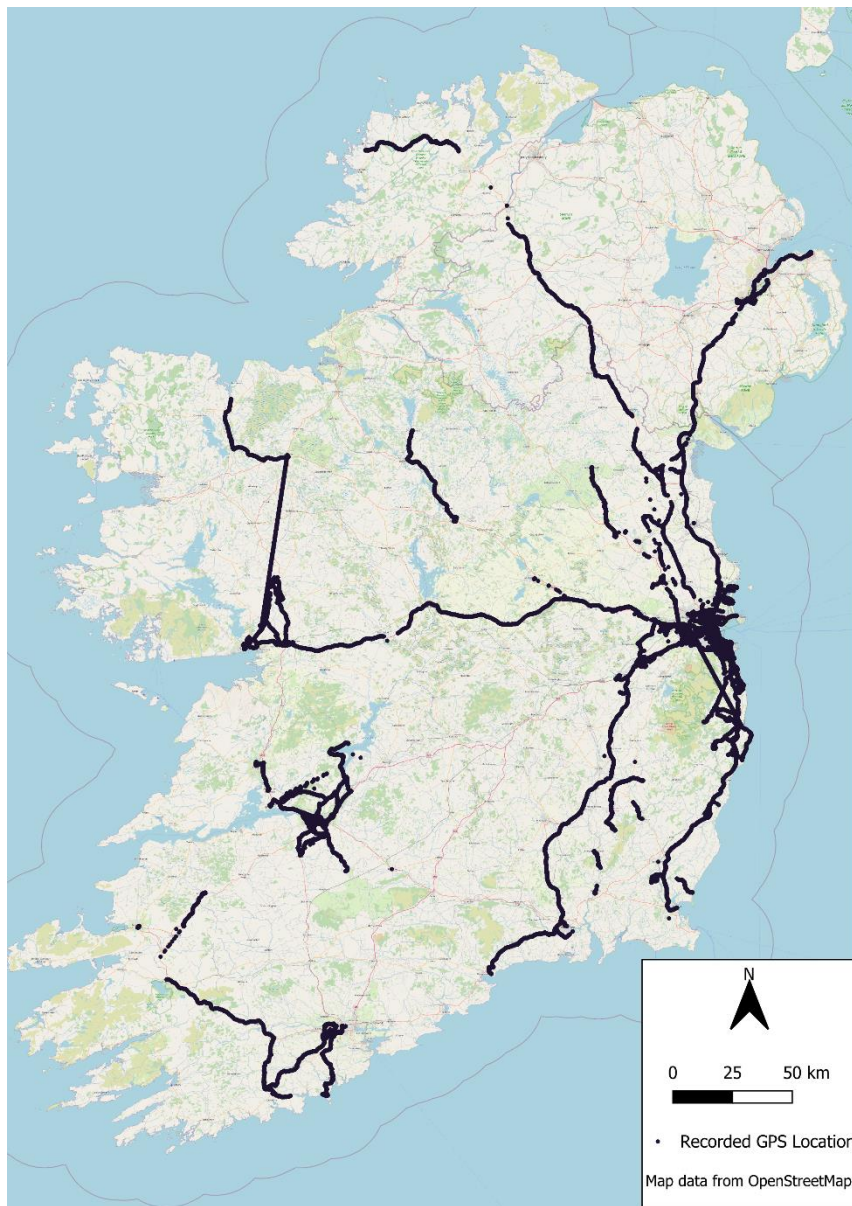


Figure 3.10. Coordinate Locations of All Trips

3.3.2.3 Driver and Vehicle Parameters

The final category of predictor variables considered were those pertaining to the vehicle and driver. Some EC models are based on just one vehicle and is thus a model specific to that vehicle (Janpoom et al., 2023; Lee et al., 2023; Sasank et al., 2023). On the other hand, a generalised model is built on data from a variety of vehicles (Ahmed et al., 2022; Alhanouti and Gauterin, 2024; Mavlonov et al., 2023). Of course, adding more

vehicles to the model adds some inherent variance because all vehicles will perform and consume energy differently. Parameters such as weight and tyre pressure that have effects on the physical resistance faced by the EV and therefore the energy consumed (Mavlonov et al. 2023; Settey et al., 2020) were not known for all trips and are therefore not included in the model. For this study, the age and battery capacity of the vehicle were considered as independent variables to capture the variance introduced by recording data in different vehicles. In addition, the sex of the driver was considered to observe any potential difference in driving behaviour on the basis of gender.

It was theorised that a driver's experience with EVs could play a role in their ECR, but not all drivers supplied these data. Therefore, this parameter could not be included in the model. Still, a correlation analysis was run on segments which included this piece of information, and experience driving EVs (in years) was found to have a low correlation of only 0.025 with average battery power (ABP). This was an order of magnitude lower than the correlation between ABP and the driver and vehicle parameters included in the model, so this excluded variable would not have improved the model significantly. Rauh et al. (2015) suggests EV driving experience is an important factor to limit range anxiety, but it does not appear to significantly decrease the day-to-day ECR.

Table 3.6 displays the vehicle information for each driver as well as the number of trips and segments supplied by each to the database.

Table 3.6. Driver and Vehicle Information

Driver	All Trips	Trips in 2 Weeks	Segments	Vehicle	Vehicle Year	Battery Size (kWh)	Sex
1	82	35	2014	Kia EV6	2022	74	Male
2	37	12	1930	Cupra Born	2023	58	Male
3	203	36	1572	Kia EV6	2024	77.4	Male
4	32	18	1294	Volkswagen ID4	2023	77	Male
5	276	38	1140	Hyundai Kona	2019	64	Male
6	65	27	996	Hyundai Kona	2023	64	Female
7	73	38	856	Hyundia Kona	2022	39.2	Male
8	10	7	850	Hyundai Ioniq	2018	28	Male
9	62	24	810	Nissan Leaf	2017	28	Male
10	25	12	793	Nissan Leaf	2021	39	Male
11	90	25	782	Hyundai Kona	2024	39.2	Male
12	11	8	781	Renault Megane Etech	2024	60	Male
13	28	26	756	Hyundai Kona	2021	64	Male
14	24	23	750	Volkswagen ID4	2021	77	Female
15	44	34	710	Hyundai Kona	2021	39.2	Female
16	31	19	709	Kia Soul	2019	64	Male
17	72	31	690	Nissan Leaf	2023	39	Female
18	14	9	688	Peugeot e208 Allure	2021	46.3	Male
19	71	36	529	Kia EV6	2024	77.4	Male
20	49	50	479	Nissan Leaf	2018	39	Female
21	46	16	461	Nissan Leaf	2019	39	Male
22	6	5	433	Hyundai Ioniq	2024	74	Male
23	26	22	400	Nissan Leaf	2019	59	Male
24	8	5	393	Nissan Leaf	2020	39	Female
25	10	5	321	Nissan Leaf	2023	59	Male
26	22	16	301	Volkswagen ID3	2021	58	Female
27	32	16	276	Hyundai Kona	2023	39.2	Female
28	8	9	241	Nissan Leaf	2018	39	Female
29	5	5	192	Hyundai Kona	2023	64	Male
30	16	8	183	Peugeot 2008	2023	46.3	Female
31	5	4	164	Volkswagen ID3	2021	58	Male
32	12	12	152	Volkswagen ID3	2020	58	Male
33	8	4	105	Nissan Leaf	2024	39	Male
34	10	10	100	Nissan Leaf	2018	39	Male
35	5	6	47	Audi Etron	2022	88	Male
36	2	2	18	Volkswagen ID3	2021	58	Male

Although vehicle make and model are not considered in the model, it is interesting to note that many drivers in the study have either a Nissan Leaf or a Hyundai Kona, so the effects of unmodelled parameters stemming from weight or body type may lean towards those vehicles in the model. Some popular vehicles such as Teslas did not share information on battery power, which is crucial to the model, so they were not included in the database.

3.3.4 Machine Learning Models

An essential step in the methodology of EC prediction is selecting the correct model to fit the data. For a prediction problem based on large amounts of data, various machine learning (ML) techniques can be implemented. Through the literature review on EC models, three major ML algorithms were identified: multiple linear regression (MLR) (De Cauwer et al., 2017; Ekici et al., 2023; Liu et al., 2018), decision tree algorithms (Nan et al., 2022; Sun et al., 2019; Ullah et al., 2022; Zhao et al., 2020), and artificial neural networks (ANNs) (Cheng et al., 2018; Cura et al; 2021; Modi et al., 2020; Vatanparvar et al; 2019). All of these methods involve using the set of predictor variables, also called features or independent variables, to estimate a target variable, in this case EC. ML models are generally well suited to sift through large datasets and find patterns in the variables, which means they have the potential to predict EC for this dataset. These methods were all tested to select the best ML model based on performance.

MLR is a relatively straightforward statistical model that assigns the linear coefficient to each predictor variable that limits the total error of the predictions (Ahmed et al., 2022). MLR is hindered by its inability to fit non-linearities between a predictor and the target variable, but it has still been shown to be effective in the literature and can provide direct insight into the importance of all predictors (De Cauwer et al., 2017; Ekici et al., 2023). Equation 3.1 shows the theory behind MLR, where Y_i is the target variable, in this case EC, β_0 is a constant term that captures the general effect of all variance not captured by the independent variables, X_i is an independent variable, and β_i is the coefficient term that when multiplied by the independent variable creates the lowest sum of squared errors between the predicted EC and actual EC.

$$Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots \quad (3.1)$$

Decision tree algorithms improve on the shortcomings of MLR by splitting features in the most effective way possible to classify or predict the output variable (Sun et al., 2019).

These algorithms have the added benefit of not putting much weight on less important features when many features are included in the model (Yavasoglu et al., 2019). The decision tree algorithms used in this paper were those observed as popular and well-performing in the literature: Extreme Gradient Boosting (XGBoost) (Nan et al., 2022; Zhao et al., 2020) and Light Gradient-Boosting Machine (LightGBM or LGBM) (Ullah et al., 2022; Zhao et al., 2020). The theory behind decision tree algorithms can best be shown through a simple example. The theory can then be applied to more complicated datasets. Figure 3.11 shows a simple example of a decision tree algorithm used for regression.

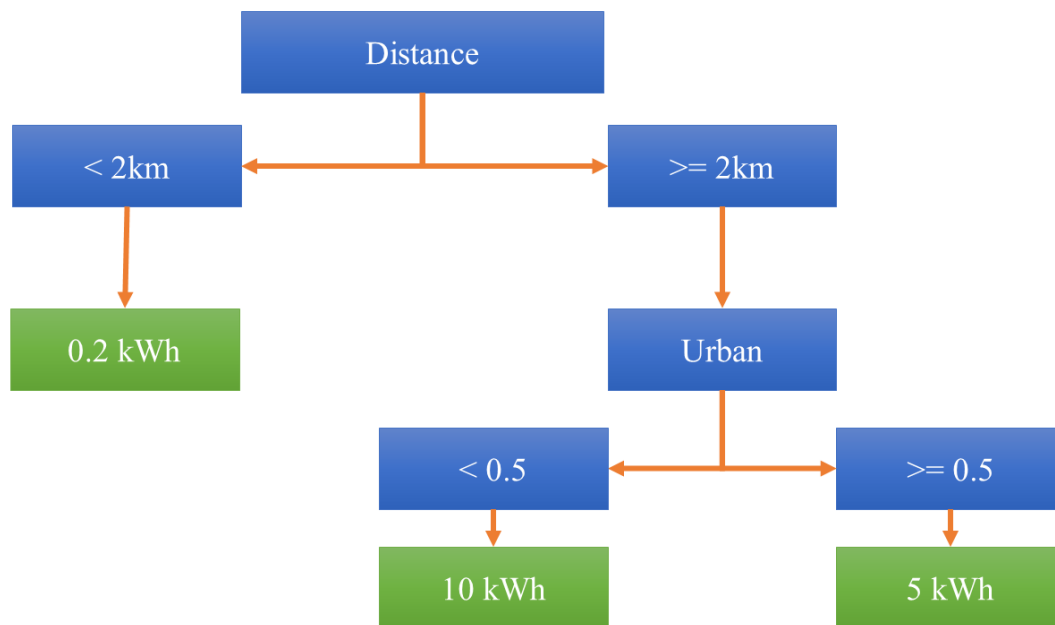


Figure 3.11. Regression Tree Example

It should be noted that Figure 3.11 does not contain actual values from the data. Decision tree algorithms split the data from the root node, which is the best predictor variable (Qi et al., 2018). In the case of Figure 3.11, the root node is distance. From there, branches sequentially divide data based on the values of features with decreasing importance. Finally, leaf nodes, seen in green in Figure 3.11, are found to predict the value of the target variable, in this case EC. The benefit of decision tree algorithms such as LightGBM and XGBoost is their ability to automatically determine feature importance and identify patterns in the data.

XGBoost is a gradient boosted decision tree algorithm widely used in ML. Boosting refers to the process of building many computationally simple models called weak learners with limited prediction accuracy and combining the predictions of these weak learners into a prediction model with low error (Friedman et al., 2000). In the case of gradient boosted decision trees, each weak learner is a unique regression tree like the one in Figure 3.11. XGBoost is a quick but thorough algorithm due to its unique method of finding the best split for each branch of the tree in a computationally efficient manner (Chen and Guestrin, 2016). The algorithm first recognises candidate split points according to the distribution of values for each parameter. It is also one of the first decision tree algorithms developed for use in sparse datasets, where many values are missing or zero (Chen and Guestrin, 2016). Additional improvements of XGBoost stem from its use of the computer itself, which is outside the scope of this thesis but can be found in Chen and Guestrin (2016).

LightGBM was developed to ease computational loads while running a gradient boosted decision tree similar to XGBoost. It accomplishes this task through a few techniques. First, the algorithm randomly drops some parameters with low information gain while creating each weak learner. Parameters with small gradients, meaning the parameters did not contribute much to minimising the loss function (error), are identified from previous weak learners and some are dropped at random. The algorithm is designed this way to improve computational efficiency by not creating branches for these dropped less significant parameters without losing all information that these parameters bring to the model (Machado et al., 2019). Unlike XGBoost which computes each possible value to split a branch, LightGBM uses histograms to bin values before finding the best value to split a branch. Another advantage of LightGBM, especially for large, sparse datasets, is its greedy bundling algorithm which merges features that never have a nonzero value simultaneously (Machado et al., 2019). The features used in this thesis are unlikely to benefit from greedy bundling, but in general it can greatly improve computational efficiency.

ANNs connect input variables to a predicted output through a series of weights and biases in hidden layers. These models have the benefit of finding complex patterns and connections between input variables while calculating the predicted output (Wei et al., 2023). In addition, after each hidden layer in an ANN there is an activation function which can help deal with nonlinearities. The general structure of an ANN can be seen in Figure 3.12. Deep learning can be achieved by including more hidden layers. The input layer is the set of independent variables, and the output layer is the target variable, in this case EC.

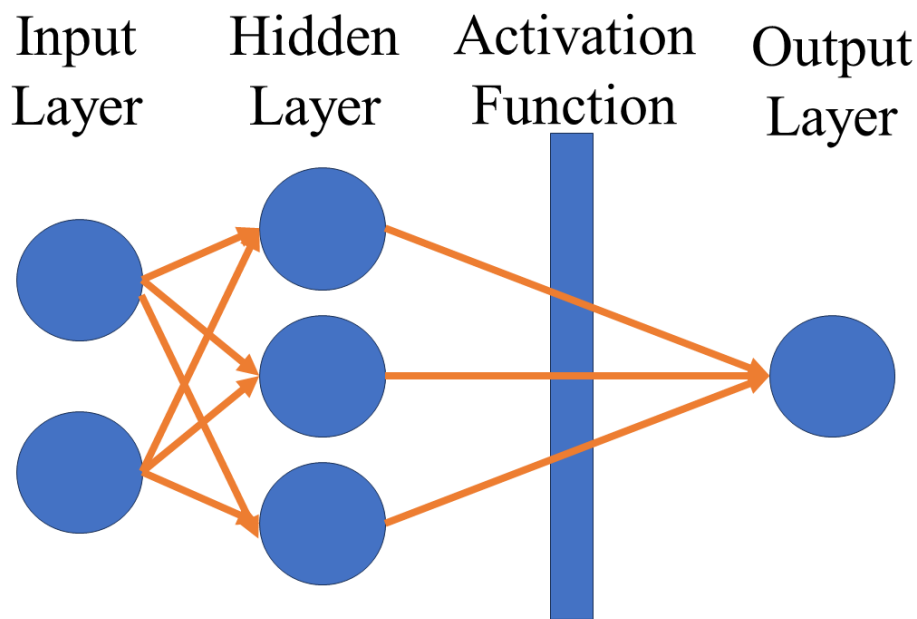


Figure 3.12. Artificial Neural Network Example

The activation function is part of the ANN which updates every epoch along with a set of weights and biases attached to each neuron in a layer which determines the predicted effects of the input layer on the output layer. In this paper, ReLU is used as the activation function which can set the importance of a feature or certain range of a feature to zero, meaning only certain values of each feature are included in the model (Cura et al., 2021; Manoharan et al., 2022). In other words, the ReLU activation function sets zero as the output of a neuron if the information is not valuable to the model. This can be useful for removing ineffective features from the model or fine-tuning heteroskedastic features of the data in ways that other models such as MLR are unable to. This means ANNs may be best suited

for big, varied data as is the case here; this also might be the reason why ANNs are the most popular model in the current literature. For example, a graph of EC of trips vs. their average velocity is shown in Figure 3.13. In this thesis, trip-level energy consumption, the amount of energy consumed over a whole trip in kilowatt-hours, is abbreviated as TEC.

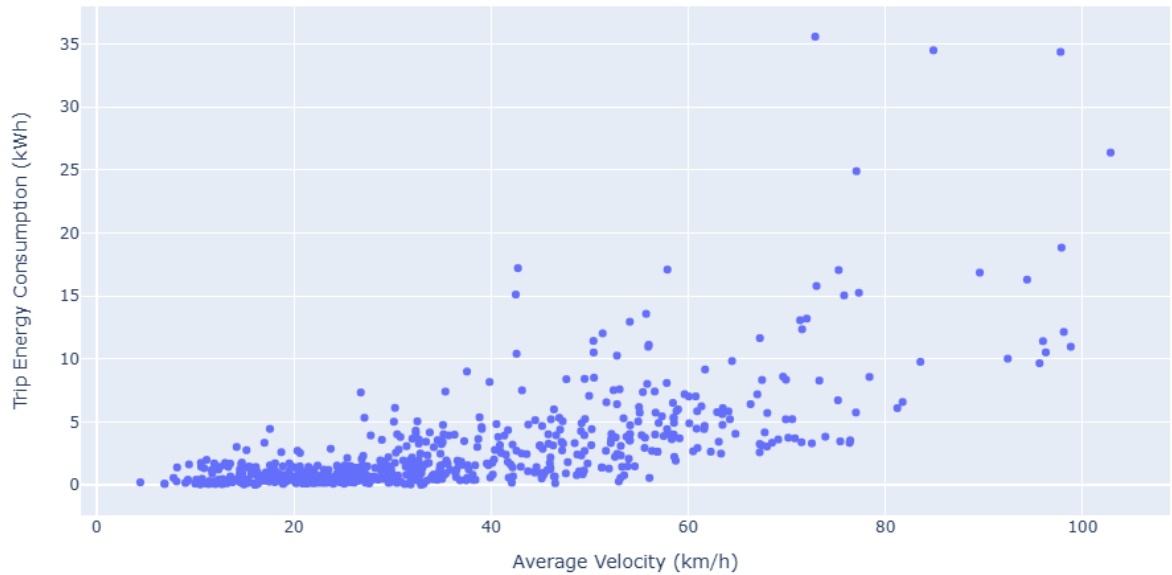


Figure 3.13. TEC vs. Average Speed of All Trips

From Figure 3.13, it can be seen that the average speed of trips below 30 km/h do not vary much in EC, but EC generally increases with increasing average speed after that threshold as seen in Figure 3.13. Figure 3.13 displays heteroskedasticity, where the variance of EC is not constant across the range of velocities. The variance in EC is larger at higher speeds. There appear to be distinct bands of EC vs. average speed at higher speeds, but it is unclear if this is a trend in the data or simply due to quirks in the available data. It was determined that because each band was made of trips by different participants, these bands are likely just quirks. Additionally, they could be due to several trips following similar speed limits on long-distance trips.

3.3.5 Model Evaluation

To compare the effectiveness of each method, it is necessary to check both the error and the goodness of fit. Many different error terms are used in the literature, but the most

common error measurements were found to be root mean square error (RMSE) and mean absolute error (MAE). RMSE is calculated using the following formula, Equation 3.2:

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_{pred} - y_{actual})^2}{n}} \quad (3.2)$$

where y_{pred} is the value of the target variable predicted by the model, y_{actual} is the empirical value of the target variable, and n is the number of samples.

MAE is calculated using Equation 3.3 on the following page:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_{pred} - y_{actual}| \quad (3.3)$$

where the terms mean the same as in Equation 3.2. To capture goodness of fit, the coefficient of determination, R^2 , is used as additional measure to compare model effectiveness. The R^2 value is found using Equation 3.4:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_{pred} - y_{actual})^2}{\sum_{i=1}^n (y_{avg} - y_{actual})^2} \quad (3.4)$$

Where again the terms retain their meaning from Equation 3.2, and y_{avg} is the average value of the target variable across all samples. MAE was the main consideration in model selection as the goal of this model is to minimise the error of TEC prediction.

3.3.6 Neural Network Hyperparameters

Generally, trial and error is the standard practice for tuning the hyperparameters of ML models (Cura et al., 2021; Kim and Cho, 2019; Li et al., 2021). Hyperparameters are values which change the model itself such as the number of layers in an ANN or the learning rate, which dictates how much new information changes the existing estimators of the model (Smith, 2017). Automated tuning of hyperparameters may be useful to optimise the usefulness of the model (Kim and Cho, 2019; Zhang et al., 2020A), but the computational

effort of these models was not high enough to necessitate any method other than trial and error.

There are five hyperparameters that were set before forming each instance of the ANN: the number of hidden layers, the loss function, the optimiser, the validation split, and the number of epochs. The number of hidden layers determines how many weights and biases will be calculated between the input and output layers. Adding more hidden layers can make the ANN a deep neural network, which can better discover and model more complex relationships between the input variables and the output EC (Wei et al., 2023). It has been suggested that neural networks require more than two hidden layers to be classified as deep (Diaz Alvarez, 2014), but adding more than two hidden layers did not decrease the error for the data in this thesis. Figures 3.14 and 3.15 show the MAE of segment-level EC (SEC) and TEC in test data using an ANN with one through five hidden layers, with the MAE shown being an average of five test-train splits. A test-train split in ML refers to splitting the data into two groups: train data, which the ML model uses to build its predictions and test data, which is not used to train the ML model but instead test the model on unseen data to ensure the model is not overfitted to the available data.

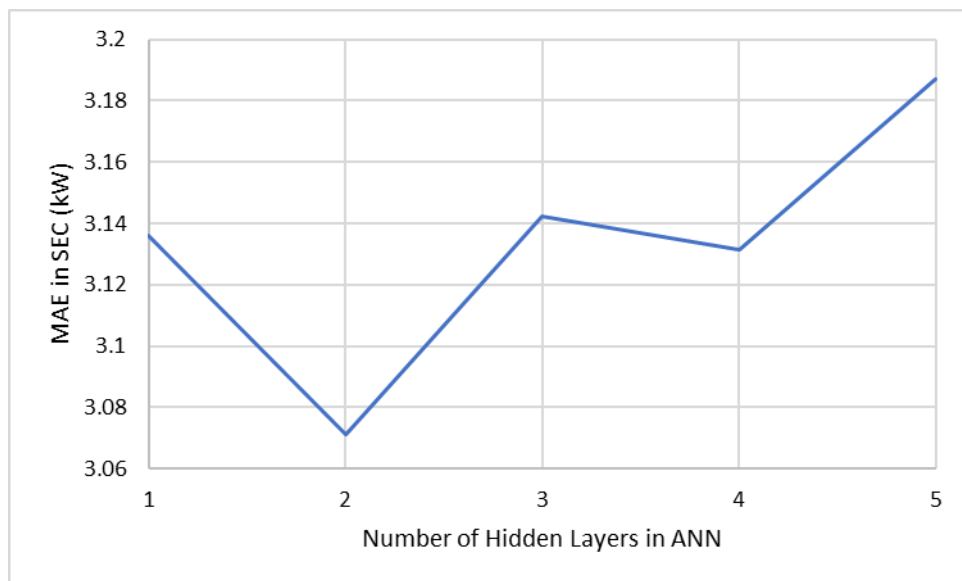


Figure 3.14. Segment MAE by Number of Hidden Layers in ANN

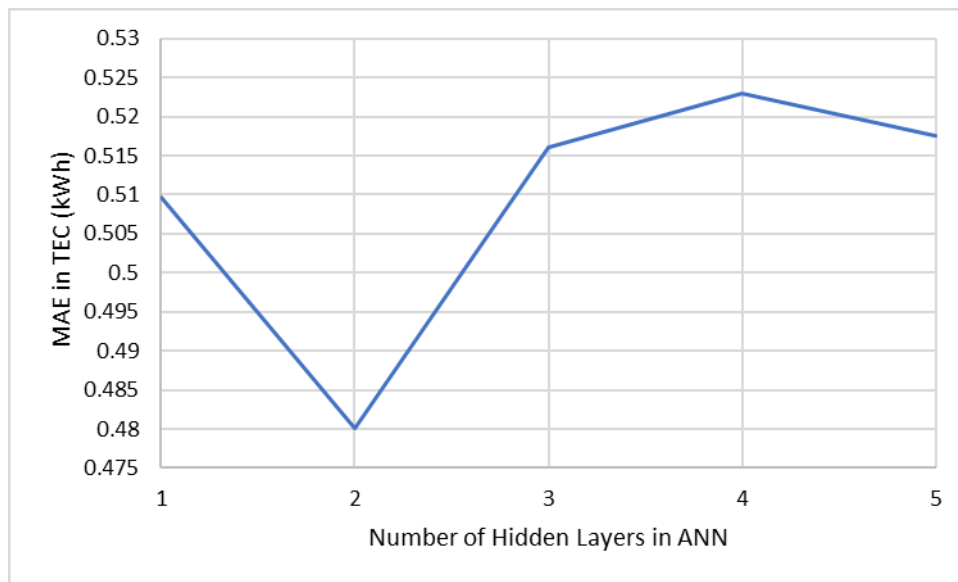


Figure 3.15. Trip MAE by Number of Hidden Layers in ANN

Figures 3.14 and 3.15 show a minimum in MAE when using two hidden layers in the ANN at both the trip and segment level. This is due to overfitting on the training data when more than two layers are used, where the neural network forms assumptions on the data that are not true for a wider dataset. On the other hand, the ANN could not find the complex relationships between variables with only one hidden layer. So, two hidden layers were used for the ANN.

The loss function was set to be MAE because minimising this value is the primary goal of the prediction model. The optimiser was set to the Adam algorithm as introduced in Kingma and Ba (2015), a quick and efficient optimiser popular for ANNs. The Adam algorithm optimises gradient descent, the method in which an ANN updates the weights and biases which lead to a prediction. The algorithm updates weights and biases using exponential moving averages of the gradient and squared gradient, maintaining a separate learning rate for each parameter which allows the weights and biases to scale according to the parameter (Kingma and Ba, 2015). This is also the recommended optimiser in the Keras library from which the neural network framework was taken (Chollet, 2015). The recommended learning rate of 0.001 for the Adam algorithm was selected for the model.

The validation split, which differs from the original test-train split of the data, is a separate branch of the training data used to ensure the model is not overfitting the data even at the training level. The validation split was set as 10% of the training data to ensure enough data could be used to train the model. The validation data was used to select the final hyperparameter and the number of epochs or iterations used to build the model. Figure 3.16 shows the importance of selecting an appropriate number of epochs used to train the ANN. Figure 3.16 tracks the progress of the neural network modelling the data, with each epoch representing one pass of gradient descent and one update of the network's weights and biases which form the model. The y-axis shows the loss function, in this case the MAE of the model at each epoch for the training and validation data.

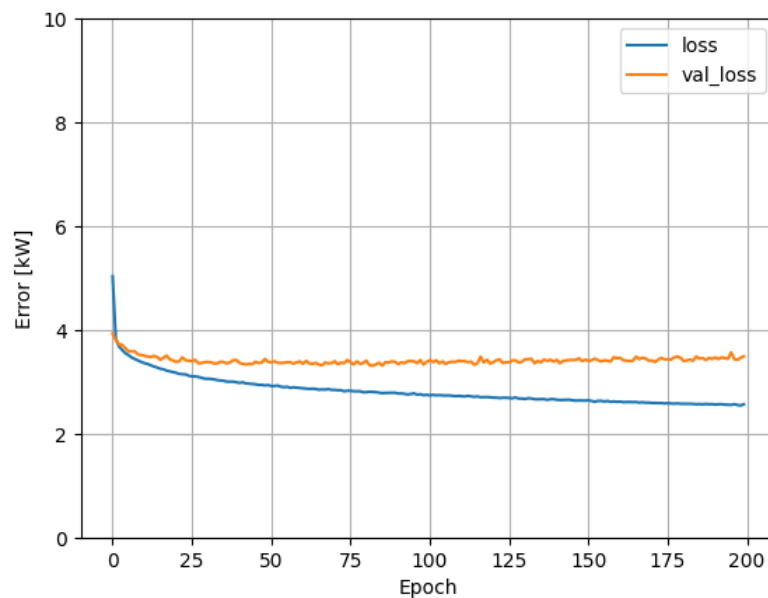


Figure 3.16. Loss and Validation Loss vs. Epoch of ANN

The number of epochs used to train an ANN determines the number of times the model is fed information to update the weights and biases. With a constant learning rate, the error of the model should continue decreasing monotonically with each epoch, as seen in Figure 3.16. The validation loss, which is the MAE on the data of each epoch which was not used to train the model, does eventually increase as the number of epochs increases. This is due to an unwanted phenomenon in ML models called overfitting. Overfitting describes the

model learning patterns in the data too well to the point where the model is not as effective at generalising to other sets of data (Diaz Alvarez, 2014; Fukushima et al., 2018). So, the number of epochs was chosen at the point where training loss was minimised without compromising validation loss, which for this model was at 80 epochs.

The hyperparameters of the XGBoost model are listed as follows: the maximum depth of a tree is seven, the learning rate is 0.1, and the number of trees is 100. The test-train split was the same for XGBoost, LightGBM, and ANN at 20% of the data being used to test the model, which is built on the 80% training data.

3.3.7 Feature Importance

Another insight that can be gained from this model, especially when comparing the trip and segment level, is the quantitative importance of the different features on EC. Many studies have attempted to study the importance of different features on EC, but they do not contain the same combination of parameters used in this paper and do not compare the features at trip and segment levels. In many cases, battery information is included as a feature whereas this model relies only on driving behaviour and the surrounding environment (Blades et al., 2024; Chen et al., 2019; Ullah et al., 2022). In order to quantify the importance of each feature in the model, the permutation feature selection method was adopted. This method randomly shuffles the values of one feature before predicting the EC again and evaluates the decrease in model accuracy with these randomised values (Fafoutellis et al., 2021; Janpoom et al., 2023; Maity and Sarkar, 2024). Thus, the features which cause the greatest increase in error when its values are shuffled are recognised as the most important features of the model. For this study, the increase in MAE was evaluated after shuffling each parameter. Additional information on feature importance can be gained through the results of MLR and by analysing the correlation matrix of the parameters.

3.3.8 Trip Segmentation

To increase the functionality of the data-driven EC model, each trip was split into segments. Trips can be segmented into MTs in various ways (Tong and Hung, 2010). Nouri and Morency (2017) presented various definitions for MTs and tested each on its ability, when MTs are pieced together to form a DC, to be representative of average driving characteristics. The definitions of MT tested include splitting the speed profile between consecutive idle periods or stops, splitting after a certain distance is covered, and splitting on event bins based on acceleration. Nouri and Morency (2017) examined how well each type of MT could represent the average statistics, including average speed and acceleration, in the database when pieced together to create a DC. The definition of MT that gave the least amount of error between the DC and average statistics of the database was found to be an MT with fixed distance.

For the segmented model in this thesis, trips were divided into fixed-distance MTs of 500m instead of 250m suggested by Nouri and Morency (2017). This change serves two purposes: lessening the computation load by half and bringing the distance closer to the average link distance of the VISUM model of 580 meters. In total, there are 269,700 links in the VISUM model of Ireland and Northern Ireland. However, many of those links are local roads which do not have any major impact on traffic in the network. For this reason, only links with a LGV volume of 100 or more cars during the AM peak simulation were examined. The number of links fitting this description is a much smaller 40,274 links. The average length of these links is 580 meters, which closely aligns with segmenting trips every 500m. Because the accepted definition of a MT in the modern literature is the speed profile between consecutive idle periods (Mayakuntla and Verma, 2018), the MTs in this thesis will be referred to as driving segments or simply segments to differentiate the methodology. Some remainder segments at the end of trips are less than 500m, but only 499 of the 22,916 segments are less than 400m.

Straub et al. (2022) perform a similar division of EC calculations. By discretising road network links into sublinks of a maximum distance, the error of EC predictions was moderately decreased. Still, the authors warn that the error decrease may not be worth the computational increase, so sublinks or segments should not be too short. The division of trips, which will in all cases be much longer than the length of a road link, should decrease the error even further. De Cauwer et al. (2017) likewise segmented the EC into links based on a computer network as an important step to increase the variety in aggregated data. De Cauwer et al. (2017) recommends against dividing trips into fixed distance or duration MTs because road characteristics cannot be adequately represented in that way, but this thesis has included road characteristics in each segment to avoid that issue. Additionally, dividing the trips by road link was infeasible for the scale of the data collected as well as for the model's intended use across the entire country of Ireland. Ping et al. (2019) segment trips by location. Maity and Sarkar (2024) divide trips into MTs of random duration to increase the amount of data to build an EC model using ANNs. Chen et al. (2019) divided trips into kinematic segments based on the frequency of data collection.

The definition of MTs being between two idle periods is problematic to use in this research for a variety of reasons. Chiefly among these reasons is that the distance and duration across MTs can vary greatly across the database. For instance, an urban trip will be made of several MTs of a few seconds which would be modelled individually whereas an interurban trip may be or include a single MT covering hours of driving. In the latter case, only one value of average speed, acceleration, and slope would be considered. The effects of varying driving behaviour could not be adequately observed for longer MTs. Petersen and Sax (2022) demonstrated that division into 500m segments produced features that could better model EC than MTs for this reason. Pan et al. (2023) divided trips into MTs to predict EC of short trips which were summed to predict TEC with some success.

3.4 Results

3.4.1 Trip-Level Model Performance

The first result to evaluate was the performance of the ML models to select the TEC model to be used in future chapters. As mentioned previously, the models were evaluated using MAE, RMSE, and R^2 to compare the performance and goodness of fit of each model. The main goal of the model is to minimise the MAE and therefore have predictions as close to the true EC as possible. RMSE supplements this measure by punishing large errors. The results in Table 3.7 are averaged across five test-train splits and show results for the 20% test data only.

Table 3.7. Performance of ML Models of TEC

Model	RMSE (kWh)	MAE (kWh)	R^2
ANN	0.835	0.480	0.964
XGBoost	1.238	0.572	0.908
MLR	1.246	0.792	0.901
LGBM	1.783	0.899	0.809

It can be seen in Table 3.7 that ANN performs the best of any model both by minimising the errors and maximising the goodness of fit. XGBoost also performs well at achieving its target of minimising RMSE. However, it appears that XGBoost had a few larger errors, which are punished greater in RMSE. MLR was the next best performing algorithm and appears to fit the data well through the R^2 score, but the linear coefficients were not able to achieve the same prediction power as the nonlinear ANN and XGBoost models. LightGBM was the worst performing model across all three measures, possibly due to prioritising computational efficiency.

3.4.2 Trip-Level Feature Importance

As the best-performing model, ANN was the model on which permutation feature importance was implemented. Figure 3.17 shows the most important features from the trip-level model.

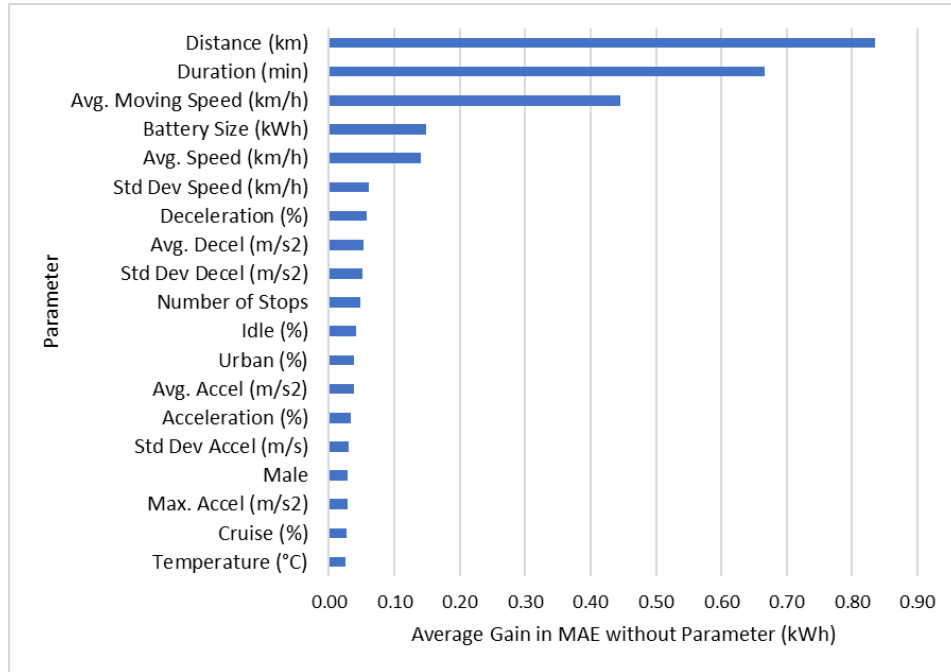


Figure 3.17. Feature Importance of Trip-Level Energy Consumption within the ANN Model

From Figure 3.17, it can be seen that the most important features are the distance and duration of the trip. This might have been expected since the database is extremely diverse, with trips from as short as 0.38km to as long as 217.3km, as seen in Table 3.8.

Table 3.8. Distance, Duration, and Energy Consumption of Trips

	Duration (min)	Distance (km)	EC (kWh)
Min	1	0.38	0.01
Max	179	217.3	35.57
Mean	24.26	17.16	2.73
Std Dev	21.80	23.21	4.01

With this dataset, it is obvious that a trip less than 1km will consume little energy. However, when the model is built on these data and relies on distance as the most important feature, it can neglect the driving behaviour which causes trips of similar distance to vary in EC. Other features such as the average speed of the trip and the vehicle's battery size do have a role in the model, but most of the model's prediction power is drawn from basic trip information.

One shortcoming of permutation feature importance is its inability to convey information about the magnitude and direction of the relationship between each parameter and TEC. Therefore, this result is supplemented by a correlation analysis between each variable and TEC and an analysis of the coefficients of the MLR model. The Pearson's correlation coefficient between each parameter and TEC are listed below in Table 3.9. The correlations are listed by magnitude, and negative correlations are shaded grey. Negative correlations mean as the parameter increases, TEC generally decreases; deceleration was recorded as negative values, meaning that through Table 3.9 it can be interpreted that steeper decelerations with higher magnitudes correlate with less TEC.

Table 3.9. Correlation between Parameters and Trip-Level Energy Consumption

Parameter	Correlation Coefficient with TEC
Distance (km)	0.972
Duration (Min)	0.791
Avg. Moving Speed (km/h)	0.720
Avg. Speed (km/h)	0.682
Cruise (%)	0.552
Std Dev Speed (km/h)	0.532
Urban (%)	-0.512
Motorway (%)	0.467
Accel (%)	-0.394
Decel (%)	-0.389
Max Accel (m/s^2)	0.367
Max Decel (m/s^2)	-0.360
Avg Accel (m/s^2)	-0.348
Avg Decel (m/s^2)	0.306
Unclassified Road (%)	0.217
Secondary Road (%)	-0.207
Residential Road (%)	-0.197
Male	0.188
Rain	0.182
Idle (%)	-0.174
Battery Size (kWh)	0.163
Vehicle Age (yrs)	-0.143
Tertiary Road (%)	-0.121
Number of Stops	0.117
Primary Road (%)	0.084
Std Dev Accel (m/s^2)	-0.082
Wind Factor (km/h)	-0.066
Std Dev Decel (m/s^2)	-0.061
Wind Speed (km/h)	0.051
Gradient (%)	0.031
Temperature ($^{\circ}\text{C}$)	-0.021

Several interesting results can be seen through Table 3.9. First, it should be noted that distance is almost perfectly linearly correlated with TEC. Some parameters which have strong presence in the literature, such as gradient and temperature, show little relation with TEC. There are other strong predictors too: higher velocities and cruise drive mode percentages are strongly correlated with increased TEC. More urban trips show less TEC, and trips that take place on motorways have higher TEC. To view in a different lens, all of

these other strong variables are simply describing trip types, which is still the major predictor of TEC. Longer trips also exhibit higher velocities, are less urban, have higher cruise percentages, and take place more on motorways. Essentially, the model is only taking into account the basic trip information and does not value driving behaviour.

Another way to see this result is through the MLR model, which clearly displays the magnitude and direction of each variable's impact on TEC. Table 3.10 shows the value of coefficients and confidence interval for each variable in the TEC model. Table 3.10 also shows the standard error of each coefficient and the p-value, which shows the probability of obtaining such a result if the parameter were not linearly proportional to TEC. For the results seen here, parameters that were not statistically significant to a 95% confidence interval were removed iteratively until all parameters were found to be statistically significant. Again, negative effects are shaded grey in Table 3.10.

Table 3.10. MLR Results for TEC

	Coefficient	Std Err	P Value	95% Confidence Interval
Constant	-1.41	0.30	0.00	(-2.00, -0.83)
Duration (Min)	-10.09	1.27	0.00	(-12.58, -7.60)
Distance (km)	40.35	1.45	0.00	(37.51, 43.20)
Number of Stops	4.26	0.75	0.00	(2.80, 5.73)
Avg. Moving Speed (km/h)	3.49	0.71	0.00	(2.09, 4.89)
Std Dev Speed (km/h)	-1.20	0.49	0.02	(-2.17, -0.23)
Avg. Accel (m/s^2)	1.71	0.65	0.01	(0.43, 2.99)
Max. Accel (m/s^2)	2.06	0.65	0.00	(0.78, 3.34)
Std Dev Accel (m/s^2)	-2.62	1.01	0.01	(-4.61, -0.63)
Primary (%)	-1.10	0.47	0.02	(-2.02, -0.19)
Battery Size (kWh)	0.90	0.22	0.00	(0.47, 1.32)

Table 3.10 again reinforces the idea that distance and duration are the driving force behind the accuracy of TEC models. The remainder of the parameters, an order of magnitude less important to the model, account for much smaller variations in EC across trips.

3.4.3 TEC Reliance on Distance and Model Alternatives

The extreme reliance of trip-level EC modelling on distance has not been thoroughly explored in the extensive EC modelling literature, but several examples were found

explicitly or implicitly. In Li et al. (2021), adding variables to a bus EC model other than distance only increased the model accuracy by 2.5% for trips over 20km. Modelling short trips was benefitted more by using more variables, but accuracy was still only increased by 6%. A random forest model with only distance achieved an R^2 of 0.859, compared to 0.926 with route characteristics, traffic conditions, and environmental conditions included. Similarly, the RMSE drops only slightly from 3.52 kWh to 3.28 kWh when other variables are included (Li et al., 2021). A study of 68 EVs in Japan mirrored this near-linear reliance of EC on trip distance. In a correlation matrix, the correlation between trip-level EC in kWh and trip distance was recorded to be 0.89, far higher than other well-noted predictor variables such as average speed, which only showed a correlation of 0.3 with EC (Ullah et al., 2022).

The heavy dependence on distance in Ullah et al. (2022) leads to an interesting result, namely that the model fits the data well when measured using an R^2 value but contains high error values. For instance, the best-fitting model gives an R^2 of 0.98, meaning almost all of the variance in EC can be explained by the dependent variables, chiefly distance. However, even this best model is left with a MAE of 13.93 kWh and an RMSE of 16.34 kWh with the mean of the target variable being 0.759 kWh (Ullah et al., 2022). Blades et al. (2024) observes that recent studies have concluded that driver behaviour has been discovered to be one of the least useful predictors of EC for electric buses, but the results may be different for EVs. A physics-based model by Donkers et al. (2020), using driving profiles through simulation in VISSIM, was found to have a MAPE of 7.8% for trips less than 5 km and 3.4% on trips between 5-20 km. This implies that differences in EC on shorter trips are more difficult to model, whereas variations in longer trips may even out as distance increases. Maity and Sarkar (2024) also demonstrates the high correlation between distance and EC. Additionally, it is shown that average speed of a trip is not a good predictor of ECR for the data of that study. Along with these results from the literature, the direct relationship between TEC and distance for the database of this thesis can be seen in Figure 3.18.

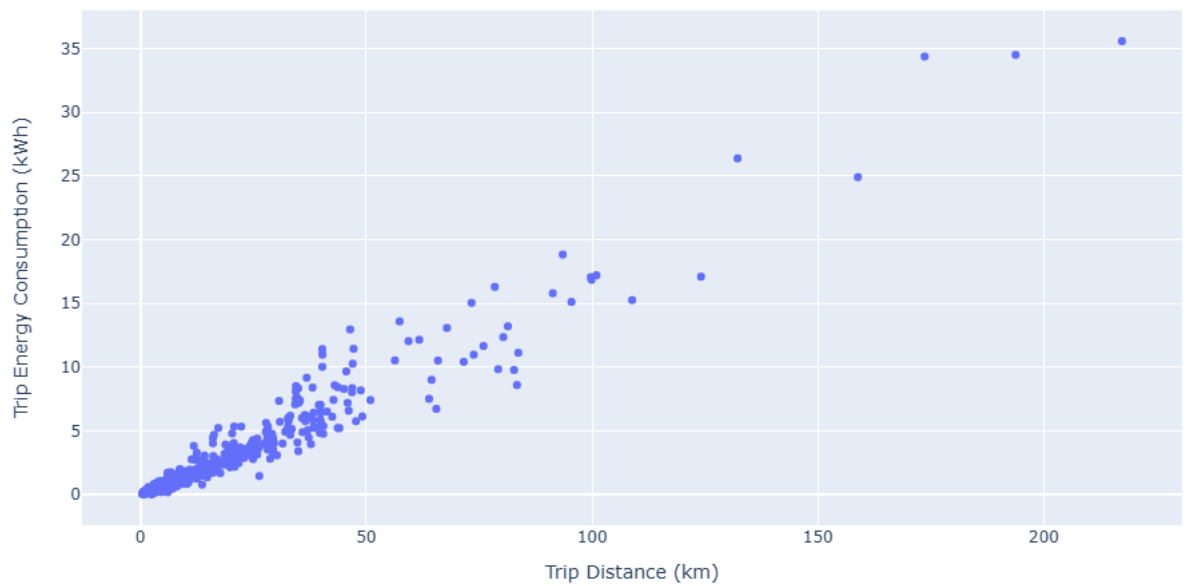


Figure 3.18. Trip Energy Consumption vs. Trip Distance of All Trips

The near-linear relationship between distance and TEC is presented in Figure 3.18. The two parameters have a Pearson's correlation coefficient of 0.972 in the recorded data. Shorter trips especially follow a linear path while slightly more variability occurs for longer trips. The heavy reliance of the TEC model on distance leaves no room for the model to analyse driving behaviour on simulated trips and find variations in TEC from driving behaviour and other parameters. The reliance on distance was also thought to bring about higher errors for longer trips, which is when estimated EC is even more important to lessen range anxiety. To visualise this, Figure 3.19 shows the predicted and actual values of TEC for the test data. Table 3.11 also analyses the difference between MAE in shorter and longer trips quantitatively.

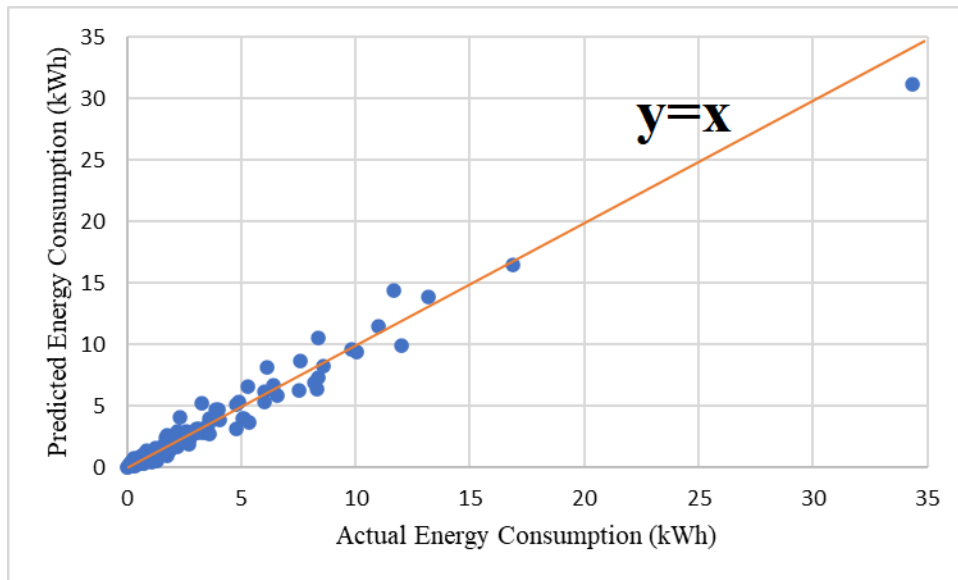


Figure 3.19. Predicted vs. Actual Energy Consumption Using a Macroscopic Model

An orange trendline of $y = x$ was added to Figure 3.19 to represent perfect predictions and better visualise errors as the vertical distance between each trip (blue dot) and the trendline. From Figure 3.19, it can be seen that a large amount of the data is clustered around 1 kWh. These are short trips in the database, and the error of these trips appear to be small. As trips get longer, the errors between predicted and actual TEC get larger in a similar manner to the variables in Figure 3.18. This trend was measured quantitatively by comparing the MAE of short trips to medium and long trips. Here trips were designated as short if the TEC was less than 2 kWh. Medium and long-range trips are grouped together as any trip with TEC greater than 2 kWh. 2 kWh was chosen as the point to split short trips because all trips below this threshold display similar characteristics. The errors are compared in Table 3.11. Table 3.11 shows errors for the test data which were not used to build the ANN during training.

Table 3.11. Error Comparison between Short and Long Trips

	Number of Trips	MAE (kWh)
Trip EC < 2 kWh	82	0.218
Trip EC ≥ 2 kWh	49	0.861

The results of Table 3.11 have a large impact on how the TEC model can be interpreted and give a warning to other EC models relying on similar measurements. The model, relying on distance and duration as features, produces less errors for short trips. The MAE on trips with TEC greater than 2kWh is more than four times greater than the MAE of short trips. The accuracy on small trips skews the overall model performance evaluations to appear better than they would be for longer trips where the information is more valuable. Without this deeper analysis, the performance of the model appeared to be at a good level of accuracy with an MAE of 0.48. Accurately predicting TEC of longer trips is necessary to reduce range anxiety of potential and current EV drivers (Ahmed et al., 2022; Mei et al., 2023; Rauh et al., 2015). Estimating TEC for small trips has a place, but the difference between high and low EC for short trips is not large whereas big differences can occur for longer trips.

So, alternatives were developed to add importance to driving behaviour variables such as speed and acceleration, and parameters other than distance without taking away from the model performance. Three methods were considered: removing distance and duration from the model, modelling ECR instead of EC, and breaking trips into smaller segments to model segment-level EC.

3.4.4 Model without Distance and ECR Models

The first method, removing distance and duration from the model, was a simple process. The ANN model was built on the same database with just those two variables excluded. The second method to model ECR was similarly accomplished by changing the target variable of the ANN model. Because ECR is calculated as the amount of power over a given distance, distance and duration were also left out of the ECR model. Two target variables were developed to measure ECR: the average battery power (ABP) of a trip in kilowatts and the TEC normalised by distance in kWh/100km (Qi et al., 2018; Yang et al., 2023). ABP is not a common measurement in the literature, but it is equivalent to a measure

of ECR normalised by time instead of distance. To calculate EC in kilowatt-hours, ECR in kilowatts is multiplied by the duration of the trip in hours. This measurement is uniquely suited for the uses of this model because it is independent of distance, fluctuates more greatly with driving behaviour, and can later be multiplied by trip duration to form a simple calculation for gathering TEC.

Table 3.12 displays the results for these secondary models along with the original ANN model, denoted by Trip EC (kWh), with distance and duration included.

Table 3.12. Alternative Models for Trip-Level Energy Consumption

Model	RMSE (Unit of Model)	MAE (Unit of Model)	R ²
Trip EC (kWh)	0.835	0.480	0.964
Trip EC (kWh) No Distance or Duration	1.929	0.900	0.771
Trip ECR - Average Battery Power (kW)	1.567	1.071	0.848
Trip ECR (kWh/100km)	4.415	3.297	0.180

The first result of note from Table 3.12 is that by only removing distance and duration from the model, MAE and RMSE are doubled. It was thought that a model built without these parameters would use other variables to fill in the gaps, but the information from other parameters cannot make up for the loss of explained variance through distance and duration. The ABP model at the trip level also falls well short of the original model in terms of error. The ECR model with EC normalised by distance has a much higher MAE of 3.297 kWh/100km and a goodness of fit of 0.180 indicating the model does not fit the data well. Although the average errors are higher in the new models, the models' target variables are different. If the target variable has more variance, it makes any model harder to fit. However, the standard deviations of all three target variables are similar. The standard deviation of EC was found to be 4.00 kWh, while the standard deviation of ECR was 3.91 kW and 4.84 kWh/100km for the respective ECR models. To see if the ABP model incorporates a more diverse list of variables, permutation feature importance was implemented on that model as seen in Figure 3.20.

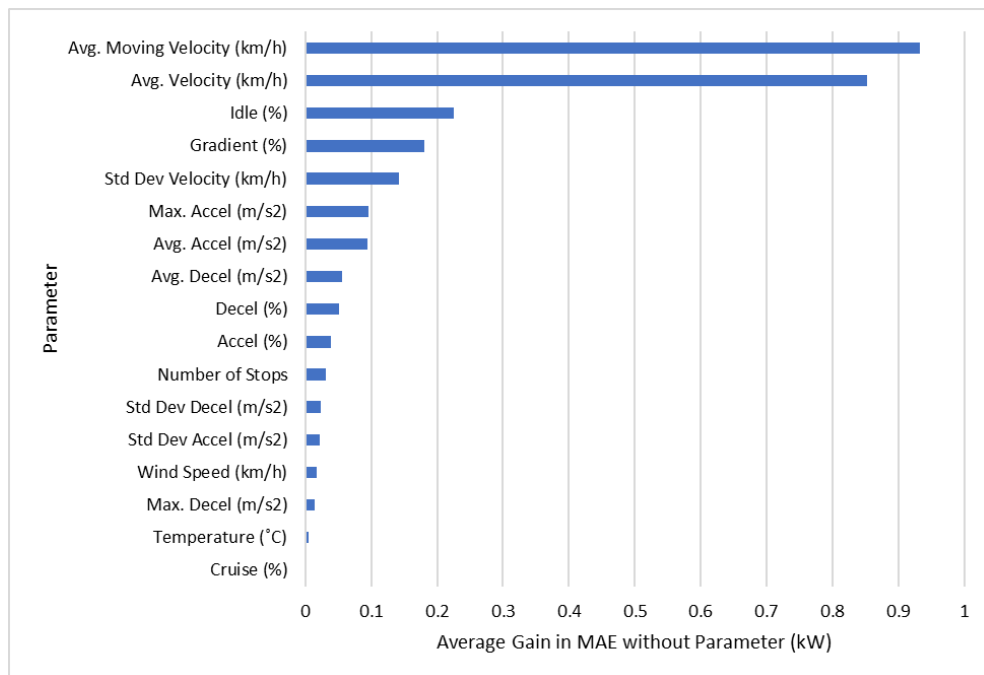


Figure 3.20. Feature Importance of Trip-Level Average Battery Power Model

Again, average speed and average moving speed top the list of important features in the battery power model as seen in Figure 3.20. Despite this, there are some promising signs that this model is at least slightly more affected by driving behaviour and route characteristics. For instance, maximum acceleration affects the model marginally, which means some level of driver aggression could be modelled. In addition, the average gradient of the trip is an important factor of this model, while it was almost completely absent from the original TEC model.

The model for ECR as a standard measure of kWh/100km falls well short of reliable prediction levels (see Table 3.12). The relative success of modelling ABP at trip level along with the failure of a standard ECR model at trip level were considered while forming the third option to divide trips into segments. It was hypothesised that the types of features included in the model can vary greatly throughout the trip but only have one value represented in a trip-level model. For example, a trip that climbs and descends many gradient changes may end up with an average gradient near zero. This trip would provide the same information on gradient to the model as a trip that travels on a flat gradient the entire trip.

While the ANN may be able to find relationships across parameters to account for some of

these differences, it is also necessary to make sure the input parameters to the model contain high quality information for the prediction of EC. For this reason, trips were discretised into 500m segments as described in the methodology section. The results of this model are discussed in the following section. Following the example outlined above, this would ensure that uphill segments are accounted for at the segment level and the model can utilise the variation of parameters during a trip.

3.4.5 Segment-Level Energy Consumption Model

Because the segment level model is so different to trip level model, the various ML models tested at the trip-level needed to be compared again to see which could best model SEC. In this chapter, SEC refers to the ABP of a segment in kilowatts. For the segment-level models, the predictor variables were not changed. They were simply calculated for each 500m segment instead of for the whole trip. Similar to the previous section, distance and duration of each segment were not included as parameters. Not all segments were exactly 500m as the final segment of each trip was composed of any remaining distance after the penultimate segment. To exclude the shortest segments which might disrupt the model, only segments over 100 m were included in the model. In total, 22,795 segments were included in the modelling process, with a mean ABP of 11.35kW and a standard deviation of 11.73kW. Table 3.13 presents the performance of the ML models.

Table 3.13. Performance of ML Models of SEC

Model	RMSE (kW)	MAE (kW)	R2
ANN	4.847	3.071	0.829
XGBoost	5.080	3.226	0.811
LightGBM	5.230	3.449	0.799
MLR	7.098	4.836	0.633

Like the TEC results, ANN was found to be the best performing model to predict SEC through all measurements. In the case of SEC, the decision tree algorithms XGBoost and LightGBM gave results much more similar to the ANN model. This result differs from Ullah et al. (2022) where XGBoost and LightGBM outperformed ANN. In both cases, MLR was

the worst-performing algorithm. The bad performance of MLR at the segment level can be explained by the input parameters. As explained in the methodology section, MLR cannot model complex, nonlinear relationships between parameters and EC. The raw amount of data present in the segment-level model, as seen in Figure 3.21, means a model needs to understand the complex relationships across input variables to fit the data well. Compared to the same parameter at the trip level seen in Figure 3.13, there is an increase in variation and nonlinearity seen in Figure 3.21. The correlation between average moving speed and EC drops from 0.720 at the trip level to 0.583 at the segment level.

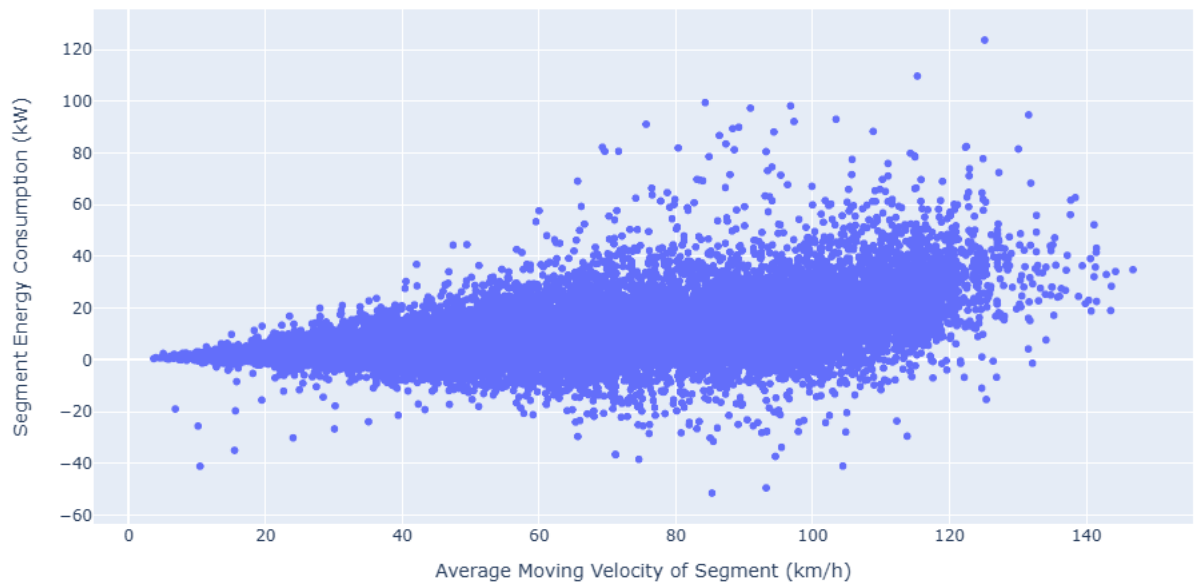


Figure 3.21. Average Moving Speed vs. Segment Energy Consumption of All Segments

Based on the results of Table 3.13, ANN was selected to be the SEC model. This model, which evaluates the ABP of segments, is referred to as both the SEC model and the ABP model. Compared to the trip level model, the MAE of 3.071 kW is much higher than 0.480 kWh. Additionally, the R^2 falls from 0.964 at the trip level to 0.829 at the segment level, indicating ANN fits the model worse at the segment level. However, this can be explained by the much greater variation in SEC compared to TEC. The standard variation of SEC is 11.73, nearly triple the standard deviation of TEC at 4.00. The evaluation of SEC

as a predictor for TEC is explored in the following section. First though, to see which parameters have the largest effect on SEC, permutation feature importance was implemented another time on this segment-level model.

The feature importance analysis of SEC in Figure 3.22 displays a completely different trend to the trip-level analyses seen in Figures 3.17 and 3.20.

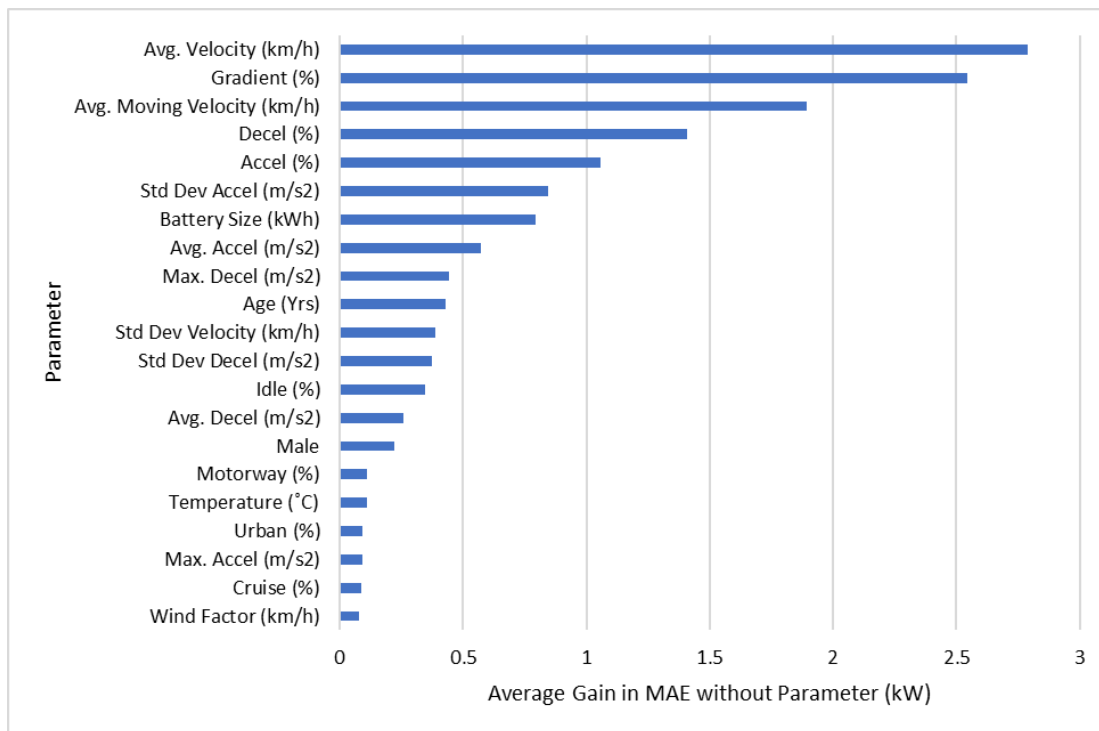


Figure 3.22. Feature Importance of Segment-Level Energy Consumption

Model

Instead of two or three parameters being the base for the entire model with all other parameters contributing little information, the SEC model takes a large variety of parameters into account. While speed remains an important parameter to the model, several previously unimportant parameters are now important to the SEC model. For example, the average gradient of the trip, which provided little information at the trip level, is the second most important feature at the segment level. This confirms the segment model's ability to obtain more information from parameters that are limited to just one average value at the trip level. Other parameters that exemplify driving behaviour, such as percentage time spent accelerating and decelerating, are also of importance to the SEC model. This result confirms

that, if the SEC model performs well at predicting TEC, the segmented model would be suitable for future uses of the model to examine the effects of driving behaviour and route choice on TEC. The Pearson's correlation coefficients between each variable and SEC are examined below in Table 3.14 to determine how each variable affects the model. As with previous tables, negative effects are shaded grey in Table 3.14.

Table 3.14. Correlation between Parameters and Segment-Level EC

Parameter	Correlation Coefficient with SEC
Avg. Moving Velocity (km/h)	0.583
Avg. Velocity (km/h)	0.581
Deceleration (%)	-0.482
Max. Decel (m/s ²)	0.453
Std Dev Decel (m/s ²)	-0.446
Avg Decel (m/s ²)	0.389
Urban (%)	-0.368
Cruise (%)	0.353
Std Dev Velocity (km/h)	-0.352
Gradient (%)	0.287
Max. Accel (m/s ²)	-0.258
Motorway (%)	0.244
Std Dev Accel (m/s ²)	-0.244
Idle (%)	-0.235
Number of Stops	-0.22
Avg. Accel (m/s ²)	-0.192
Secondary (%)	-0.189
Battery Size (kWh)	0.164
Vehicle Age (yrs)	-0.147
Male	0.122
Residential (%)	-0.075
Tertiary (%)	-0.05
Temperature (°C)	-0.042
Accel (%)	0.034
Primary (%)	0.033
Unclassified (%)	0.019
Rain	-0.015
Wind Factor (km/h)	-0.007
Wind Speed (km/h)	0.003

Interestingly, acceleration parameters such as average acceleration, which are known to increase EC as the values increase (Lee et al., 2023), show the opposite relationship with

SEC here. This is likely due to the fact that the segments with the highest SEC are high-speed segments on interurban trips, which are unlikely to have large changes in speed and therefore high accelerations. Also of note is the fact that gradient, whilst the second most important feature in the model, has only the tenth highest correlation with SEC. This demonstrates the importance of using advanced ML algorithms which can take into account relationships between parameters and the nonlinearities of data especially when large datasets are used. Because of the high level of variance and nonlinearity, MLR coefficients are not helpful to understand the effects of these parameters on SEC.

3.4.6 Final Model Selection: Trip Energy Consumption as a Summation of Segments

To incorporate the segment-level EC predictions into the trip-level model, an SEC prediction term was developed. The SEC prediction term adds weights to each segment of a trip based on the percentage of trip duration of that segment to find the ABP over the entire trip. Then, this weighted sum is multiplied by the trip duration in hours to get a value of TEC in kilowatt-hours. This equation can be seen below in Equation 3.5:

$$TEC = t * \sum_{i=1}^n SEC_n * \frac{t_n}{t} \quad (3.5)$$

where TEC is the trip-level energy consumption (kWh), SEC_n is the energy consumption rate (kW) of segment n, t is the duration of trip in hours, and t_n is the duration of segment n in hours. This TEC prediction term was added to the original model as an additional parameter. The prediction term scales the weighted segment-level ECRs in kW by duration of the trip to find TEC. This ensures that trip duration is an essential parameter to the trip model, but the variations throughout a trip are captured at the segment level and included in the prediction term as well. This allows the model to capture variation along a trip instead of relying on distance as the most important parameter. Permutation feature importance, seen in Figure 3.23, showed the new prediction term based on SEC was the most important feature, greatly surpassing distance.

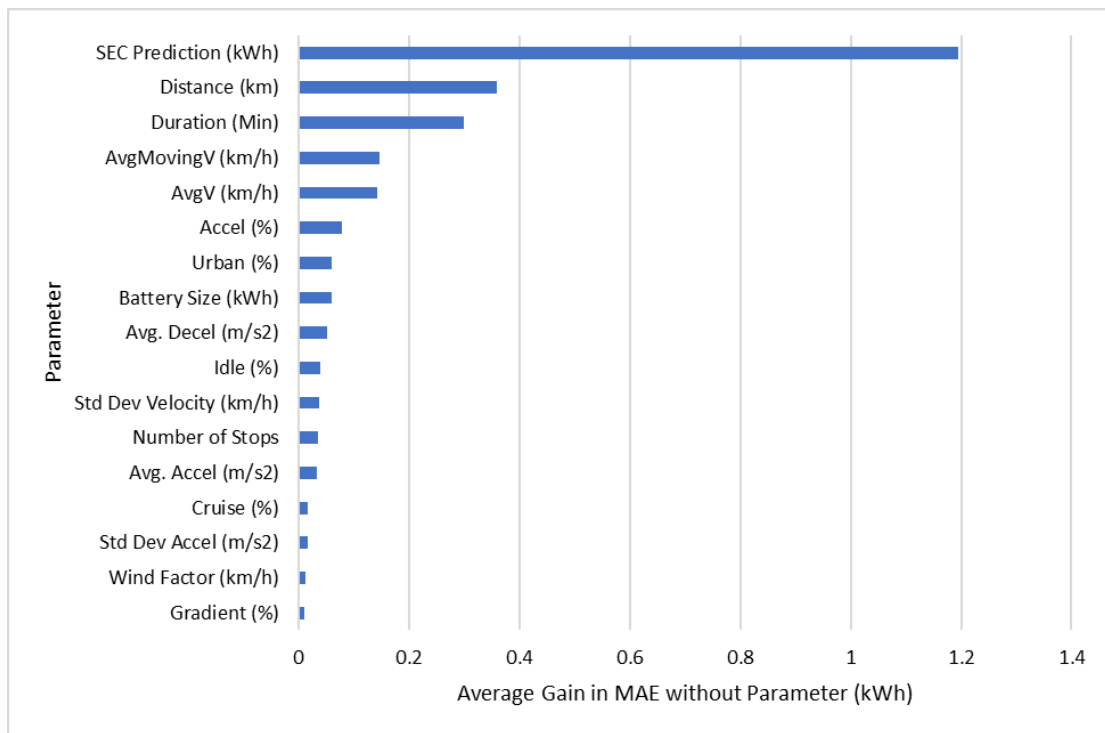


Figure 3.23. Feature Importance of Trip-Level EC Model with Segment Predictions

It was supposed that, even though the SEC model predictions are based on the same parameters used in the TEC model without distance or duration, some parameters would be captured better at smaller distances. For example, the gradient can change drastically over a trip as the vehicle travels up and down slopes. However, at the trip level, only one average value of gradient over the entire trip was considered. By adding the SEC prediction term to the model, these parameters could exhibit their wider range of impacts by being captured at the segment level first. This hypothesis appears to be proven through Figure 3.23, where the effects of many parameters that were adding little information to the trip-level model are utilised to increase the prediction power of the model when incorporated at the segment level. The rigidity of having one value of each parameter per trip also explains why the trip-level ECR (kWh/100km) model performed poorly originally.

Because the SEC predictions for TEC are in themselves a TEC model, it was possible to use Equation 3.5 as a TEC model. The TEC predicted by adding SEC performed better than the ANN model with other parameters included, as can be seen in Table 3.15. The

results of the original TEC model without SEC as a variable are also included for reference in Table 3.15. While other errors are shown only for test data, the measurements for the final model include all trips as the test-train split was performed at the segment level when building the SEC model.

Table 3.15. Comparison of Final Model to Previous Models

Model	RMSE (kWh)	MAE (kWh)	R2
Final Model	0.408	0.225	0.995
ANN with SEC Prediction	0.586	0.370	0.982
ANN without SEC Prediction	0.835	0.480	0.964

As Table 3.15 shows, the benefit of discretising data into segments can greatly improve trip-level modelling performance. Adding SEC predictions to the trip-level model decreased errors, especially large errors as seen in the RMSE, and improved the goodness of fit compared to the original TEC model. By using only the segment-level EC predictions to model TEC, the errors are less than half of the original model and the goodness of fit approaches a perfect value of one. For this reason, a two-step model of an ANN to predict SEC and Equation 3.5 to combine SEC into TEC was selected as the final TEC model. It was also considered to add a constant term to the model in case the SEC predictions were consistently over- or under-estimating TEC. For this, an MLR model was developed for TEC with just the SEC predictions and a constant as the input variables. The MLR model suggested a constant of 0.02, but the confidence interval spanned either side of zero, so a constant was not included in the model.

To see if the final model achieves the goal of minimising errors on medium and long trips, the MAE of short and long trips were separated and shown in Table 3.16.

Table 3.16. Error Comparison between Short and Long Trips of Final Model

	Number of Trips	MAE (kWh)
Trip EC < 2 kWh	412	0.122
Trip EC ≥ 2 kWh	241	0.403

The results of Table 3.16 show a decrease in MAE for both short and long trips compared to the original model shown in Table 3.11. Importantly, the MAE of longer trips was decreased by 53%, even greater than the decrease of MAE for short trips at 44%. This result shows that not only does the model accuracy increase by modelling segments but also the accuracy for TEC of longer trips is specifically increased as the model is not as reliant on distance and duration. The ability of the segment-level model to consider a wider range of energy-influencing parameters increases the overall accuracy of the model, especially for medium- and long-range trips. The accuracy of the final model across all trip lengths is illustrated in Figure 3.24.

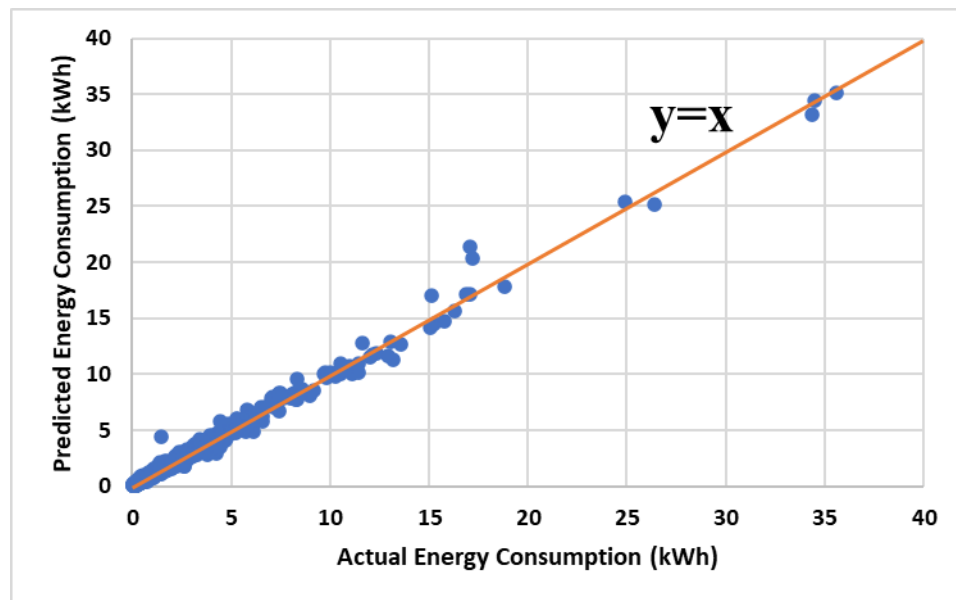


Figure 3.24. Predicted vs. Actual Energy Consumption for Final Model

Compared to Figure 3.19, the errors on Figure 3.24 show a much higher degree of homoskedasticity. The errors of the final model are more evenly distributed over the entire range of trip distances and ECs as described in Tables 3.11 and 3.16. The results shown in Table 3.16 and Figure 3.24 demonstrate that a major issue of the original model, the failure to accurately model TEC on longer trips, was resolved through the adoption of the final model of aggregating segment-level energy consumption.

To visualise the presence and importance of segment-level variation in ECR, Figure 3.25 was developed.

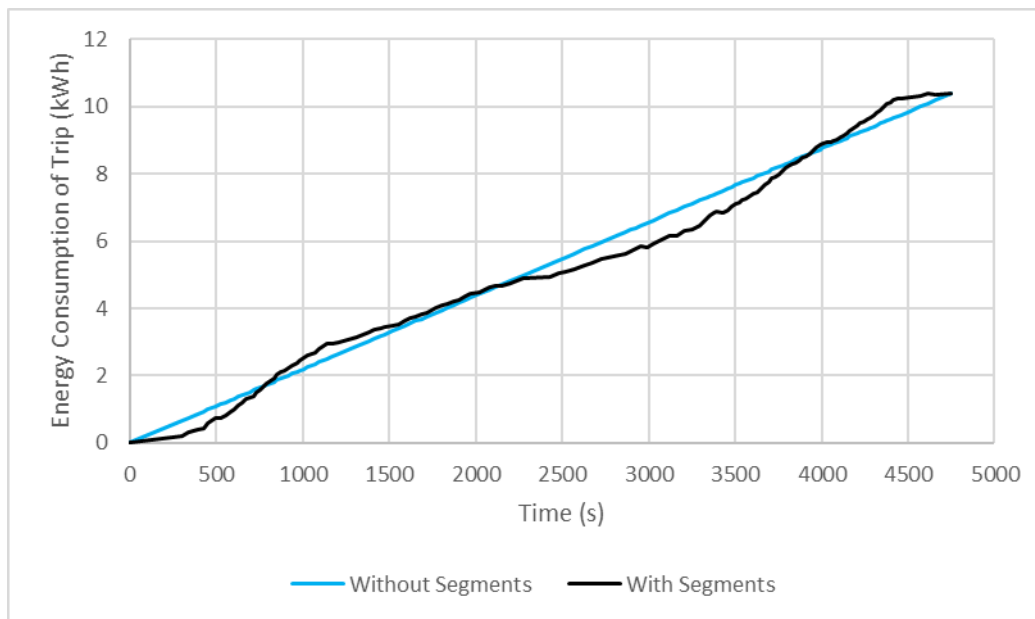


Figure 3.25. Variation of ECR along a Trip

Figure 3.25 shows the application of the final model to an 80-minute trip in the database. The black line shows the variation in ECR from segment to segment compared to a constant ECR for the entire trip shown in blue. Figure 3.25 demonstrates the amount of variation in ECR that occurs throughout a long-distance trip and the importance of considering the influence of shifting parameters through the trip.

3.5 Conclusion

This chapter explored the state of the literature for modelling the energy consumption of electric vehicles and developed a model that combines the EC predictions for each 500m segment of a trip to form the trip-level EC prediction. Various data-driven models to predict EC and ECR at the trip level and segment level were developed before finding the final model with low errors across the entire range of trips. ANNs outperformed other ML models for modelling both SEC and TEC.

In the early stages of modelling, it was found that trip-level EC could be predicted to a reasonable degree by only knowing the trip's distance. Although an interesting and important result, a model relying heavily on distance would not be useful for the later parts of this thesis where driving behaviour will be analysed. Therefore, three different strategies

to lessen the reliance on distance were developed to drive the research further. First, the model was evaluated without distance to see if an accurate prediction could be made without that parameter. Then, the model was run to predict ECR which would be independent of distance, and a distance could be added later to find EC. Finally, the trips were divided into segments of equal length and the model was adapted to measure segment-level ECR. This final method of segmentation resulted in the best prediction model for trip-level EC (TEC). The result is a two-step model where segment-level EC (SEC) is predicted using a neural network, then the SEC of all segments in a trip are added together to form a prediction of TEC.

The summation of all SEC predictions was a better predictor of TEC than a model that added all other trip-level variables, including distance and duration, to the SEC predictions. The results highlight a potential shortcoming in other TEC models, especially data-driven models. Namely, data-driven models such as Li et al. (2021) and Ullah et al. (2022) may have a tendency to rely on distance as a predictor of TEC. The results demonstrated that when a database with a wide range of trips is considered, distance shows a near-linear relationship with TEC. Models built on this data prioritise trip distance as a parameter, which could limit the variation explained by driving behaviour and other parameters for longer trips. It has been shown by dividing the trip into short segments and modelling the ECR of each segment, a more diverse range of parameters affecting EC can be considered. From there, combining the SECs of each trip was shown to result in a better prediction of TEC than using trip-level variables.

The major contributions can be summarised as highlighting the need for data-driven models to discretise trips in order to capture a wider range of parameters to model the changing ECR along a trip. This goes hand-in-hand with the necessity of similar models to ensure that the model is not reliant on the distance and duration of a trip, which can result in high errors for medium- and long-range trips. This chapter also showed that data-driven

models at the mesoscopic level can perform very well in estimating TEC, with the MAE as low as 0.225 kWh across the 653 trips in the database. The mesoscopic level has a unique advantage of being able to model driving behaviour such as acceleration rates and drive mode percentages that a microscopic model cannot. At the same time, the mesoscopic level was shown to reduce error by more than half compared to a macroscopic model using the same parameters.

Chapter 4. Driving Behaviour Effects on Energy Consumption and Sensitivity Analysis

One of the research questions addressed in this thesis is how drivers can impact their EC through their driving behaviours and route choice. In the previous chapter, a model was built to predict the EC using a DC and characteristics of the driver, vehicle, and road environment. Although permutation feature importance was used to gauge how important parameters are to the model's predictions, the effect of changing input parameter values on EC was not considered. It is imperative to policy makers and EV drivers that the most effective strategies to limit EC are found. In this chapter, the focus is placed on how driving behaviour can influence EC so that drivers can know how to extend their vehicle's range and limit the need to charge.

There are many ways to define driving behaviour. Driving behaviour, also called driver behaviour, is generally a term used when categorising drivers' actions while driving a vehicle, commonly including rates of acceleration, braking, and speeding (Ellison et al., 2015). This chapter begins with a literature review on understanding and classifying driving behaviour. Three methods to classify driving behaviour are selected from the literature and tested. The clustering method is common classification tool used to group driving behaviour based on parameter values using a ML clustering algorithm. The limit-based method is derived from the European Union's legislative real driving emission (RDE) test, which places limits on how aggressive or passive DCs are allowed to be during testing based on acceleration and speed values (Regulation 2016/646). The third method is the empirical method, where recorded data is sorted by ABP, and trends are examined across all parameters.

The results of the empirical method encourage a more thorough investigation into the impact of each parameter through a detailed sensitivity analysis (SA). The SA is preceded

by a literature review into state-of-the-art SA techniques and their applications on similar models. Through the literature review, the Sobol' method is selected for the SA in this chapter. Results from the SA help to better understand the effect of parameters on EC across different driving scenarios.

The methods are compared and validated first using the real EV data used to build the EC model in Chapter 3 and again using synthetic driving segments that are made from different cruising speeds and rates of acceleration and deceleration. The goal of this chapter is to find a classification method that produces driving behaviour classes that exhibit different levels of EC. This way, EV drivers and policymakers will have scientific evidence of what behaviours to encourage in order to limit EC and extend EV range. Additionally, it is essential to know how driving behaviour affects EC across different styles of driving, of which knowledge is lacking in the literature. This is achieved by comparing the effect of driving behaviour on EC across different speeds. There are several related but distinct sections of this thesis that are outlined here in order of their placement in this chapter for clarity:

1. Literature review on driving behaviour classification
2. Methodology of clustering, limit-based, and empirical methods
3. Literature review on sensitivity analysis
4. Sensitivity analysis on ABP model
5. Sensitivity analysis on ABP model within speed intervals
6. Analysis of driving behaviour and EC on synthetic segments

4.1 Literature Review on Driving Behaviour

This section covers how various studies have defined and classified driving behaviours, and how these driving behaviours impact EC. The studies described here are included in Appendix D, which helps to easily compare factors like methodology across studies.

4.1.1 Driving Behaviour Definition and EC

It has been established that driving (or driver) behaviour has a noticeable effect on an EV's EC and, by extension, its range. Generally, driving behaviour is classified by levels of aggression. Energy efficient driving behaviour can be referred to as eco-driving. Fafoutellis et al. (2021) define eco-driving as “a set of choices and behaviours adopted by drivers that are connected with an energy-efficient way of using a vehicle.”

Driving behaviour can manifest in many ways both during and before a trip. Fafoutellis et al. (2021) considered the following eco-driving behaviours: route choice, driving style, weight reduction, HVAC use, vehicle maintenance, and vehicle selection. Ma et al. (2024) consolidates this list into driving behaviour, vehicle selection, and route choice. In a study relating driving behaviour to road safety, it was determined that road characteristics and environment have a major effect on driving behaviour (Ellison et al., 2015). Faria et al. (2017) and Faria et al. (2020) also demonstrated how environmental characteristics such as rain and road type have a major effect on driving behaviour. Chu et al. (2023) differentiated between short- and long-term driving behaviour where short-term is mostly guided by the vehicle's environment and long-term by the driver's personality and habits. In this chapter, the focus is placed on driving behaviour and its effect on EC.

The results of Iclodean et al. (2025) demonstrate the complexity of ascribing changes in EC to driving behaviour when categorised into the typical brackets of passive, moderate, and aggressive driving. At higher average speeds of 40-80 km/h, normal driving behaviour consistently consumed the most energy. This is because aggressive driving behaviour includes both sharply accelerating, which increases EC, and sharply decelerating, which can decrease EC due to more energy being recovered through regenerative braking (Iclodean et al., 2025). At lower speeds around 20 km/h, however, aggressive driving behaviour consumed the most energy because the energy recovered by regenerative braking is not as significant at low speeds (Iclodean et al., 2025). Al-Wreikat et al. (2021) also demonstrated

how short DCs may be more influenced by differences in driving behaviour than long DCs. These results highlight the fact that common categories of driving behaviour can demonstrate differences in EC, but the magnitude and direction of these differences are highly situational.

Galvin (2017) showed that EC increases when the EV is driven more aggressively, with EC increasing at higher speeds and more significantly with higher acceleration. Donkers et al. (2020) showed that aggressive drivers consume more energy at higher speeds but save energy at low speeds because the journey duration will be less. The study also demonstrated how speed oscillations increase EC more for aggressive drivers. Additionally, aerodynamic resistance was found to be the dominant force on motorways, while rolling and acceleration forces had a larger effect on urban trips.

EV drivers are known to have different characteristics from ICEV drivers. EV drivers can accelerate faster than ICEV drivers and can choose to decelerate quicker (Xu et al., 2021). Wang (2020), through survey statements from the US, concluded that EV drivers may be calmer and more eco-friendly than ICEV drivers. Rolim et al. (2012) discovered that 75% of early EV adopters in their study reported speeding less after switching to an EV in order to drive more economically. The same study found that most new EV drivers adapt to their vehicle within the first six months. Similarly, Rauh et al. (2015) observed that more experienced EV drivers are less likely to experience range anxiety. Huang et al. (2020) suggest that novice drivers tend to be more cautious and therefore less energy efficient. Zhang et al. (2020B) depicted the variable usage of EV charging based on driver demographics.

One reason EV drivers may drive differently is because of regenerative braking found exclusively in EVs. Regenerative braking has potential to improve energy economy in everyday DCs by recovering mechanical energy while braking for use to power the vehicle (Ruan et al., 2016). Regenerative braking produces energy for the EV at speeds

above a minimum threshold and increases regeneration rate until reaching a maximum at a higher speed threshold. Genikomsakis and Mitrentsis (2017) estimated the minimum speed for regeneration to 5 km/h with maximum regeneration at speeds higher than 17 km/h. Lee et al. (2023) found that regenerated energy is similar during motorway and urban driving because although motorway driving includes less deceleration, the deceleration occurs at higher speeds and generates more energy per second.

Some studies have examined the effect of driver feedback to encourage more energy efficient driving. Berzi et al. (2016) suggested that EVs are designed to communicate with the drivers on factors such as regenerative braking so drivers can change their behaviour advantageously. Caulfield et al. (2014) found that drivers of an ICE vehicle created less emissions when given advice on minimising emissions through driving styles. Meseguer et al. (2017) proposed a phone app to track driving habits and promote energy efficient driving styles. Lu et al. (2019) demonstrates how the smoothing of acceleration and deceleration from smart technologies like adaptive cruise control can have a major beneficial effect on EC.

Studies use the terms driving behaviour and driver behaviour interchangeably (Ellison et al., 2015). Although in specific circumstances driver behaviour may be used to define a subset of overall driving behaviour that can be described by actions or tendencies of a driver while driving (Carmona et al., 2015). For this thesis, the terms are not differentiated, and driving behaviour is used to describe speed- and acceleration- based parameters such as average speed, average acceleration, and time spent accelerating. It can be seen in Appendix D that most studies use the term driving behaviour.

4.1.2 Driving Behaviour Identification and Categorisation

Fafoutellis et al. (2021) outlined four steps for driving behaviour analysis: data collection, understanding the data, modelling, and policy/decision making. These steps seem easy to follow, but there are no standard methods for dividing driving behaviour into groups

or classifying a driver into a behavioural group. This section reviews the methods previously used in the literature for both steps. Ma et al. (2024) identified clustering, classification, and regression as common ML techniques to predict energy efficient driving, with each method having a distinct use case.

Clustering is a method commonly used in the literature to group driving behaviour (Wu et al., 2023). Yu et al. (2012) provided an early example of driving pattern recognition based on EV driving. This study attempted to predict a trip's energy usage and therefore range based on gradient and traffic clusters formed by mountain and fuzzy c-means clustering (Yu et al., 2012). Hu et al. (2022) characterise driving behaviour from characteristic parameters (CPs) of 26 DCs into aggressive, standard, and cautious after using PCA to reduce the dimensions of the CPs and K-means clustering to group the CPs. Liu et al. (2022) used the same ML techniques to identify driving conditions before using a neural network to predict energy consumption. Ping et al. (2019) used different ML clustering methods on speed and acceleration to group driving behaviour, but there was no obvious division between the three resulting groups. Still, the groups exhibited distinct EC patterns due to their driving properties.

Different studies measure and classify driver aggressiveness through other means. The various methods can be found in Appendix D, but several important studies are presented here in more detail. Bar et al. (2011) divided driving behaviour into 5 categories (aggressive, anxious, economical, keen, and sedate) based mainly on how driving speed relates to advised speed as well as acceleration speeds around stop signs and bends. Aggressive and keen drivers behaved similarly as did economical and sedate drivers, indicating division into three driving behaviour categories may be more suitable than five. Yang et al. (2018) used acceleration and throttle opening to divide driving behaviour into six levels of aggressiveness, divided by standard deviations of a normal distribution. Sun et al. (2017) developed a ML model to both classify driving behaviour and assign drivers to

groups based on acceleration and speed. Cheng et al. (2018) represent the scale of driver aggressiveness by acceleration, pedal angle, and vehicle speed. Pedal angle was able to be monitored in Cheng et al. (2018) because an original testing rig was developed, and pedal angle is measured as the angle between the initial pedal position and the position while driving. Huang et al. (2020) investigated the effects of throttle and brake pedal frequency and intensity on EC. Frequency had a greater effect on EC, and the combination of high intensity and low frequency was found to be the most energy efficient. Wu et al. (2023) used SVMs to characterise driving behaviour at the MT level into aggressive, normal, and conservative styles. Before dividing the dataset, locally linear embedding was implemented to reduce the dimensionality of several kinematic CPs from speed, acceleration, and jerk.

Zhang et al. (2020A) divided driving behaviour into four categories based on high or low acceleration combined with high or low deceleration, with the 95th and 5th percentile values used to categorise high and low values. The categories with high acceleration showed higher EC, and both categories with low acceleration exhibited similar average EC. Al-Wreikat et al. (2021) used a measurement called relative positive acceleration to determine the aggressiveness of a driver during a trip. Xing et al. (2020) found that drivers with higher EC exhibited the following driving behaviours: high mean speed, high mean acceleration, high mean jerk, longer periods of acceleration compared to constant speed, longer space headways and divided drivers into three categories based on this data. Chen et al. (2019) invoked a wide range of parameters to describe driving behaviour including percent time of various drive modes (acceleration, decelerating, and cruising).

Still other studies use different methods to classify driving behaviour. Wu et al. (2015) first divided trips by purpose before analysing driving behaviour. Hong et al. (2021) used a dynamic time warping algorithm to group together similar segments of recorded trip data. Lai et al. (2023) used ANNs to find driver preferences using empirical data. The study found major differences in average EC between drivers operating similar EVs.

Another step of driving behaviour analysis is modelling, which can also vary between studies. Yavasoglu et al. (2019) simulated different driving behaviours by scaling the duration of a DC to increase or decrease acceleration. Xu et al. (2021) developed an EV-specific car-following model for microsimulation including different acceleration and following behaviour for timid and aggressive drivers. Jiménez et al. (2018) modelled EC of aggressive drivers by increasing the abruptness of turns, acceleration, and deceleration as recorded by smartphone inertial data. Adding these driving behaviour parameters increased the model accuracy by over 50%. Nan et al. (2022) incorporated driving behaviour into an EC model through the parameters of speed, acceleration, acceleration variance, and position of the brake and acceleration pedals.

A review on driving behaviour detection using ML techniques revealed that many parameters involve a relationship to another vehicle or the road environment such as time headway to the preceding vehicle and number of lane changes. However, some parameters such as speed, acceleration, and jerk can contribute to classifying aggressive driving and only require data from the instrumented vehicle (Alkinani et al., 2020).

The literature lacks a conclusive study on how driving behaviours affect EC and what strategies, if any, drivers can employ to lower their EC across varying trips. Iclodean et al. (2025), published after the work in this chapter was conducted, provides some of the first insights into how typical driving behaviour categories affect EC across different trips. The results show that typical classification techniques fail to encapsulate the complex relationship between driving behaviour and EC. The results of Iclodean et al. (2025) are also limited to a small amount of DCs and rates of acceleration and deceleration. Therefore, there is a gap in the literature to find a driving behaviour classification method that can be shown to consistently produce differences in EC across different trip types. In addition, the results of this chapter can help expand on the current literature surrounding EV driving behaviour

by providing a review of current classification techniques and how they relate to EC and ECR for real trip data and modelled driving segments.

This review has presented many methods for classifying driving behaviour. Some of these methods are unable to be implemented in this thesis because they used parameters such as driving pedal angle which are not available here. Other methods provide different ways to group drivers based on speed and acceleration statistics, which is similar to a clustering technique. Clustering is tested in this thesis due to its popularity in the literature. Two additional methods were identified during the literature search for testing. The first is the limit-based method used by Al-Wreikat et al., which uses limits from the RDE test to classify driving behaviour based on speed and acceleration (2021). This method stands out as it can be used to compare DCs with varying distances and speeds, and EC was shown to significantly increase for aggressive drivers (Al-Wreikat et al., 2021). The other identified technique is from Xing et al. (2020), where real trip data was sorted by energy consumption and grouped by percentile into thirds. Then, driving behaviour statistics like average speed and acceleration were compared across the groups to see which driving behaviours contributed to low and high EC empirically (Xing et al., 2020). This method has the advantage of defining driving behaviours using already known results of EC. These two methods, along with clustering, will be implemented in the following sections of this chapter to see if any of the methods is able to classify driving behaviour in a way that directly relates to EC.

4.2 Clustering Method

The first method to categorise driving behaviour presented here is the clustering method. Clustering is a very popular ML method that groups datapoints based on chosen criteria. Clustering is an unsupervised ML method (Bozorgi, 2017), meaning that no information is given beforehand on how groups should be assigned. Instead, the clustering algorithm autonomously groups datapoints based on minimising its cost function in the

context of the given criteria (Abdel-Wahed et al., 2023). Minimising the cost function generally means assigning datapoints to groups so that, depending on the clustering algorithm used, the distance between datapoints in a cluster or the distance between datapoints and the centroid of the cluster is minimised (Abdel-Wahed et al., 2023).

For this thesis, the K-means algorithm was chosen to cluster driving behaviour. K-means clustering was chosen due to its simplicity and computational efficiency (Ye et al., 2023; Zhao et al., 2020). K-means clustering has also been shown to be a highly effective clustering method in a comparison study (Abdel-Wahed et al., 2023). In this thesis, K-means clustering was implemented in Python through the open-source library scikit-learn (Pedregosa et al., 2011). The K-means algorithm clusters datapoints to minimise the following cost function, Equation 4.1:

$$F(S, \mathbf{C}) = \sum_{k=1}^k \sum_{i=1}^{n \text{ in } S_k} (|y_i - c_k|)^2 \quad (4.1)$$

Where S_k is the set of datapoints within the k-th cluster, $\mathbf{C}$ is the set of clusters, k is the number of clusters chosen either by the user or through an algorithm, y_i is the location of datapoint i , and c_k is the centroid of the cluster to which i has been assigned (Kodinariya and Makwana, 2013). Here, “location” refers to the point occupied by a datapoint or cluster within an n -dimensional space (Kodinariya and Makwana, 2013).

In more tangible terms, the K-means algorithm modifies the locations of centroids and places datapoints into k number of sets to minimise the total squared error between the location of the datapoints and the location of the centroid of the cluster that datapoint was assigned to. The fact that the error term is squared in Equation 4.1 means that the K-means algorithm is sensitive to outliers (Abdel-Wahed et al., 2023). The number of clusters, k , chosen for the application of this thesis is three to align with the common method in the literature to divide driving behaviour into three groups (Al-Wreikat et al., 2021; Hu et al.,

2022; Iclodean et al., 2025) as well as the results of Bar et al. (2011), which indicated that more than three groups of behaviour may result in groups with similar behaviours.

Although the number of clusters used in this thesis is three to align with common categorisations of driving behaviour in the literature, there are alternate ways of selecting clusters. A popular method in statistics is to optimise the number of clusters based on the statistical properties of the dataset (Zhou et al., 2016B). In this case, the optimal number of clusters can be chosen where the error is not significantly reduced by adding another cluster (Syakur et al., 2018). The elbow method, as seen in Section 3.3.1, is a popular technique for identifying this inflection point, and, in this technique, error is defined as the total distance between points in a cluster and the cluster's centroid (Syakur et al., 2018).

The K-means algorithm iteratively moves centroids and reassigns datapoints to clusters based on the errors found in Equation 4.1 (Kodinariya and Makwana, 2013). This means the results of K-means clustering can be heavily influenced by the starting point of each cluster's centroid as the centroid can be caught in a local minimum during the iterative process (Ye et al., 2023). Because of this, several algorithms to optimise the starting location of each centroid have been developed. One such method is K-means++ (Arthur and Vassilvitskii, 2006), which improves on random selection of starting clusters by probabilistically choosing centroid locations. In K-means++, the first centroid location is chosen at random. Then, the probability of any being chosen as the location of the next centroid is given by Equation 4.2:

$$P(c = x) = \frac{dist(x - c_1)^2}{\sum dist(x_i - c_1)^2} \quad (4.2)$$

where x is any datapoint, c_1 is the first centroid, and x_i denotes that all datapoints are summed in the divisor of Equation 4.2. Equation 4.2 probabilistically assigns the next centroid location so that the location is less likely to be near a previously placed centroid and more likely to be near a high number of datapoints (Arthur and Vassilvitskii, 2006).

The final step before clustering is to determine which parameters will be used to group datapoints into clusters. Hu et al. (2022) utilised PCA before clustering to reduce the dimensionality of parameters before clustering, but PCA distances parameters from physical meaning and consequently limits the interpretability of results. Because K-means clustering is an unsupervised classification algorithm (Bozorgi, 2017), the algorithm does not assign a label to characterise the driving behaviour contained within each cluster. To be able to apply meaningful labels to the resulting clusters, it is useful to limit the number of parameters used to form the clusters (Jeon et al., 2023). A recent review on driving behaviours found that acceleration and speed are found twice as many times as other parameters to characterise eco-driving behaviour in the literature (Ma et al., 2024). For this reason, average speed and acceleration were chosen to divide trips into clusters in the clustering method applied here. Labels can easily be placed on the resulting clusters based on the relative difference of these parameters between clusters.

To demonstrate the clustering method, the 22,916 segments used to build the ABP model in Chapter 3 were clustered, and the results are shown in Figure 4.1.

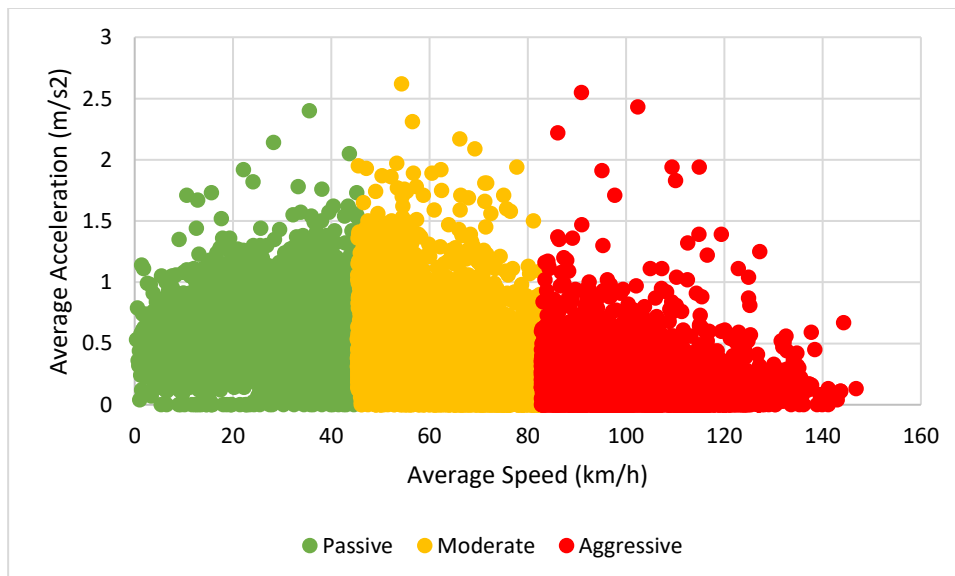


Figure 4.1. Clustering Method Results on Real Trip Segments

Figure 4.1 shows that the clusters were divided by average speed rather than average acceleration. Because there are no natural breaks in the distribution of these parameters, the

clusters are formed in a way that only differentiates between low, medium, and high speeds. Average acceleration peaks around where the passive and moderate clusters intersect and has lower values in the aggressive driving cluster. Clusters with higher average velocities were labelled as more aggressive because higher velocities lead to higher EC (Galvin, 2017). Clustering on the complete driving database means that driving behaviour classification was defined by speed of travel rather than a more complex representation of driving behaviour. To determine the effectiveness of the clustering method, it is necessary to evaluate the average EC and ECR of each cluster, are shown in Table 4.1.

Table 4.1. EC and ECR of Real Trip Segments Averaged by Cluster

Cluster	EC (kWh)	ECR (kWh/km)
Passive	0.06	0.13
Moderate	0.07	0.15
Aggressive	0.09	0.18

Ideally, the clustering method can differentiate segments by ECR, which is EC normalised by distance travelled. If the method can do that, it means that method can reliably relate driving behaviour to ECR over any trip type. EC, on the other hand, is heavily influenced by trip type. Trips and segments with higher speed will have a higher EC than trips with lower velocities regardless of acceleration behaviour. In this context, Table 4.1 presents mixed results. While the average ECR differs somewhat between each cluster, so does average EC. To determine whether the differences in EC and ECR are statistically significant, analysis of variance (ANOVA) tests are implemented here.

ANOVA tests are used in the field of statistics to determine whether the means of different groups are significantly different, with a null hypothesis that the population means of the groups are the same (Kim, 2017). ANOVA testing was implemented here in Python through the SciPy library (Virtanen et al., 2020). In this case, ANOVA tests will help to determine whether the differences in EC and ECR between the passive, moderate, and aggressive clusters are significant compared to the variance in EC and ECR within the

clusters. In this test, the null hypothesis is that the mean of EC and ECR between each cluster is the same. If the p-value is lower than $5.00\text{E-}2$, it means this null hypothesis can be rejected with 95% confidence. Then, the alternative hypothesis is accepted, which is that at the mean of at least one cluster is statistically different. If the p-value is higher than $5.00\text{E-}2$, it means that, due to high variance within the clusters, any differences in mean EC and ECR in Table 4.1 are not statistically significant. The results of ANOVA testing for the clustering of real trip segments by driving behaviour are shown in Table 4.2, where the f-statistic is the result of the test and the p-value determines how statistically different the means of EC and ECR are between clusters.

Table 4.2. ANOVA Test Results on Clustering Method Applied to Real Trip Segments

Metric	F-Statistic	P-Value
EC (kWh)	35.00	6.68E-16
ECR (kWh/km)	30.65	5.11E-14

Table 4.2 shows that, despite the small numerical differences in EC and ECR between each cluster, the clusters result in different means of EC and ECR that are statistically significant. However, as previously mentioned, these differences may be due to the types of segments tested. Additionally, the ANOVA tests only determine whether the mean of one cluster is different than the means of the others, so two clusters may not be statistically different.

Therefore, Tables 4.1 and 4.2 do not reveal whether the clustering method succeeds at differentiating driving behaviour of similar segment types. Further investigation is needed by applying the clustering method to a more limited sample where average speed does not vary as highly. This chapter concludes with such a study performed on the clustering method and the other two methods presented in the following sections.

4.3 Limit-based Method

The next method considered to categorise driving behaviour is the limit-based method. This method was used in Al-Wreikat et al. (2021) to classify driving behaviour as passive, moderate, or aggressive. The method is based on the European Union's RDE test, which is a required test that measures emissions of ICEVs during real trips (Regulation 2016/646). Within the specifications of the RDE exist dynamic limits that the vehicle must meet for a trip to be considered valid (Regulation 2016/646). The RDE procedure outlines both a lower boundary limit, which determines whether a trip is too passive, and an upper boundary limit, which determines whether a trip is too aggressive. The RDE test was also studied in Edwards (2022) and Huang et al. (2021) as it relates to measuring driver aggression. This method was selected for testing in this thesis because the limits scale with changing speeds and was shown in Al-Wreikat et al. (2021) to be a good indicator of EC for both short and long trips.

Before the limits are discussed, it is necessary to say that a trip is divided into speed bins before the limits are applied as detailed in the RDE test. The speed bins were made to represent different types of driving: urban, rural, and motorway. Every second of data with a speed less than 60 km/h is put in the urban speed bin, every second of data between 60 and 90 km/h is put in the rural speed bin, and every second of data above 90 km/h is put in the motorway speed bin. The data of each speed bin must be within the boundary limits for a trip to be considered valid in the RDE test (Regulation 2016/646).

Relative positive acceleration (RPA) is used to determine the lower boundary limit. In the RDE test, any trip below this limit is considered too passive. RPA is calculated using the formula in Equation 4.3:

$$RPA \left(\frac{m}{s^2} \right) = \frac{\sum_{t=1}^t (V * a^+)}{dist} \quad (4.3)$$

In other words, the product of speed and acceleration are summed for any second of data with positive acceleration. This summed value is then divided by distance travelled to obtain RPA. The limit of RPA is set by Equation 4.4 (Regulation 2016/646):

$$\begin{aligned} \text{if } v_{avg,i} \leq 94.05 \frac{km}{h} : \quad RPA_i &> \left((-0.0016 * v_{avg,i}) + 0.1755 \right) \\ \text{if } v_{avg,i} > 94.05 \frac{km}{h} : \quad RPA_i &> 0.025 \end{aligned} \quad (4.4)$$

where $v_{avg,i}$ is the average speed within speed bin i and RPA_i is the RPA of speed bin i . Trips with an RPA below this limit in any speed bin are determined to be too passive as it means the vehicle is not putting enough stress on the powertrain through high acceleration rates at high speeds.

At the other end, the aggressive limit is found by taking the 95th percentile value of speed times positive acceleration, denoted here as VA^{95th} . The product of speed and acceleration is taken for all seconds of data with positive acceleration above 0.1 m/s² within a speed bin, and the 95th percentile value of this product is the VA^{95th} . The aggressive limit for VA^{95th} is set by Equation 4.5 (Regulation 2016/646):

$$\begin{aligned} \text{if } v_{avg,i} \leq 74.6 \frac{km}{h} : \quad VA^{95th}_i &< \left((0.136 * v_{avg,i}) + 14.44 \right) \\ \text{if } v_{avg,i} > 74.6 \frac{km}{h} : \quad VA^{95th}_i &< \left((-0.097 * v_{avg,i}) + 31.635 \right) \end{aligned} \quad (4.5)$$

where $v_{avg,i}$ is the average speed within speed bin i and VA^{95th}_i is the VA^{95th} in speed bin i . In the RDE, a trip is considered too aggressive if the VA^{95th} in any speed bin is above this limit.

To better understand the RDE test's limits, the data from the 653 real trips used to build the EC model in the previous chapter was compared with both the passive and aggressive limits, shown in Figures 4.2 and 4.3 below. Because the RDE test is built around whole trips, trips are analysed instead of segments in this method.

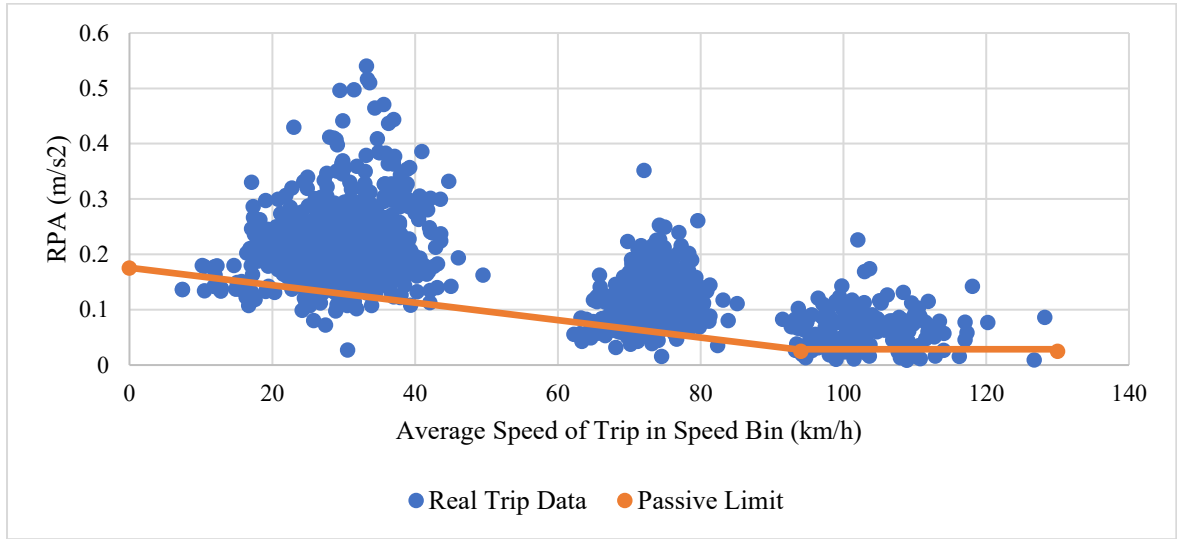


Figure 4.2. RPA Passive Limit Method Applied to Real Trips

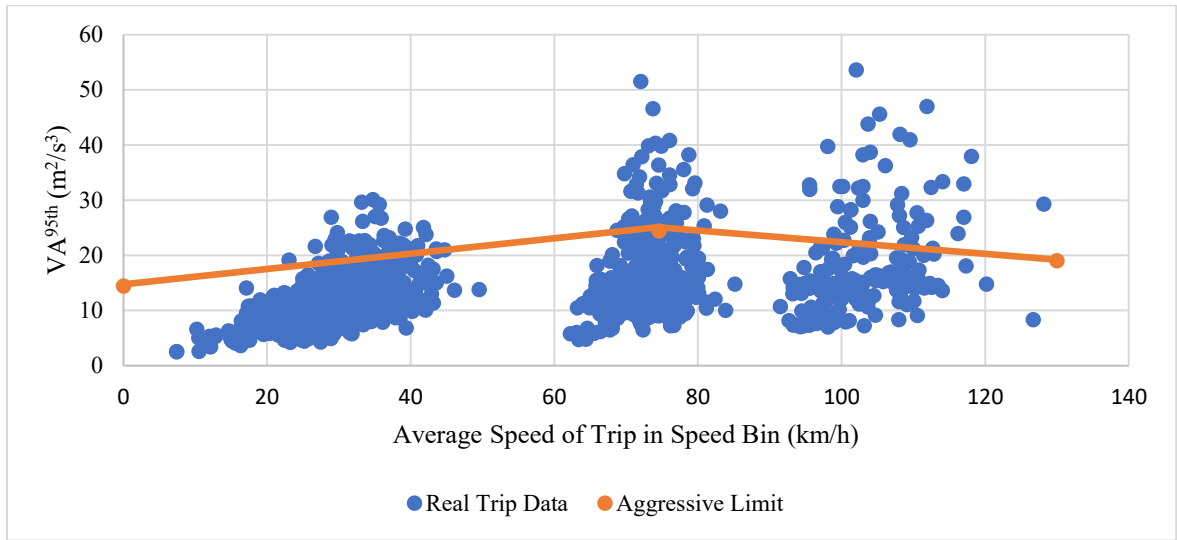


Figure 4.3. VA^{95th} Aggressive Limit Method Applied to Real Trips

In Figures 4.2 and 4.3, the limits described by the piecewise Equations 4.4 and 4.5 are displayed in orange. The data from the real-world trips are divided into three distinct sections in both Figure 4.2 and Figure 4.3 because the data within each trip was divided into three speed bins: urban, rural, and motorway. The average speed, RPA, and VA^{95th} was calculated for each speed bin of each trip. Just 7.76% of the trips were found to be below the passive limit in Figure 4.2 while 12.62% of trips were above the aggressive limit in Figure 4.3. These low percentages indicate that the limits set in the RDE test are a good indicator of extreme passive and aggressive driving behaviour.

In Al-Wreikat et al. (2021), trips are determined to be aggressive if the trip VA^{95th} is within 25% of the aggressive limit, and trips are passive if the trip RPA is within 25% of the passive limit. All trips neither passive nor aggressive are classified as moderate. Trips above the aggressive limit or below the passive limit are not studied in Al-Wreikat et al. (2021), but Figures 4.2 and 4.3 demonstrate that many real EV trips exceed these limits and therefore should not be excluded from study. Al-Wreikat et al. (2021) also makes no mention of how speed bins are incorporated into this classification process, so a standard method is needed for the application of a limit-based classification method in this thesis.

The limit-based method of classification developed here creates one metric for trip aggressiveness based on the RPA and VA^{95th} in each speed bin of each trip. First, for each trip the difference between RPA and the passive limit is averaged across all speed bins. The difference between VA^{95th} and the aggressive limit is also averaged across the speed bins. These differences are compared across all trips to get one measure of driver aggression, seen in Equation 4.6. This term was designed to be a simple way to avoid the limitations of the methodology presented in Al-Wreikat et al. (2021), which did not account for both RPA and VA^{95th} terms of the same trip or trips which exceeded the passive or aggressive limits while maintaining the structure of the limits of the RDE test. The term is based on the relative difference in a trip's proximity to the passive and aggressive limits compared to the mean values of all trips. This aggression measurement can improve on studies that use the RDE test's limits to classify driver aggression such as Al-Wreikat et al. (2021), Huang et al. (2021), and Edwards (2022) by combining a trip's relationship to both the aggressive and passive limit into one term, seen below:

$$\begin{aligned}
 Diff_{A,i} &= VA^{95th}_i - Aggressive\ Limit_i \\
 Diff_{P,i} &= RPA_i - Passive\ Limit_i \\
 Aggression &= \frac{\frac{1}{n}\sum(Diff_{A,i}) - \frac{1}{m}\sum\frac{1}{n}\sum(Diff_{A,i})}{\left|\frac{1}{m}\sum\frac{1}{n}\sum(Diff_{A,i})\right|} + \frac{\frac{1}{n}\sum(Diff_{P,i}) - \frac{1}{m}\sum\frac{1}{n}\sum(Diff_{P,i})}{\left|\frac{1}{m}\sum\frac{1}{n}\sum(Diff_{P,i})\right|}
 \end{aligned} \tag{4.6}$$

where n is the number of speed bins for a trip, m is the total number of trips, the aggressive and passive limits are calculated using Equations 4.4 and 4.5 using the average speed of speed bin i .

This aggression term ranks driver aggression across all trips by calculating the relative difference of the aggression of one trip to the average aggression of all trips, and the same calculation is done for passiveness with the results combined into one term. Driving behaviour beyond the passive and aggressive limits can be included in classification through Equation 4.6. The trips can then be sorted and grouped into equal thirds based on this aggression value, with the highest third classified as aggressive, the middle third moderate, and the lowest third passive. This method classifies driving behaviour based on the relative difference to both the passive and aggressive limits in all speed bins, and it can be used to classify driving behaviour across DCs with different distances and average velocities.

As with the clustering method, the limit-based method was applied to real driving data. The RDE test's limits are designed to fit a trip, so here driving behaviour is grouped by trip rather than segment. The aggression measure was calculated for each of the 653 trips. The trips were then sorted by aggression measurement and grouped into equal thirds. The results are shown in Figure 4.4, which relates aggression to ECR. Alternatively, other percentile metrics could have been used to divide these categories. This could be done to widen the pool of moderate drivers and ensure that more extreme driving behaviours are classified as passive or aggressive.

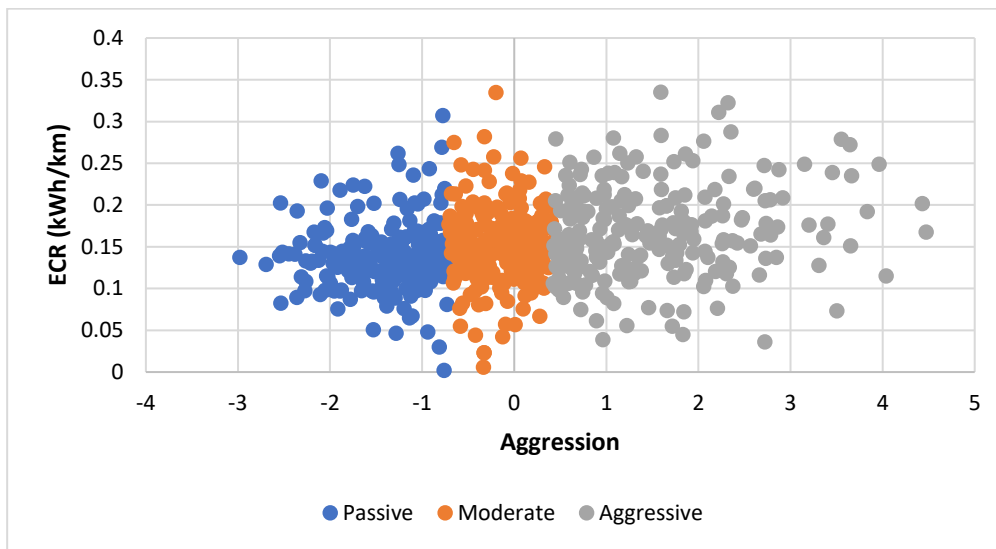


Figure 4.4. ECR by Aggression of Real Trips: Limit-Based Method

Figure 4.4 shows that the aggression term shows some positive correlation with ECR. That is, trips with a higher value of aggression tend to have higher ECR. The moderate drivers tend to be clustered around zero aggression, indicating that this metric succeeds at fitting the aggression rating to the sample by measuring only the relative differences to the passive and aggressive limits to the average of all trips. There tend to be more extreme outliers at the aggressive end, which can be explained by the fact that more trips exceeded the aggressive limit in Figure 4.3 than exceeded the passive limit in Figure 4.2. The aggression term was found to have a higher Pearson's correlation to ECR in kWh/km (0.26) than to EC in kW (0.17), indicating that it may be a good classifier of ECR regardless of the total EC. The Pearson's correlation coefficient of aggression to ABP was even higher at 0.37, meaning it is a moderate predictor of ABP according to Ji et al. (2022). ABP is a measurement of ECR normalised by time rather than distance, which means this aggression term can be a good indicator of how driving behaviour relates to EC across the same trip duration. The results of the classification method in terms of EC and ECR are presented below in Table 4.3.

Table 4.3. EC and ECR of Real Trips by Limit-Based Group

Group	EC (kWh)	ECR (kWh/km)
Passive	2.26	0.14
Moderate	2.41	0.15
Aggressive	4.22	0.17

Table 4.3 shows similar results to Table 4.1 from the clustering method. The EC values in Table 4.3 differ greatly from those in Table 4.1 because the limit-based method classifies trips while the clustering method was used on segments. In both cases, both EC and ECR increase with increasing aggressiveness. In this case, higher EC means generally longer trips which are likely to include higher speeds and in turn higher ECR. As with the clustering method, the method's ability to differentiate driving behaviour in a way that correlates to ECR is inconclusive. However, a more in-depth look at how the limit-based method performs among similar trips is required later in the chapter.

As with the clustering method in Section 4.2, ANOVA testing can help to determine whether the differences in EC and ECR are statistically significant between the groups formed through the limit-based method. Table 4.4 presents the results of ANOVA testing on the limit-based classification of real trips.

Table 4.4. ANOVA Test Results on Limit-Based Method Applied to Real Trips

Metric	F-Statistic	P Value
EC (kWh)	9.38	9.60E-05
ECR (kWh/km)	20.21	2.97E-09

As with Table 4.2, the results of Table 4.4 show that the means between groups in both EC and ECR are statistically significant at above 99% confidence. The p-values in Table 4.4 are lower than those in Table 4.2 mostly due to the lower sample size of trips compared to segments. Still, there is not enough information about distinguishing between EC and ECR on similar trips, so more investigation is needed later in this chapter.

4.4 Empirical Method

The final method considered to classify driving behaviour is the empirical method. The empirical method involves sorting all trips by ABP into three groups and comparing the driving behaviour statistics of each group. Unlike other classification methods, the empirical method uses EC results to retrofit classifications onto driving behaviours. This method is taken from Xing et al. (2020), and it was selected because it directly relates driving behaviour to EC and uses real data to categorise driving behaviour.

Unlike the previous methods, the empirical method does not require an extensive methodology. Real driving segments are sorted from least to most ABP, grouped into equal thirds in this sorted order, and the average statistics of each group are compared. Since the ABP model in Chapter 3 was built at the segment level, here segments, rather than full trips as in Xing et al. (2020), are sorted by ABP. Figure 4.5 displays the 22,916 real driving segments used to build the ABP model grouped into light, moderate, and heavy ABP. Here, the light ABP group is the third of segments that has the least ABP, the moderate ABP group has the middle third of ABP values, and the heavy ABP group has the highest ABP values.

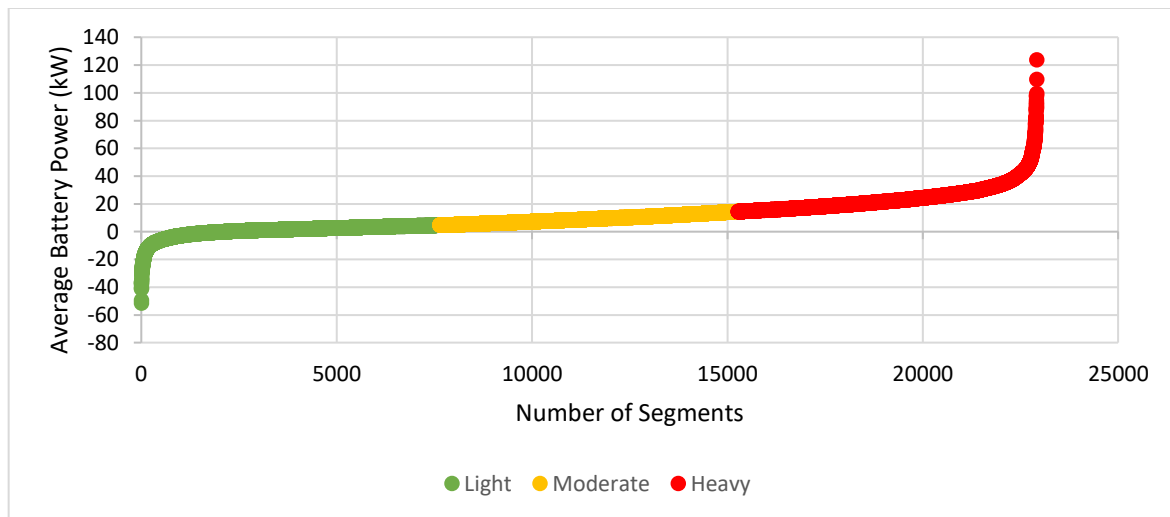


Figure 4.5. Segment Average Battery Power by Empirical Method Group

Figure 4.5 shows that ABP generally increases moderately from around 0 to 30 kW, but there are some extreme values of ABP for both light and heavy energy consumers. To compare the driving behaviour between groups, the average statistics of all parameters for

each group are shown in Table 4.5 below, which is sorted by relative difference between groups.

Table 4.5. Average Statistics of Parameter by Group

Parameter	Average Battery Power		
	Light	Moderate	Heavy
Gradient (%)	-0.56	0.06	0.52
Idle Time (%)	9.35	1.15	0.03
Number of Stops	0.75	0.10	0.00
Urban (%)	46.75	19.32	1.66
Motorway (%)	7.93	17.21	35.41
Maximum Deceleration (m/s ²)	-1.40	-0.81	-0.34
Std Dev Deceleration (m/s ²)	0.36	0.22	0.08
Std Dev Speed (km/h)	9.95	5.73	3.06
Deceleration Time (%)	33.59	22.53	11.38
Std Dev Acceleration (m/s ²)	0.30	0.20	0.12
Maximum Acceleration (m/s ²)	1.18	0.79	0.47
Average Deceleration (m/s ²)	-0.51	-0.36	-0.20
Average Speed (km/h)	42.28	65.84	93.85
Average Moving Speed (km/h)	44.48	66.21	93.86
Wind Factor (km/h)	0.03	0.02	0.01
Secondary Road (%)	39.25	32.52	17.17
Cruise Time (%)	33.30	51.61	67.24
Residential Road (%)	14.95	11.34	7.85
Average Acceleration (m/s ²)	0.44	0.34	0.24
Tertiary Road (%)	15.94	12.98	10.71
Primary Road (%)	3.92	6.26	6.33
Rain (mm)	0.08	0.09	0.07
Vehicle Age (Yr)	2.76	2.60	1.94
Unclassified Road (%)	11.41	14.06	15.17
Male	0.71	0.78	0.85
Acceleration Time (%)	23.76	24.71	21.35
Usable Battery (kWh)	55.07	55.01	60.51
Temperature (°C)	14.24	14.15	13.75
Wind Speed (km/h)	2.04	2.03	2.05

A few conclusions can be drawn from the data presented in Table 4.5. First, the table shows that gradient has a noticeable effect on average ABP. The segments with highest ABP tended to be on uphill segments while the lowest ABP segments tended to be going downhill. The differences across groups in several parameters all contribute to showing that the highest ABP segments tend to be on motorways cruising at high speeds while the lowest ABP

segments are low speed urban driving with higher acceleration and deceleration due to more frequent stops. For instance, the heavy ABP group is characterised by low idle time, high cruise time, low number of stops, low percent urban, high percent motorway, low maximum, average, and std dev deceleration, low maximum, average, and std dev acceleration, high average and moving speed, and low std dev speed. All of these values point toward heavy ABP characterised by high-speed motorway journeys, meaning little information can be drawn on how driving behaviour can affect EC along similar driving segments. Therefore, Table 4.5 reveals how, when all segments are compared together, the segments with the highest ABP tend to be on high-speed roads while the lowest tend to be more urban segments with lower speeds.

One other insight from Table 4.5 is that the percent time spent decelerating (11.38) is on average significantly lower than the time spent accelerating (21.35) for the heavy ABP group, and, within the light ABP group, time spent decelerating (33.59) is significantly higher than time spent accelerating (23.76). This is an example of the power regained through regenerative braking. The ability of EVs to recover energy through regenerative braking means that segments with higher deceleration tend to have lower ABP. Driver and vehicle-related parameters show some differences across groups, but the differences are likely due to the data used in this thesis rather than being replicable across other studies.

The results of the empirical method presented here give little usable output for EV drivers or policymakers. There is no way to classify driving behaviour along similar segments using the data presented in Table 4.5 as the ABP groups appear to be divided mostly by type of driving segment (urban or motorway) rather than driving behaviour that could be altered along those segments to decrease EC. To classify driving behaviour empirically, more information is needed on how parameters affect ABP within different speed ranges. While the empirical data could be divided into speed intervals and the empirical method could be repeated for each speed interval, the findings would risk being

less robust due to a more limited sample size. Instead, a SA is introduced in the following section so that the effect on ABP of the complete range of input variables can be tested at different speed intervals. The results and information from the SA will enable a driving behaviour categorisation method for different trip types.

4.5 Literature Review on Sensitivity Analysis

A SA is conducted as part of this chapter to better understand how changing values of parameters influence EC. This literature review of the current state of the SA field was conducted to ensure the best SA method for the purposes of this thesis was selected. Some information on the studies included in this literature is presented in Appendix E for ease of comparison.

4.5.1 Sensitivity Analysis Introduction and Terms

A detailed review on SA techniques is not currently available for the field of EV EC. The search was expanded to the literature of ICEVs, other energy systems, and neural network models in any field as SA techniques are applicable to all areas of research. First, the definition of SA is given as the study of how much variance in an output or target variable of a model can be attributed to the different input or explanatory variables (Tsvetkova and Ouarda, 2021). SA aids in the understanding of a research topic by helping to explore how input variables affect the target variable. SA can also help researchers reduce the dimensionality of a model by removing any insignificant variables (Razavi et al., 2021). Braband et al. (2022) outlines the steps required to setup a SA, highlighting the importance of determining which parameters will be investigated and learning their distribution.

Tsvetkova and Ouarda (2021) draw an important distinction between two types of SA: local and global. Wei (2013) asserts that although both types of SA are commonly used, a global sensitivity analysis (GSA) generally is preferred for its reliability and robustness despite its increased computational cost. A local sensitivity analysis only examines the change in parameters from one point in the input variable space: the base case, which is

generally the mean value of each input parameter (Wei, 2013). The input variable space refers to all possible values of that variable, which can be determined empirically or theoretically (Wei, 2013). A GSA analyses the sensitivity to a variable across its entire space. The most common local SA is called one-at-a-time sensitivity analysis (OATSA), which is what most researchers mean historically when referring to SA (Tsvetkova and Ouarda, 2021). OATSA is generally performed by changing the value of one parameter in the model at a time while the rest of the parameters are held at a constant value in a base case (Tsvetkova and Ouarda, 2021).

GSA, on the other hand, is a more recent branch of SA that considers the full range of the input parameters (Tsvetkova and Ouarda, 2021; Wei, 2013). Local SAs monitor how the output changes when a parameter deviates from the base case, but there is no base case in GSA. GSA has the major advantage of being applicable to any model, whereas local SA is limited to being most effective when applied to a linear model (Tsvetkova and Ouarda, 2021). The limitation of local SA to linear models is because changing the input variable by, for example, 1% would change the output by the same amount regardless of where the base case exists in the space of the input variable. On the other hand, in a nonlinear model changing the input variable by 1% could result in smaller or larger changes to the output depending on where a base case sits in the input space due to the nonlinear relationship between the input and output variables. Alternatively, GSA accounts for the sensitivity of the model to input variables across the entire space of the input variable, making it more useful for nonlinear models such as ANNs (Tsvetkova and Ouarda, 2021).

4.5.2 Global Sensitivity Analysis Methods

Razavi et al. (2021) outline four major methods for implementing GSA. Wei (2013) adds another method, meta-model-based SA, which creates a simplified meta-model that eases computational loads when calculating sensitivity by aggregating the effects of less influential parameters. This method is not considered for the SA in this paper as it adds an

unnecessary level of complexity when understanding the effects of input variables. The main purpose of a meta-model is to simplify the calculations of the SA, but the ABP model of this thesis does not reach a level of complexity that necessitates a meta-model (Coutts and Yokomizo, 2014). The SA can be implemented on the modelled parameters with minimal computational time without a meta-model.

Derivative-based methods are a logical extension of local SA. A derivative-based GSA analyses the effect of changing inputs over a number of points in the input space, not just one base case (Razavi et al., 2021). The structure of ANNs allows for the calculation of partial derivatives of the output variable for each of the input variables, making derivative-based SA achievable for ANNs (Ricotti and Zio, 1999). The Morris method is a systematic derivative-based method that is generally used to reduce model dimensionality (Tsvetkova and Ouarda, 2021). The Morris method is computationally light but does not present a deep quantitative analysis on the explained variance by input variables (Wei, 2013). Abburu et al. (2024) describes the steps of the Morris method in detail. An example derivative-based SA is found in Ben Chaim et al. (2013). One shortcoming of the Morris method is that it does not provide information on the most significant interactions to the model (Kioutsioukis et al., 2004).

Distribution-based methods usually analyse how the variance of the output variable can be attributed to input variables or interactions between input variables (Razavi et al., 2021; Wei, 2013). Distribution-based methods generally output both first order effects, which consider only how an input directly effects the output, and total effects, which also consider the significance of an input's correlations with other inputs and nonlinear relationships with the output (Wei, 2013). Other distribution-based GSA methods look at the skew and kurtosis of the output distribution rather than its variance (Razavi et al., 2021). Still other methods model the difference between the output distribution of the full model and the distribution when one or more input variables are fixed (Razavi et al., 2021). A

simple distribution-based method is observing the scatterplots between each independent variable and the output variable. Saltelli et al. (2008) theorised that all SA techniques aim to systematically measure the quantitative information that can be found qualitatively in scatterplots.

There are several distribution-based methods popular in state-of-the-art literature that rank input parameters based on their contribution to nonlinear models. The Fourier amplitude sensitivity test is another method to compute sensitivity indices efficiently (Culkier et al., 1978; Tsvetkova and Ouarda, 2021). Shapley values and effects are yet another method of developing sensitivity indices. Shapley values are derived from game theory and applied to GSA to determine how much explained variance in the output variable can be attributed to each input variable (Owen, 2014; Shapley, 1953; Tsvetkova and Ouarda, 2021). Song et al. (2016) demonstrated how Shapley values can fully and efficiently model the effects of interactions between variables. Finally, the PAWN method, proposed in Pianosi and Wagener (2015), aims to further improve efficiency by analysing the sensitivity of the output variable's cumulative distribution function rather than its probability distribution function (Tsvetkova and Ouarda, 2021).

The Sobol' method is a popular distribution-based method that measures sensitivity indices which rank the input variables on their importance to the model (Tsvetkova and Ouarda, 2021). The Sobol' method utilises analysis of variance tests to measure sensitivity in a manner that can account for interaction and dependencies between input variables (Sobol', 1990; Tsvetkova and Ouarda, 2021). One limitation of the Sobol' method is that it does not indicate which direction the target variable moves when an input variable is increased or decreased, but that information can be found graphically or with further analysis (Abburu et al., 2024). The Sobol' method has been used frequently in the EV literature and related fields (Botshekan et al., 2019; Lai et al., 2021; Vepsäläinen et al., 2019).

Variogram-based methods are similar to derivative-based methods as they involve testing the model at different points in the entire input space (Razavi et al., 2021). They are meant to improve upon distribution-based and derivative-based methods by capturing the structure of the output distribution and covariance instead of assuming a random distribution. Razavi and Gupta (2016) provides an example case of variogram-based GSA for reference.

Regression-based methods historically use coefficients from linear regression such as MLR to create a SA (Razavi et al., 2021). These methods generally have the same limitation that was noted previously for local SA, where the SA relies on the model fitting well to the data with linear relationships. More recent approaches to regression-based GSA have been applied to nonlinear models through analysing the output variance that occurs when input variables are changed (Razavi et al., 2021). The permutation feature importance method implemented in the previous chapter is one example of regression-based GSA.

4.5.3 Uncertainty Analysis

An uncertainty analysis is a topic related to SA where the uncertainty of the output is measured against changes in inputs. This technique is usually accomplished using Monte Carlo simulations of input variables to find the distribution of the output variable for the range of inputs (Braband et al., 2022). Jadali et al. (2021) runs an uncertainty analysis by probabilistically creating 100,000 iterations of input data, assuming each of the variables followed a normal distribution. Gawlikowski et al. (2023) notes that uncertainty in ANNs may be computed as part of building the ANN or afterwards using external analysis methods. Raghu (2019) recommends using external measures of uncertainty as combining the development and uncertainty analysis of an ANN may bias the uncertainty results. Ricotti and Zio (1999) used five equally spaced input values, instead of Monte Carlo simulation, to perform an uncertainty analysis on an ANN. Braband et al. (2022) states that uncertainty analysis does not provide the necessary information to quantify the influence of input

parameters on the output. So, though uncertainly analysis is a straightforward process to understand how an input variable may affect the output, it does not sufficiently quantify variable effects or interactions and therefore will not be used in this thesis.

4.5.4 Sensitivity Analysis for Energy Consumption of Electric Vehicles

Following the general review of SA methods in the previous sections, this subsection covers the applications of SA in the EV literature. As seen in Appendix E, many studies in the field of EV EC utilise local OATSA (Guimarães et al., 2019; Iora and Tribioli, 2019; Ji et al., 2022; Miri et al., 2021). Asef and Lapthorn (2021) found the number of papers in automotive engineering applied OATSA was more than double the amount of papers that used GSA, although recent papers showed a more equal share. Yi and Bauer (2017) provides a global OATSA using partial derivatives of EC relating to five environmental parameters. The SA is repeated over the full range of recorded speed values for a Nissan Leaf, from 0 to 128 km/h.

One crucial aspect to developing a SA is determining the input values on which the model will be tested to perform the SA. Both the number and distribution of samples used for SA must be determined. This is simple for OATSA, where the values are generally just a few deviations from a base case at the mean value (Wei, 2013). For GSA, on the other hand, a finite number of values need to be chosen to represent the entire input space of each parameter. Generally, around 1,000 samples are generated for GSA (Asef and Lapthorn, 2021; Lai et al., 2021; Schäfers et al., 2024; Vepsäläinen et al., 2019), although that number can vary to be significantly higher (Teymouri et al., 2023) or lower (Huang et al., 2022; Pal et al., 2017). Tran et al. (2024) use Monte Carlo simulation to create fifty thousand different wind scenarios which are added to a dataset of three real trips to discover the sensitivity of EC to wind speed and direction. Asamer et al. (2016) manually created a set of input cases to create a SA using the Sobol' method. The SA was iteratively computed over a few different vehicle speeds and elevation profiles (Asamer et al., 2016). It can be helpful to repeat SA to

cover different scenarios such as different speeds (Asamer et al., 2016) or area type (Schäfers et al., 2024).

The range of input variables used during SA should be determined in the context of why the SA is being implemented (Wei, 2013). For example, if the SA is being used to make a policy choice on speed limits, the range of inputs should reflect only the possible speed limits that are being considered for the policy. Wei (2013) suggests Latin-hypercube sampling (LHS) as a popular and efficient method for generating input samples to use in GSA. LHS is a quasi-random selection method that ensures the variables are selected evenly across the entire input space, as opposed to fully random Monte Carlo sampling (Schäfers et al., 2024). Both Monte Carlo sampling (Asef and Lapthorn; Pal et al., 2017; Vepsäläinen et al., 2019) and LHS (Huang et al., 2022; Schäfers et al., 2024) are used in state-of-the-art research to sample input values. Sobol' sampling is a quasi-random method similar to LHS, but Sobol' sampling is unique in that the values are selected to be equally spaced across all dimensions (Sobol', 1967).

4.6 Sobol' Sensitivity Analysis

4.6.1 Sobol' Method Implementation

Following the review of the literature, the Sobol' method was selected for the SA implemented in this section. This method was selected due to its ability to quantify the effects of input parameters independently and the effects of interactions with other variables and for its frequent use in state-of-the-art literature. Interactions are particularly important for an ANN since, as opposed to linear models like MLR, interactions between input variables can significantly affect the model's predictions (Tsang et al., 2018). The outputs of the Sobol' method include three different sets of unitless sensitivity indices which all contain important information on how the parameters affect the output (Vepsäläinen et al., 2018):

1. Total-order effects, marked as ST, quantify how much variance of the target variable is explained by variance of the input variable with all interactions included

2. First-order effects, S1, quantify the variance explained by the variance of an input variable independently (i.e. without interactions)
3. Second-order effects, S2, quantify the variance explained by the interaction of a pair of variables

The Sobol' method conducted in this chapter was performed in Python using the open-source library SALib (Herman and Usher, 2017; Iwanaga et al., 2022). The Sobol' SA methodology, provided by SALib in this thesis, is a variance decomposition method which quantifies sensitivity as the amount of variance in the output explained by the variance of an input or the interactions of inputs (Lai et al., 2021; Saltelli et al., 2010; Tosin et al., 2020). Mathematically, the model output Y is decomposed into the sum of its parts as seen in Equation 4.7 below (Tosin et al., 2020):

$$Y = M_0 + \sum_i^n M_i(X_i) + \sum_{i,j}^n M_{ij}(X_i, X_j) + \dots + M_{i\dots n}(X_i, \dots, X_n) \quad (4.7)$$

Expanding on Equation 4.7, the following terms can be understood through Equation 4.8 (Tosin et al., 2020):

$$\begin{aligned} M_0 &= E(Y) \\ M_i(X_i) &= E(Y|X_i) - M_0 \\ M_{ij}(X_i, X_j) &= E(Y|X_i, X_j) - M_0 - M_i - M_j \end{aligned} \quad (4.8)$$

That is, M_0 is the expected value of Y (i.e. the mean value of Y), and the subsequent terms are expected values of Y conditional on a value of one or more input variables X subtracted by the previous expected values. The variance can then be decomposed into a sum of these conditional variances as seen in Equation 4.9, where u denotes any subset of X (Tosin et al., 2020):

$$Var(Y) = \sum_u Var(M_u(X_u)) \quad (4.9)$$

From equation 4.9, the Sobol' nth-order index can be calculated from the following Equation 4.10:

$$Sn_{i,(j),...} = \frac{Var\left(M_{i,(j),...}(X_{i,(j),...})\right)}{Var(Y)} \quad (4.10)$$

where the first-order Sobol' index $S1_i$ would be computed for a single variable X_i and the second-order index $S2_{ij}$ for the interaction between variables X_i and X_j . Further orders are calculated until all possible interactions have been computed (Tosin et al., 2020). The total-order index ST_i is the sum of all sensitivity indices of the variables X_i (Tosin et al., 2020). In practice, the variances in Equation 4.10 are estimated through sampled values (Saltelli et al., 2010).

An important step of SA is sampling a finite set of input values to be modelled. As mentioned in the literature review of Section 4.5.4, the sampling methods of LHS, Monte Carlo sampling, and Sobol' sampling are all broadly similar methods to sample n input values from a range of input values. While Monte Carlo sampling is truly random, LHS and Sobol' sampling are quasi-random methods that guarantee samples are taken evenly across the range of input variables. Due to the similarity between these sampling methods, the Sobol' method was selected due to its ease of implementation within SALib. Sobol' sampling has been shown to be superior to LHS and Monte Carlo when building a SA (Renardy et al., 2021).

Sobol' sequencing or sampling is a low-discrepancy sampling method meant to improve on random sampling (Saltelli et al., 2010). The resulting samples of the Sobol' method are as equally distributed as possible across the input space (Saltelli et al., 2010). The sampled values are used to construct a matrix of size $N \times 2d$ (Saltelli et al., 2010). N is the number of quasi-random samples, and d is the number of input variables, 29 for the ABP model of this thesis. Here, N is chosen as 1,024 to align with the majority of state-of-the-art SAs. The matrix is split into matrices **A** and **B**, where each is a matrix of 1024 independently

sampled inputs (Saltelli et al., 2010). To shift input values and determine the sensitivity of the model, a matrix $\mathbf{AB}^i$ is developed for every column by replacing the i -th column of matrix $\mathbf{A}$ with the i -th column of matrix $\mathbf{B}$ (Saltelli et al., 2010). This process created a total of 61,440 inputs to the model on which to test model sensitivity, with each input representing the average statistics of a 500m segment.

Although the uniform distribution is the default setting in SALib and commonly assumed in the literature (Kioutsioukis et al., 2004; Lai et al., 2021), it is essential to the SA that the samples are representative of realistic data through accurate distributions. The distributions of the input variables used during sampling directly affects the resulting sensitivity indices. SALib allows each input variable to be distributed independently through the following distributions: uniform, logarithmic uniform, triangular, normal, truncated normal, and lognormal. The best-fit distributions for each variable were determined manually through scatterplots and descriptive statistics, and each variable closely fit either the normal, uniform, or truncated normal distributions. Each of these distributions are defined by distinct parameters, called params in this thesis to avoid confusion with the input parameters of the ABP model.

As mentioned previously, each modelled parameter could be reasonably assumed to be a random variable described by a uniform, normal, or truncated distribution. For SALib to generate samples, each input parameter must be defined with a distribution and that distribution's associated params. A random variable with a uniform distribution is equally likely to have any value within a certain range (Forbes et al., 2011). This range is defined in Salib by a minimum (Param 1) and maximum value (Param 2) (Herman et al., 2025). A normal distribution describes a parameter distributed about a mean (Param 1) with a standard deviation (Param 2) (Forbes et al., 2011). A truncated normal distribution describes a random variable that is normally distributed but confined to a certain range usually due to physical limitations (Burkardt, 2014). Truncated normal distributions were required for some

parameters of the ABP model that could not physically be above or below zero. Truncated normal distributions are defined by four params in SALib, so N/A values are common in Tables 4.6 and 4.7 because normal and uniform distributions are defined by only two params. (Herman et al., 2025). The params that define each distribution type are explained in Table 4.6, while the distributions used for sampling each input parameter to the model are displayed in Table 4.7. Table 4.7 can be used to understand the input space of each variable that is analysed in the Sobol' GSA.

Table 4.6. Distribution Params

Distribution	Param 1	Param 2	Param 3	Param 4
Normal	Mean	Std Dev	-	-
Uniform	Min	Max	-	-
Truncated Normal	Min	Max	Mean	Std Dev

Table 4.7. Parameter Distributions for Sobol' Sampling

Parameter	Distribution	Param 1	Param 2	Param 3	Param 4
Gradient (%)	Normal	0	1.65	N/A	N/A
Number of Stops	Uniform	0	10	N/A	N/A
Rain (mm)	Uniform	0	1	N/A	N/A
Temperature (°C)	Normal	14.05	4.00	N/A	N/A
Wind Speed (km/h)	Uniform	0	10	N/A	N/A
Wind Factor (km/h)	Uniform	-5	5	N/A	N/A
Average Speed (km/h)	Uniform	0	128	N/A	N/A
Std Dev Speed (km/h)	Truncated Normal	0	43.05	6.25	6.42
Average Acceleration (m/s ²)	Truncated Normal	0	2.62	0.34	0.28
Maximum Acceleration (m/s ²)	Truncated Normal	0	6.11	0.81	0.80
Std Dev Acceleration (m/s ²)	Truncated Normal	0	1.68	0.21	0.22
Average Deceleration (m/s ²)	Truncated Normal	-2.41	0	-0.35	0.29
Maximum Deceleration (m/s ²)	Truncated Normal	-6.67	0	-0.85	0.84
Std Dev Deceleration (m/s ²)	Truncated Normal	0	1.89	0.22	0.23
Idle Time (%)	Truncated Normal	0	0.75	0.04	0.12
Cruise Time (%)	Uniform	0.05	1	N/A	N/A
Acceleration Time (%)	Uniform	0	0.8	N/A	N/A
Deceleration Time (%)	Uniform	0	0.8	N/A	N/A
Urban (%)	Uniform	0	1	N/A	N/A
Primary Road (%)	Uniform	0	1	N/A	N/A
Secondary Road (%)	Uniform	0	1	N/A	N/A
Tertiary Road (%)	Uniform	0	1	N/A	N/A
Unclassified Road (%)	Uniform	0	1	N/A	N/A
Motorway (%)	Uniform	0	1	N/A	N/A

Residential Road (%)	Uniform	0	1	N/A	N/A
Usable Battery (kWh)	Uniform	28	88	N/A	N/A
Male	Uniform	0	1	N/A	N/A
Vehicle Age (Yr)	Uniform	0	7	N/A	N/A

Because the parameters average speed and average moving speed are correlated with a Pearson's value of 0.996 in the collected data, the inputs for average moving speed were made equal to the average speed values rather than being sampled. This was done to ensure the input values for SA did not contain unrealistically high differences between these two parameters. The resulting effect is that average moving speed is not analysed. No other logical limitations were placed on the Sobol' samples to, for instance, ensure that drive mode percentages sum to 100%. This is a limitation of the Sobol' SA methodology when applied to the ABP model.

The results of permutation feature importance in the previous chapter and the empirical method in Section 4.4 indicate that the two most impactful components of a driving segment on ABP are the average speed and gradient. So far, all analysis on factors affecting EC has been with all types of trips or segments grouped together. It is not debated that higher speeds result in higher EC (Galvin, 2017). There has been no method implemented to look further into the model and differentiate how parameters affect EC under different driving conditions, such as at high speeds versus low speeds. These differentiations can be made by limiting the distributions of input variables during Sobol' sampling. In this way, SA can be repeated under different ranges of speed as seen in Asamer et al. (2016). But first, a SA is conducted on the entire range of inputs, with the results explored in the following subsection.

4.6.2 Sensitivity Analysis on Entire Range of Inputs

The Sobol' methodology described in the previous section was implemented through SALib. The results of the SA are shown below in Table 4.8. Both the first-order effects, S1, and total-order effects, ST, are included in the table. The S1 values calculate what fraction

of the variance in the output variable (ABP) is removed by fixing the value of that parameter (Nossent et al., 2011). ST estimates the fraction of variance in ABP that can be explained by one input parameter and that parameter's interactions with all other parameters.

Table 4.8. Sobol' Sensitivity Indices

Parameter	S1	ST
Average Speed (km/h)	0.30	0.58
Acceleration Time (%)	0.12	0.20
Deceleration Time (%)	0.09	0.18
Gradient (%)	0.07	0.12
Usable Battery (kWh)	0.00	0.06
Number of Stops	0.00	0.06
Wind Speed (km/h)	0.00	0.04
Wind Factor (km/h)	0.01	0.04
Vehicle Age (Yr)	0.00	0.04
Average Acceleration (m/s ²)	0.01	0.03
Std Dev Acceleration (m/s ²)	0.02	0.03
Temperature (°C)	0.01	0.03
Rain (mm)	0.01	0.03
Std Dev Speed (km/h)	0.00	0.03
Average Deceleration (m/s ²)	0.00	0.02
Primary Road (%)	0.00	0.02
Idle Time (%)	0.01	0.02
Tertiary Road (%)	0.00	0.02
Std Dev Deceleration (m/s ²)	0.00	0.01
Male	0.00	0.01
Cruise Time (%)	0.01	0.01
Residential Road (%)	0.00	0.01
Unclassified Road (%)	0.00	0.01
Urban (%)	0.00	0.01
Maximum Deceleration (m/s ²)	0.00	0.01
Secondary Road (%)	0.00	0.01
Motorway (%)	0.00	0.01
Maximum Acceleration (m/s ²)	0.00	0.01

Table 4.8 shows that only four parameters explain 58% of variance in ABP without even accounting for variable interaction. As previous results indicated, both average speed and gradient have a significant effect on a segment's ABP. Along with these two parameters, only the percent time spent accelerating and time spent decelerating account for more than

2% of ABP variance through the S1 term. Even when accounting for interactions as in the ST term, these four variables are most significant. Second-order effects were also calculated in this SA, but there are too many results to be tabulated here. The 23 interactions with the highest S2 values were all interactions of other parameters with average speed. No other interaction terms were significant.

Because the SA accounts for the complete range of possible inputs, it may provide a better insight into how different parameters affect ABP without being biased toward the sample. As discussed previously, Sobol' sampling generated 61,400 sample driving segments. These samples were generated using the sampling process and distributions described in Section 4.6.1. To understand how ABP is affected by the model's parameters across their total input space, the empirical method from Section 4.4 is applied on the Sobol' samples and seen in Table 4.9. As in Table 4.5, the parameters in Table 4.9 are listed in order of the highest difference between groups.

Table 4.9. Average Statistics of Modelled Parameter by ABP Group

Parameter	Average Battery Power		
	Low	Medium	High
Wind Factor (km/h)	-0.25	-0.05	0.26
Gradient (%)	0.08	0.06	-0.12
Std Dev Speed (km/h)	8.24	7.57	8.65
Number of Stops	5.01	5.11	4.78
Wind Speed (km/h)	4.79	4.94	5.20
Average Acceleration (m/s ²)	0.39	0.39	0.42
Acceleration Time (%)	38.47	40.87	40.11
Std Dev Deceleration (m/s ²)	0.30	0.28	0.29
Average Deceleration (m/s ²)	-0.42	-0.43	-0.40
Male	0.48	0.51	0.51
Idle Time (%)	10.42	10.31	10.84
Unclassified Road (%)	48.77	49.41	51.69
Rain (mm)	0.49	0.51	0.49
Tertiary Road (%)	50.69	49.29	50.46
Vehicle Age (Yr)	3.62	3.47	3.45
Temperature (°C)	13.80	14.02	14.47
Motorway (%)	50.77	49.34	50.12
Usable Battery (kWh)	58.02	57.11	58.64
Maximum Deceleration (m/s ²)	-1.07	-1.10	-1.11
Urban (%)	51.04	49.72	48.98
Deceleration Time (%)	39.07	39.92	40.72
Secondary Road (%)	49.68	49.29	50.43
Primary Road (%)	50.94	50.07	49.68
Residential Road (%)	50.64	50.60	49.39
Maximum Acceleration (m/s ²)	1.02	1.04	1.04
Cruise Time (%)	51.58	52.54	52.81
Average Moving Speed (km/h)	63.53	63.70	62.85
Average Speed (km/h)	62.90	63.07	62.23
Std Dev Acceleration (m/s ²)	0.28	0.28	0.28

Table 4.9 shows there are no significant differences between average parameter values in any ABP group. This shows the extent of how much interaction is considered in the ABP model's predictions. In the ABP model, it is not enough to change one parameter value and expect ABP to change. Instead, a complex relationship between parameters means that EC can vary significantly depending on the exact combination of parameter values.

The results of this SA, chiefly the importance of average speed and its interactions, prompt further SAs within shortened speed intervals as in Asamer et al. (2016). Thus far,

most of the insight into the model has shown that EC increases at high speeds because all types of segments, from urban traffic to high speeds on the motorway, have been analysed together. The information that policymakers or EV drivers can gain from this is limited, as speed cannot be changed wildly due to safety concerns, traffic, land use context, and other reasons. In the following section, the SA is conducted within limited speed intervals, so only segments at similar speeds are compared. This way, driving behavioural differences can be analysed contextually to see which behaviours and driving environments most affect EC within small speed intervals. The SA was divided by speed due to the high interactions of other parameters with speed and the high sensitivity of ABP to speed.

4.6.3 Sensitivity Analysis within Speed Intervals

The process for conducting a SA on speed intervals is almost identical to the SA of the previous subsection. The only change is that during Sobol' sampling, the speed range is limited. The sampling was repeated for 10 km/h speed intervals from 0 to 130 km/h, and sensitivity indices were computed for each interval. The distribution of speed is assumed to be uniform throughout each 10 km/h interval. For example, each value between 0 and 10 km/h was equally likely to be sampled for the lowest speed interval. This SA allows the understanding of how parameters affect EC when the variance of EC is limited to these speed intervals.

For this section, interaction terms are not investigated. As the previous SA revealed, no interaction terms are significant to the model other than interactions with average speed. This section determines how sensitive the model is to the input parameters at different speeds. Therefore, only each variable's independent effect on the model is considered in this section. The interaction terms and total-order effects were calculated during the SA but not presented here to simplify the outputs.

The S1 of each parameter are displayed in the following graphs. The parameters are grouped by type to ensure better readability, and Figures 4.6-4.10 show the S1 of each

parameter through all speed intervals. The datapoint for a speed interval is placed at the minimum speed of that interval. For example, a parameter's S1 in the 0-10 km/h range is placed on the figure at 0 km/h.

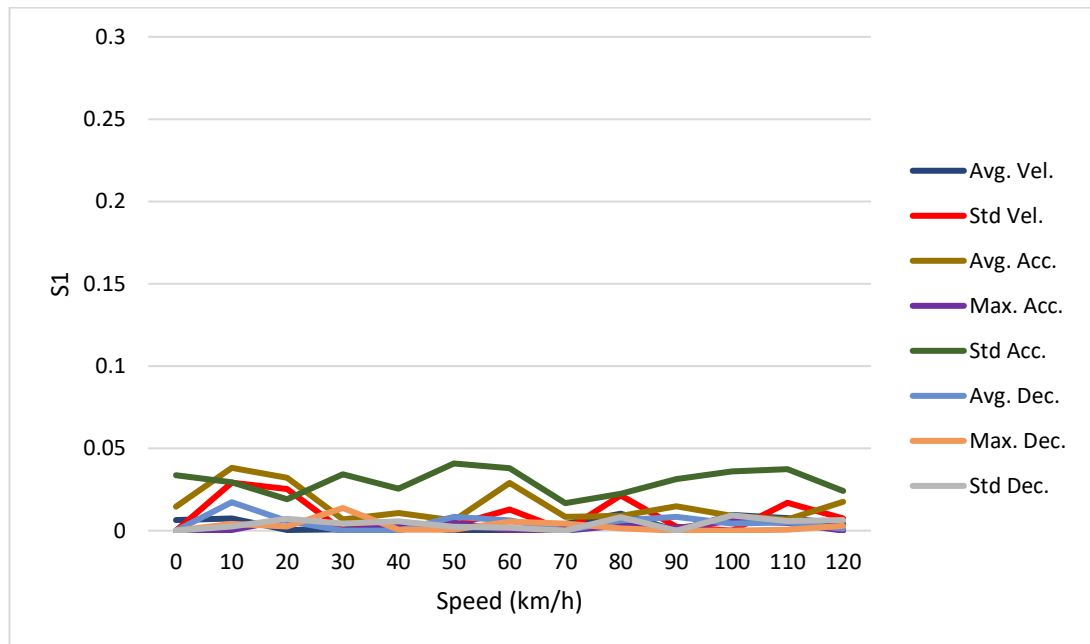


Figure 4.6. ABP Sensitivity to Driving Behaviour Variables by Speed Interval

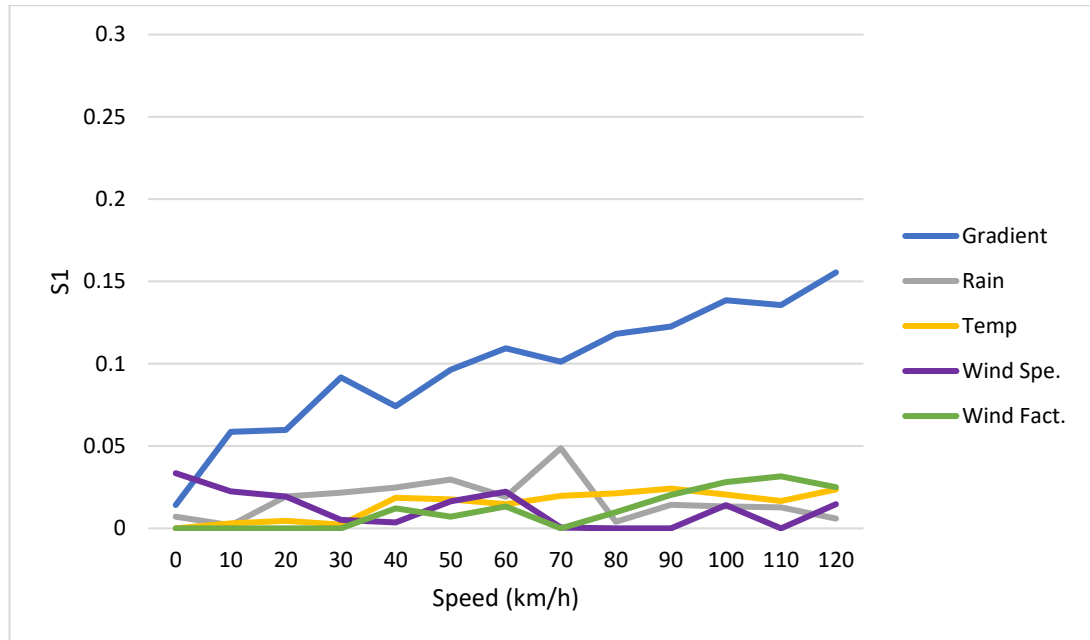


Figure 4.7. ABP Sensitivity to Weather and Slope Variables by Speed Interval

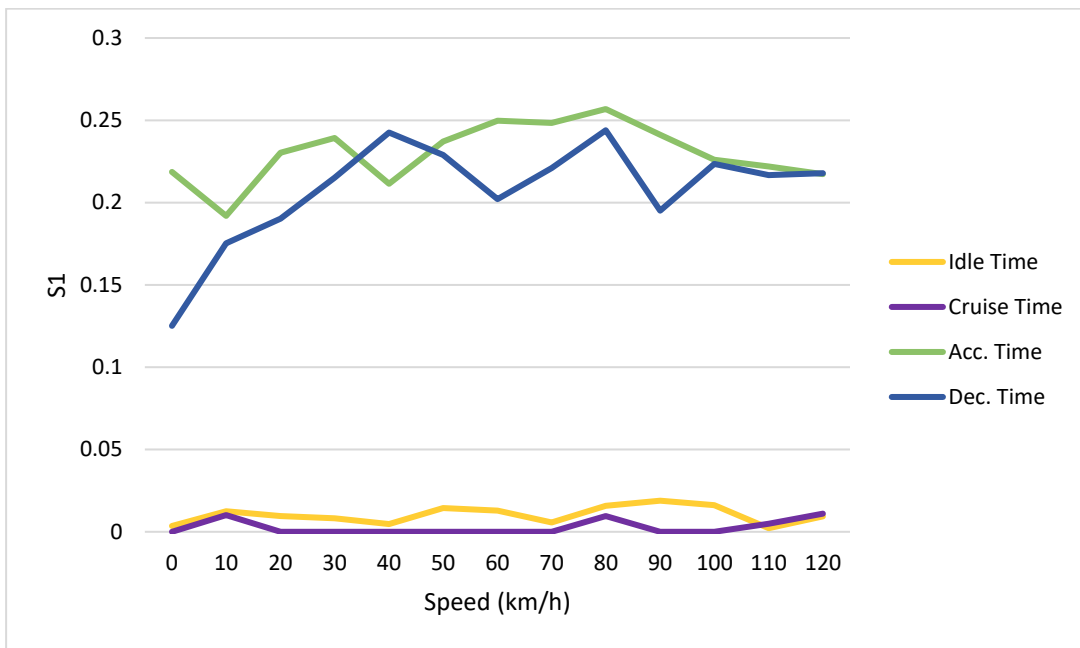


Figure 4.8. ABP Sensitivity to Drive Mode Variables by Speed Interval

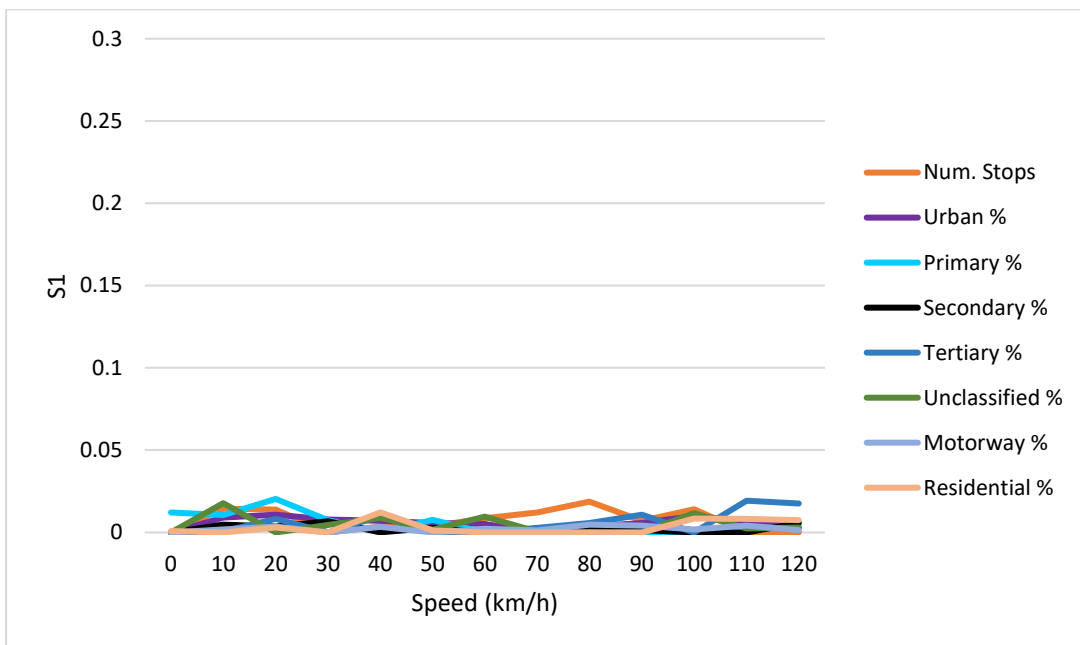


Figure 4.9. ABP Sensitivity to Road Type Variables by Speed Interval

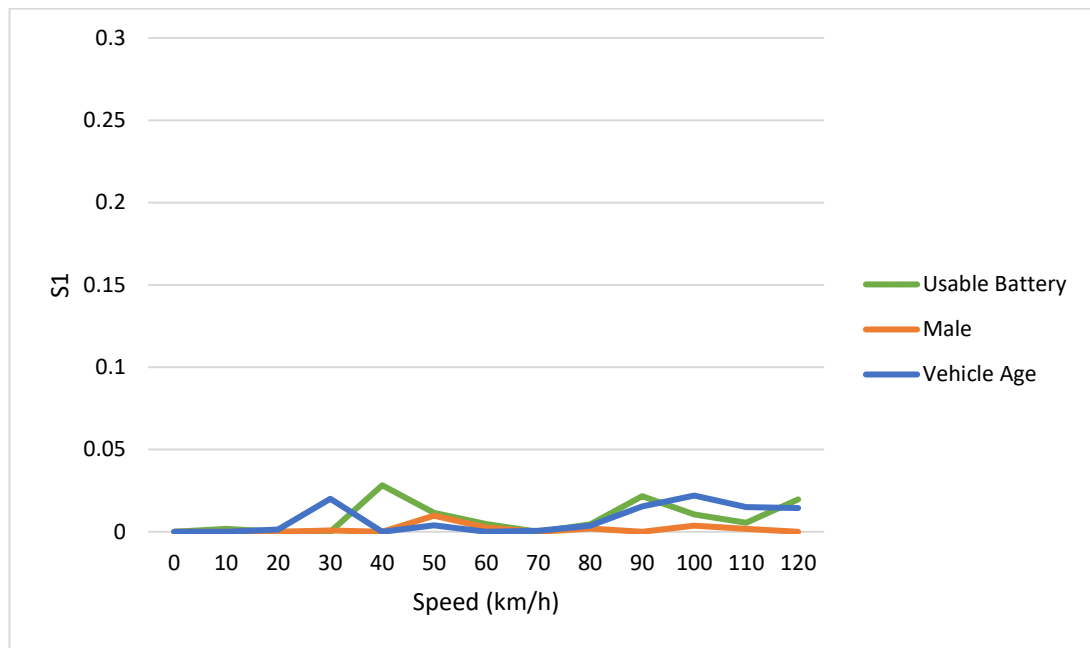


Figure 4.10. ABP Sensitivity to Vehicle and Driver Characteristics by Speed Interval

Figures 4.6-4.10 show that most variables have little effect on the model at any speed range when interactions are not accounted for. The three parameters that stand out are road gradient in Figure 4.7 and percent time spent accelerating and decelerating in Figure 4.8. The effects of time spent accelerating and decelerating are relatively stable across all speed intervals, but the effect of gradient on ABP increases as speed increases. The increasing effect of gradient can be interpreted twofold: positive gradients increase the amount of resistive force which adds additional strains to the powertrain at higher speeds and negative gradients increase the effect of regenerative braking, which increases power recovered at higher speeds.

Of the driving behaviours in Figure 4.6, the std. dev. of acceleration had the largest effect on ABP for most speed ranges. This indicates that smoother accelerating sequences can help reduce ABP by a small yet significant margin relative to highly varied acceleration. At some lower speeds, average acceleration also affects ABP meaning that ABP is more sensitive to high rates of acceleration at these lower speeds. The amount of time spent cruising or idling has little effect in this SA, as seen in Figure 4.8, because the percent time

of each driving mode in the Sobol' samples were not required to sum to one. In realistic driving segments, the drive mode percentages are interdependent, so the time spent cruising would directly affect the time spent accelerating and decelerating. The results here show how accelerating and decelerating are the drive modes most responsible for changes in ABP.

In Figure 4.7, the weather parameters show little consistent trends across speed intervals. The results do indicate that wind factor, which takes into account the direction of wind relative to the direction of travel, may explain more variance in ABP at high speeds relative to the wind speed parameter, which is a scalar parameter. This result demonstrates that future researchers should include wind direction in their data collection and modelling whenever possible. The rain parameter shows a high peak at 70 km/h, which is likely left over noise from a particular trip used to build the model. None of the road type parameters, driver characteristics, or vehicle characteristics were shown in Figures 4.9 and 4.10 to have a significant effect on ABP at any speed range.

To better understand the quantitative effect the most significant parameters have on EC, further investigation was required through another SA. By investigating the scatterplots from the Sobol' samples of each speed range, it was found that the three parameters all have a linear effect on ABP in all speed ranges. Percent time spent accelerating and road gradient both linearly increased ABP, while percent time decelerating linearly decreased ABP. For each of these parameters at each speed interval, the linear coefficient of the scatterplot between the parameter and predicted ABP was calculated. Example scatterplots with linear equations for time spent accelerating and decelerating are seen in Figures 4.11 and 4.12, with these figures showing results for the 20-30 km/h range. These coefficients were then compared to the average EC of all Sobol' samples in each speed interval through the standard relative error formula to obtain a relative effect on ABP. The results, shown in Figure 4.13, give a measurement of how much a 10% change in time spent accelerating or decelerating, and a 1% change in slope for a segment can cause an increase or decrease in ABP relative

to the average EC in that speed interval. These percent changes were chosen to showcase moderate differences in these parameters, but the effects seen in Figure 4.13 would scale with any percentage change in the parameters.

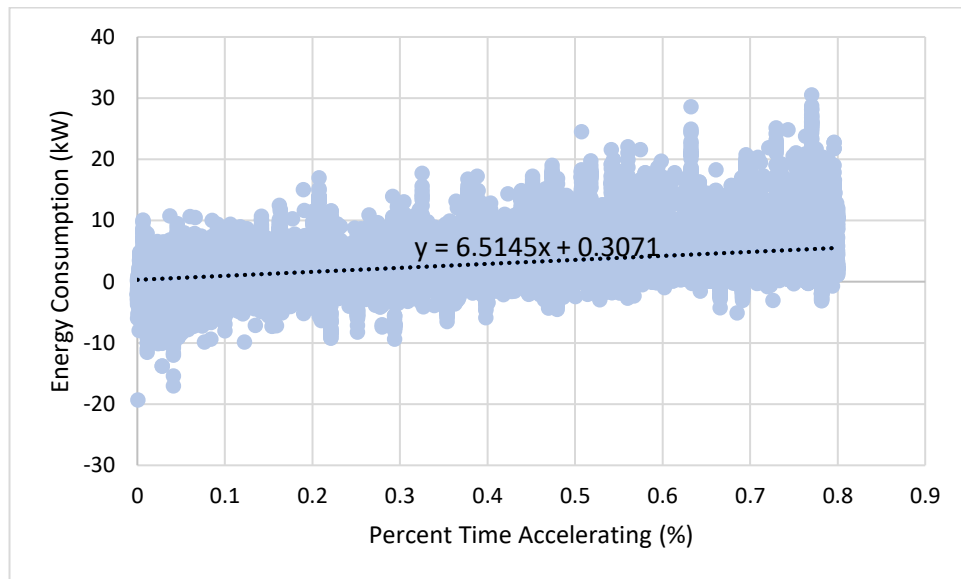


Figure 4.11. Linear Trendline between EC and Acceleration Time in the 20-30 km/h Interval

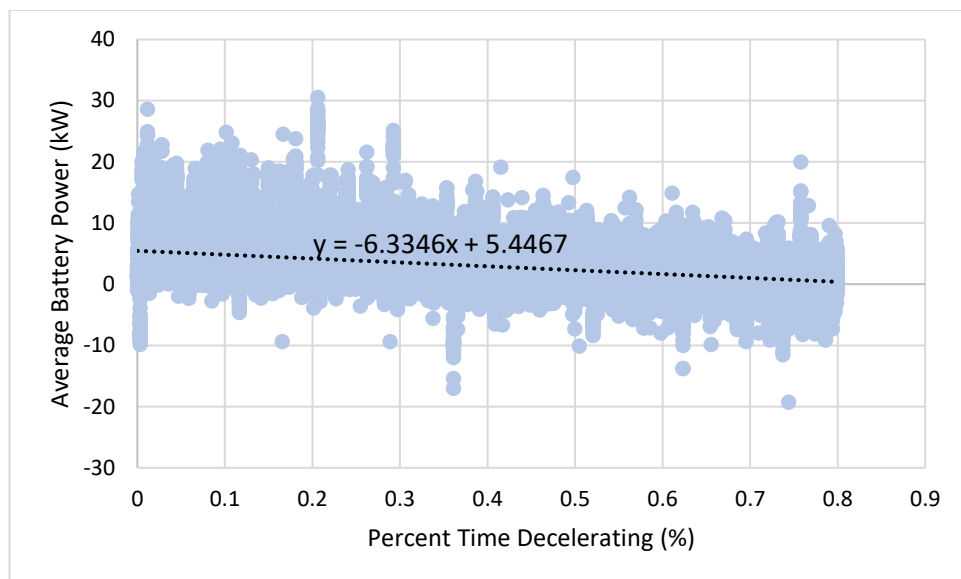


Figure 4.12. Linear Trendline between EC and Deceleration Time in the 20-30 km/h Interval

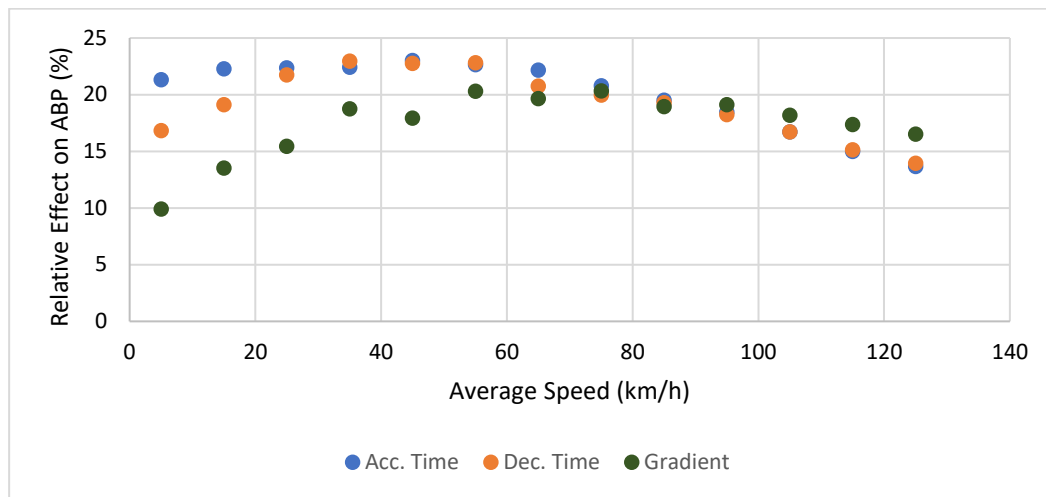


Figure 4.13. Sensitivity of ABP to Acceleration, Deceleration, and Gradient by Speed Interval

Several interesting conclusions can be drawn from Figure 4.13. First, the results show that at speeds lower than 20 km/h, acceleration has a larger effect on ABP than deceleration. This reflects the fact that regenerative braking is less effective at these low speeds. A 10% increase in the amount of accelerating and decelerating periods have more or less equal effects on ABP at speeds higher than 20 km/h. This means that drivers could significantly lower EC by reducing time spent accelerating and maximising time spent decelerating, although the effect of acceleration and deceleration rates on EC are not directly accounted for in these figures. The effect of gradient on EC is largest at moderate-high speeds, where downhill gradients can increase the effect of regenerative braking further and uphill gradients require even more power from the vehicle. The relative effects of all three parameters on ABP taper at higher speeds where air resistance is the dominant resistive force (Asamer et al., 2016).

A final piece of this SA is to consider is the effect of speed on ABP. In the SA on the entire range of speed in Section 4.6.2, ABP was most sensitive to average speed. To quantify the effect of average speed on ABP, the average ABP was calculated for all Sobol' samples in each speed interval. The average ABP of each interval are shown in Figure 4.14.

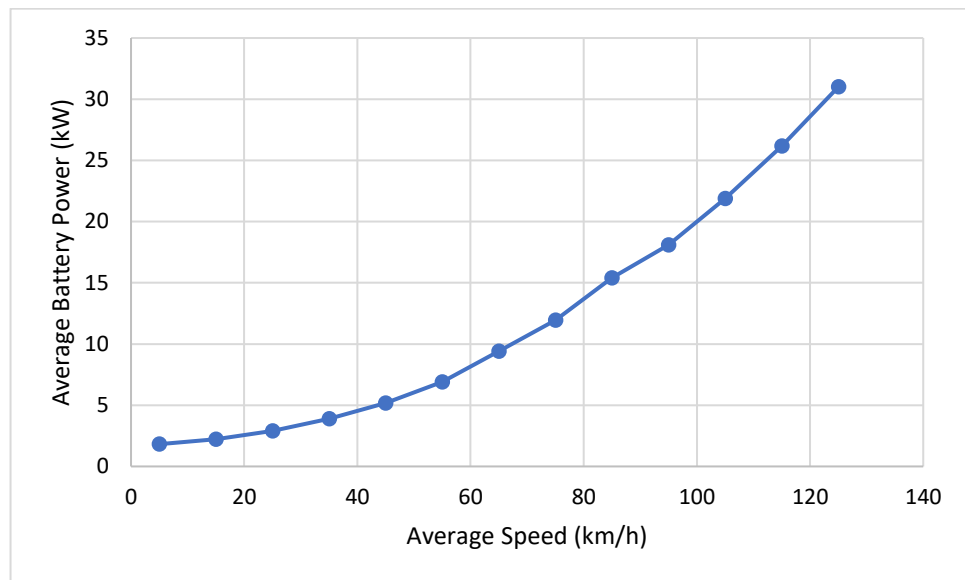


Figure 4.14. Modelled Average Battery Power by Speed Interval

Figure 4.14 shows that, on average, ABP strictly increases with increasing speed, resembling an exponential function. In this case, the ABP at zero speed would still be positive due to auxiliary power consumption. The increase in ABP from zero speed increases steadily at low speeds but increases more significantly at higher speeds where the amount of power required to overcome resistive forces such as air resistance increases. However, traveling at higher speeds can partly cancel out the higher ABP by spending less time to cover the same distance (Galvin, 2017). Therefore, it is useful to consider ECR, which normalises EC by distance. The average ECR, found by comparing the average EC to the amount of time taken to travel 1km at the average speed of each interval, for the different speed intervals are displayed in Figure 4.15.

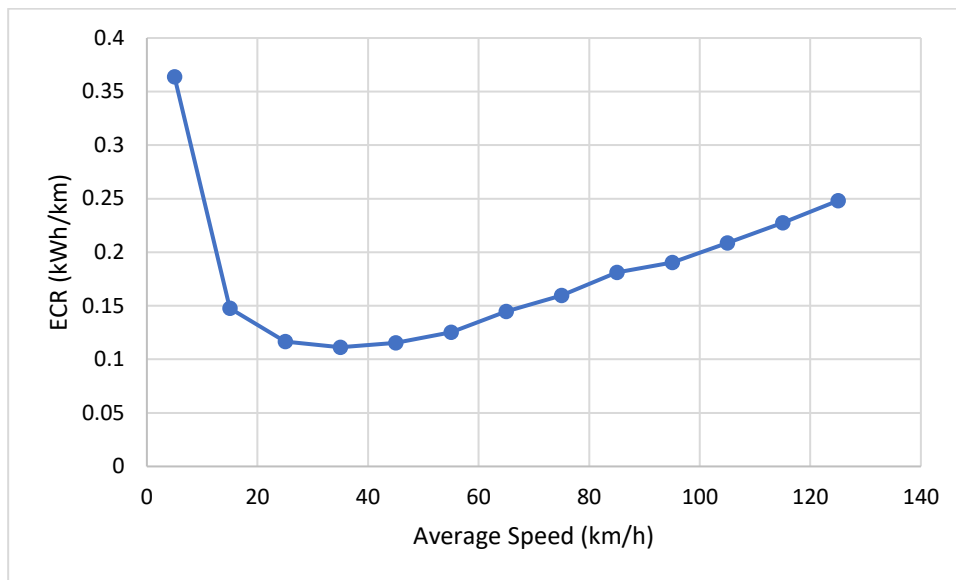


Figure 4.15. Modelled Average ECR by Speed Interval

Figure 4.15 shows a wide range of speeds, from 20-60 km/h, can be used to achieve near-optimal ECR. At speeds lower than 20 km/h, the relatively greater amount of time spent to cover the same distance at higher speeds causes the ECR to be higher. The ECR increases steadily at speeds higher than 60 km/h.

To verify these results of the SA's modelled ABP and ECR by speed interval, the ABP and ECR by speed interval of real driving data are presented here. The average ABP of real driving segments within each 10 km/h was calculated, and the results are shown in Figures 4.16 and 4.17.

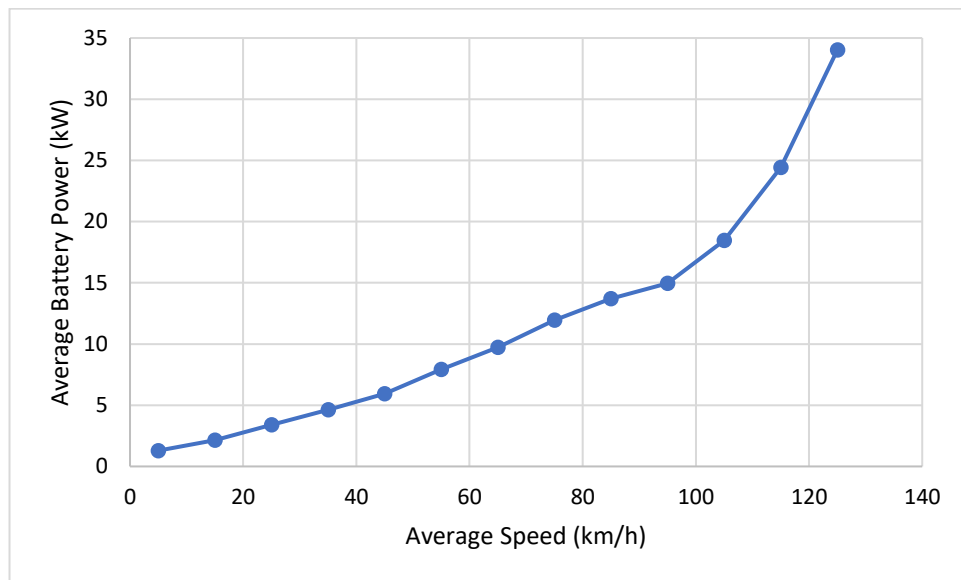


Figure 4.16. Actual Average Battery Power by Speed Interval

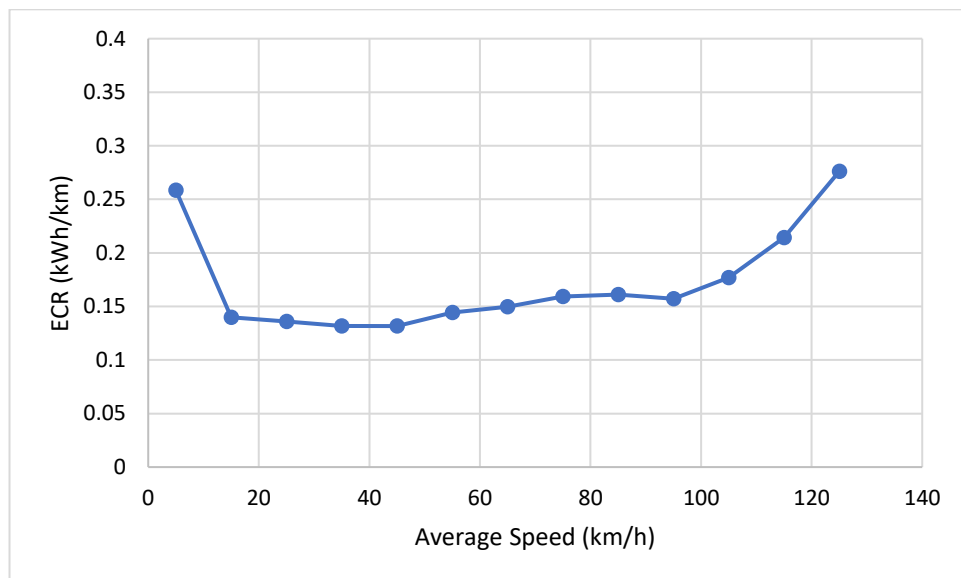


Figure 4.17. Actual Average ECR by Speed Interval

Figures 4.16 and 4.17 of actual data display most of the same trends as Figures 4.14 and 4.15 with results from the SA. Noticeably, the increase in both ABP and ECR increase even more significantly at speeds over 100 km/h in the real data. Also, the best-performing speed zone in terms of ECR shifts back to the 10-50 km/h range in actual data, 10 km/h slower than the modelled results. The extreme ECR in the 0-10 km/h interval shown in Figure 4.15 is somewhat lessened in Figure 4.17 although it is still much higher than subsequent speed intervals with higher speeds. The ECR in the 80-100 km/h range in

actual data is lower than the modelled data, indicating that a more significant increase in ECR occurs at speeds higher than 100 km/h.

The results of Figures 4.14-4.17 can inform drivers and policymakers alike that EVs use the least power per kilometre within a moderate speed range. Excessively high speeds, especially above 100 km/h, lead to higher EC to cover the same distance, meaning EVs are more likely to need to recharge on long distance trips when travelling fast. Additionally, heavy traffic conditions causing excessively slow speeds should be avoided.

As this SA relates to driving behaviour, the results are limited. The Sobol' SA proved that the complex relationships between parameters in the ABP model mean that the effect of any one parameter on ABP is dependent on the trip conditions. The initial SA showed average speed of a segment has the most influence on ABP. The pairwise interactions with the highest sensitivity indices were other parameters interacting with average speed. By repeating the SA across different speed ranges, it was determined that gradient and percent time spent accelerating and decelerating had the highest impact on ABP across all speed intervals. While this information is important, it fails to help form a driving behaviour classification method based on empirical and modelled results. What remains to be tested is the effect of changing acceleration rates, deceleration rates, and cruising speed on sample segments with all other parameters held equal. This testing of driving behaviour through synthesised segments is shown in the subsequent section.

4.7 Results of Driving Behaviour Classification on Sample Segments

4.7.1 Sample Segment Methodology

By testing different cruising velocities, acceleration rates, and deceleration rates, the effect of these driving behaviour parameters on ABP can be compared directly across similar driving segments. The abundance of samples in the Sobol' SA proved to be helpful in showing that EC can be high or low for any parameter value depending on the context. On the other hand, this meant the results did not include meaningful information on how drivers

can alter their driving behaviour to lower EC. In this section, short driving segments are constructed synthetically with varying rates of acceleration and deceleration at varying speeds so that the effect of driving behaviour on EC during a realistic driving segment can be known. The segments are composed of one acceleration period followed by a cruising period and a deceleration period so that the driving behaviour for each period can be altered and the results compared.

The methodology for developing these sample segments was simplified so that comparisons could be made directly between different rates of acceleration and deceleration and cruising velocities. A speed profile was built using an algorithm with a variable starting speed, acceleration rate, deceleration rate, and total distance for travel. The speed profile starts at the given starting speed and accelerates to a speed 10 km/h higher than the starting speed at the given acceleration rate. A distance of deceleration is calculated to know how much distance will be travelled in order to decelerate at the given rate from the cruising speed back to the starting speed. The speed profile continues at the cruising speed until the desired distance is reached, accounting for the necessary distance of deceleration. A sample segment can be seen in Figure 4.18, where a segment is built at a starting speed of 40 km/h, an acceleration and deceleration rate of 2 km/h/s, and a total distance of 500 metres.

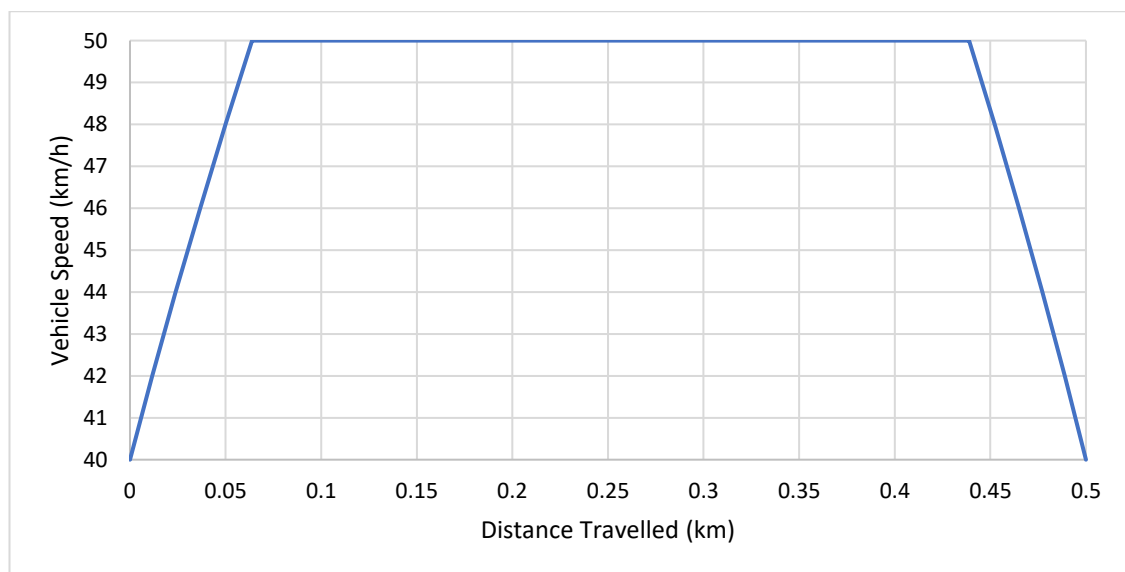


Figure 4.18. Sample Driving Segment Example

To test how different acceleration rates, deceleration rates, and cruising speeds affect EC, different values for each variable need to be tested. For this section, tests were divided into moderate speeds, referred to as urban segments, and high speeds, called motorway segments. Urban segments were built to be 500 m, while motorway segments were set to 1 km to allow for more time to accelerate and decelerate. Segments were developed for each starting speed with each possible combination of acceleration and deceleration rate. For example, a unique combination of an acceleration rate of 0.14 m/s^2 and a deceleration rate of 1.39 m/s^2 is one of the 100 synthetic segments built at each starting speed. The chosen acceleration and deceleration rates, shown in m/s^2 , are increments of 0.5 when measured in km/h/s from 0.5 to 5 km/h/s which corresponds to the most common acceleration rates seen in the data. Table 4.10 shows the values for these parameters used to create segments. The sample segments were then plugged into the ABP model, with all non-driving variables held at the mean value from the data used to build the model. The exceptions are percent time in urban areas, which was set to 1 for urban segments and 0 for motorway segments, percent time on secondary roads, which was set to 1 for urban segments and 0 for motorway segments, and percent time on motorways, which was set to 0 for urban segments and 1 for motorway segments.

Table 4.10. Parameter Values for Sample Segments

Accel/Decel Rate (m/s^2)	Segment Type	Starting Speed (km/h)
0.14	Urban	20
0.28		30
0.42		40
0.56		50
0.69	Motorway	80
0.83		90
0.97		100
1.11		110
1.25		
1.39		

4.7.2 Sample Segment Results

Because each segment within each segment type covers the same total distance, EC can be compared in kilowatt-hours as the total amount of energy consumed to cover that distance. This EC measurement is equivalent to ECR since the distance is always the same. First, results from the urban segments are compared. The EC of the urban segments compared to average speed can be seen in Figure 4.19 below.

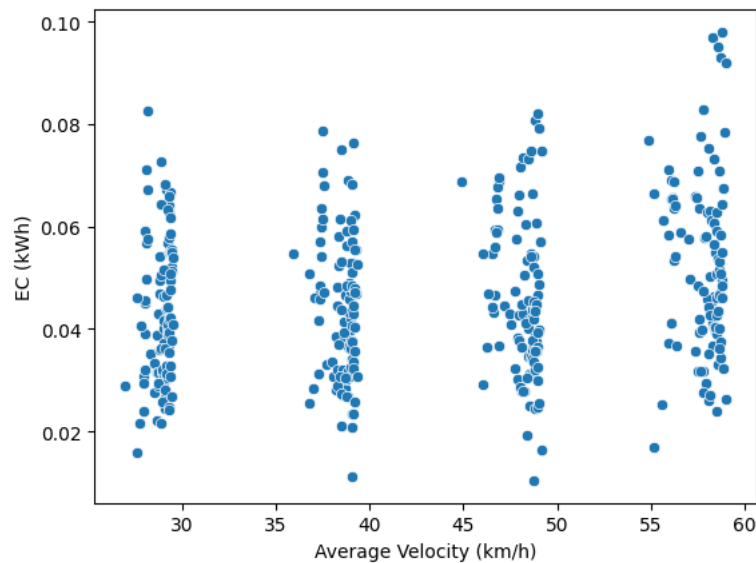


Figure 4.19. EC vs. Average Speed for Urban Sample Segments

It can be seen in Figure 4.19 that average speed has almost no effect on the range of EC observed in the urban segments. Despite the fact that velocities were tested in a large 30 km/h range, the EC does not appear to be affected by changes in average speed, except for a few outlying segments with high EC at the highest velocities. As in Sections 4.2 and 4.3, ANOVA testing is implemented here to determine whether the EC differs significantly depending on the starting speed of the segment. When the first three groups, with a starting speed of 20, 30, and 40 km/h are compared, the null hypothesis cannot be rejected with the resulting p-value of 0.385, so the mean EC of these groups are not statistically significant. When the group of segments starting at 50 km/h is added to the comparison, the p-value is found to be 0.00004, so the null hypothesis can be rejected. As a result, the mean value of

EC for the segments starting at 50 km/h is higher than the other urban segment groups when compared with an ANOVA test.

Despite this, EC shows a significant range within each group of velocities. These differences, then, must be caused by the rates of acceleration and deceleration. The effect of these rates on EC are shown in Figure 4.20 which is a heatmap where redder datapoints have higher EC and bluer datapoints have lower EC. For clarity, only segments starting at 40 km/h are shown in Figure 4.20.

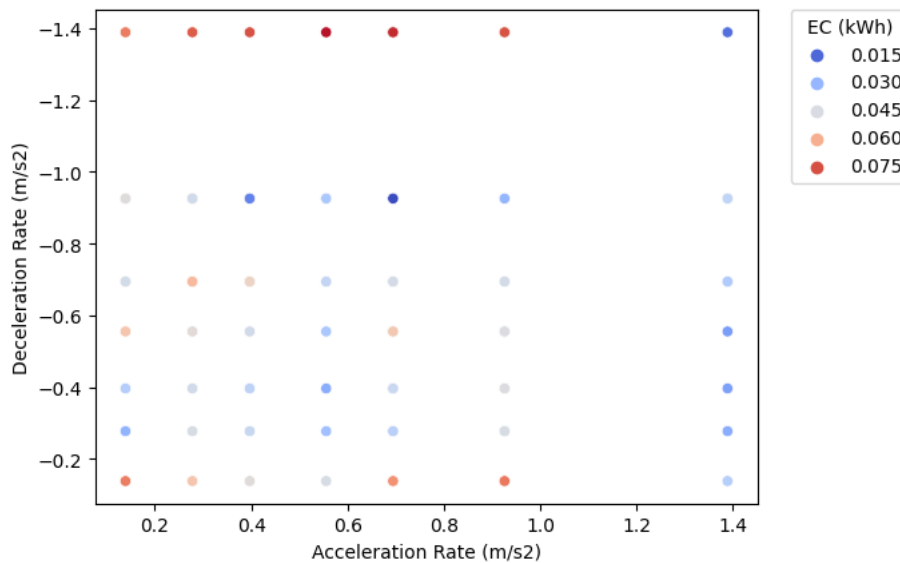


**Figure 4.20. EC by Acceleration and Deceleration Rate for the Urban Sample
Segments Starting at 40 km/h**

Figure 4.20 displays the complex relationship between acceleration and deceleration rates. The highest instances of EC occur when deceleration rate is highest, at the top of Figure 4.20. The second-highest EC grouping is found when deceleration rate is lowest, at the bottom of the figure. This detail reaffirms the importance of a steady deceleration rate to maximise the benefits of regenerative braking. Higher acceleration rates also tend to lower EC. This indicates that at lower speeds, such as in urban areas, it could be useful to accelerate sharply in order to maximise the amount of time spent cruising. Higher acceleration rates than the rates tested here would likely lead to higher EC, based on findings in the literature (Galvin, 2017; Sweeting et al., 2011). To better understand the relationships between driving

behaviours during urban driving segments, the empirical method of Section 4.4 was applied. The results are shown in Table 4.11, where only variables that differ between segments are shown. In Table 4.11 all urban segments are considered.

Table 4.11. Average Statistics of Parameters by EC Group on Urban Segments

Parameter	Energy Consumption		
	Low	Moderate	High
Segment Duration (Sec)	46.92	45.13	42.77
Average Speed (km/h)	41.32	43.11	45.57
Std Dev Speed (km/h)	2.87	2.88	2.97
Average Acceleration (m/s ²)	0.78	0.71	0.59
Maximum Acceleration (m/s ²)	0.80	0.81	0.68
Std Dev Acceleration (m/s ²)	0.04	0.14	0.14
Average Deceleration (m/s ²)	-0.59	-0.68	-0.81
Maximum Deceleration (m/s ²)	-0.65	-0.78	-0.85
Std Dev Deceleration (m/s ²)	0.10	0.15	0.07
Cruise Time (%)	0.80	0.80	0.82
Acceleration Time (%)	0.08	0.10	0.11
Deceleration Time (%)	0.12	0.10	0.07
EC (kWh)	0.03	0.04	0.06

Table 4.11 shows that an increase in average speed does minorly increase EC on average for urban segments. However, the rates of acceleration and deceleration have a more significant effect. On average, EC increases with lower acceleration rates and higher deceleration rates. The increase of ABP with lower acceleration rates is likely due to the application of ABP modelling to the specific segments outlined in this section. In this case, higher acceleration rates lead to less time spent accelerating, which, as seen in the SA, has a significant effect on EC. Additionally, Table 4.11 confirms the finding from Figure 4.20 that slower, more moderate deceleration rates are likely to decrease EC.

These parameter values would likely change if applied to segments where the amount of time spent accelerating was held constant. The segments were developed in such a way to mirror a real-world scenario where a driver needs to accelerate 10 km/h, such as a change in speed limit or an increase in traffic flow. For this reason, acceleration rate is the

independent variable, and time spent accelerating is dependent on that. The same is true for the deceleration rate.

The results from the motorway segments will now be analysed. The first comparison among motorway segments, Figure 4.21, shows the difference in EC across different average velocities.

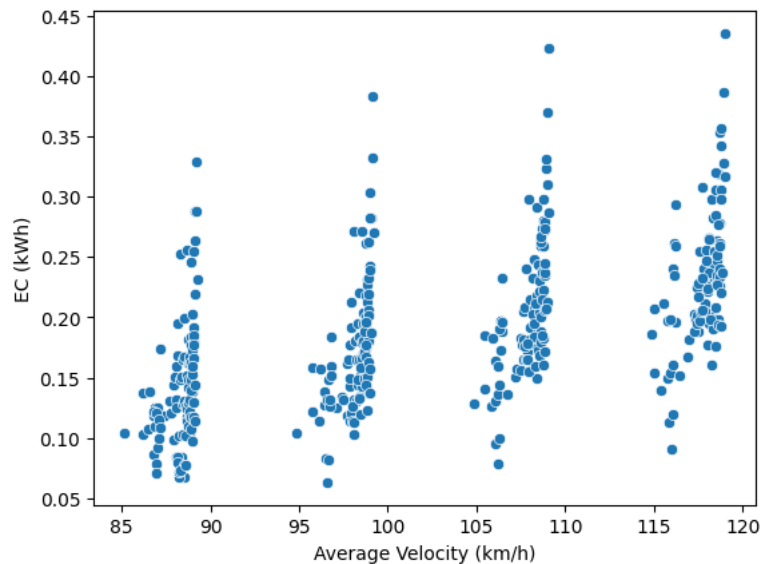


Figure 4.21. EC vs. Average Speed for Motorway Sample Segments

Figure 4.21 shows a stark difference to Figure 4.19, where almost no difference was noted between segments with different average velocities. In Figure 4.21, EC increases significantly on average as speed increases as confirmed through ANOVA tests, which resulted in p-values close to zero. This result reflects the findings of Figures 4.15 and 4.17, which showed average ECR increased markedly at speeds over 100 km/h. To investigate how acceleration and deceleration rates affect EC on the motorway segments, a heatmap is displayed in Figure 4.22. Figure 4.22 includes only segments with a starting speed of 100 km/h to improve its readability.

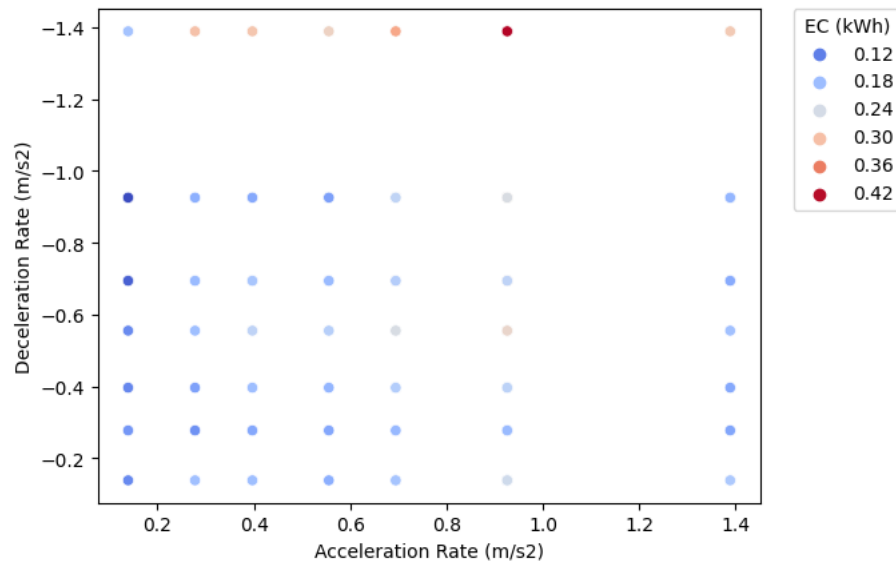


Figure 4.22. EC by Acceleration and Deceleration Rate for Motorway Sample Segments Starting at 100 km/h

For the motorway segments, the highest EC occurs when deceleration rate is the highest. Again, this is related to the energy regained through regenerative braking. EC can be lowered by decelerating at a lower rate, which increases the amount of energy recovered. Such a result has been reported in Yang et al. (2023), who also found that high deceleration rates can decrease the amount of energy recovered through regenerative braking and therefore increase EC. Khanra et al. (2018) also reported similar results, which Szumska (2025) suggests could be due to mechanical braking systems which are utilised only at higher deceleration. Unlike the urban segments, it does not appear that a low rate of deceleration negatively impacts EC. This indicates that at higher speeds, the range of deceleration rate to reach the maximum regenerative braking power is wider and more forgiving. The effect of acceleration rates on EC is less clear, although the lowest values of EC correspond with the lowest rates of acceleration. This may be due to the increase in resistive forces by air drag at higher speeds causing the energy needed to accelerate to increase significantly. It may also be due in part to the fact that lower acceleration rates result in a lower average speed because it takes more time to reach the cruising speed.

Table 4.12 is introduced here to gain a deeper understanding of how different parameters affect EC for these motorway segments.

Table 4.12 Average Statistics of Parameters by EC Group on Motorway Segments

Parameter	Energy Consumption		
	Low	Moderate	High
Duration (Sec)	38.80	35.47	33.62
Average Speed (km/h)	94.75	104.20	110.10
Std Dev Speed (km/h)	3.18	3.12	2.97
Average Acceleration (m/s ²)	0.61	0.69	0.78
Maximum Acceleration (m/s ²)	0.65	0.74	0.90
Std Dev Acceleration (m/s ²)	0.06	0.08	0.18
Average Deceleration (m/s ²)	-0.52	-0.65	-0.90
Maximum Deceleration (m/s ²)	-0.60	-0.74	-0.95
Std Dev Deceleration (m/s ²)	0.11	0.13	0.08
Cruise Time (%)	0.78	0.75	0.78
Acceleration Time (%)	0.09	0.13	0.13
Deceleration Time (%)	0.13	0.12	0.10
EC (kWh)	0.12	0.18	0.26

Table 4.12 displays different trends to the results of the urban sections in Table 4.11. For the motorway segments, average speed shows a more significant impact on EC. In Table 4.12, higher average speed increases significantly for the higher EC groups. Unlike the urban segments, average EC increases as both acceleration and deceleration rates increase. The effect of acceleration rate seems to be reversed on the motorway segments compared to the urban segments. The possible reasons for this are noted in the previous paragraph. In addition, the trend observed in Figure 4.22 that low deceleration rates do not cause high EC are confirmed as the average rate of deceleration for high EC is higher in the motorway segments than in the urban segments.

4.7.3 Application of Driving Behaviour Classification Methods to Sample Segments

As a final step in this chapter, the driving behaviour classification methods described in Sections 4.2-4.4 are applied to these sample segments to see if the resulting driving behaviour classes can help predict changes in EC. The methods are applied to the urban and motorway segments separately, and the average EC for each resulting group is calculated.

The results are shown for the limit-based method in Figure 4.23 and for the clustering method in Figure 4.24.

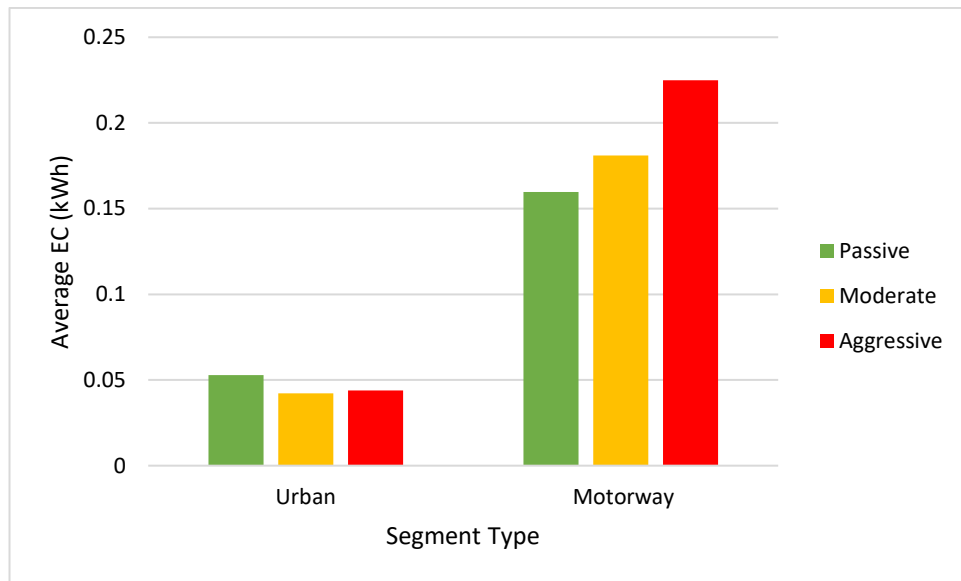


Figure 4.23. Application of Limit-Based Method to Sample Segments

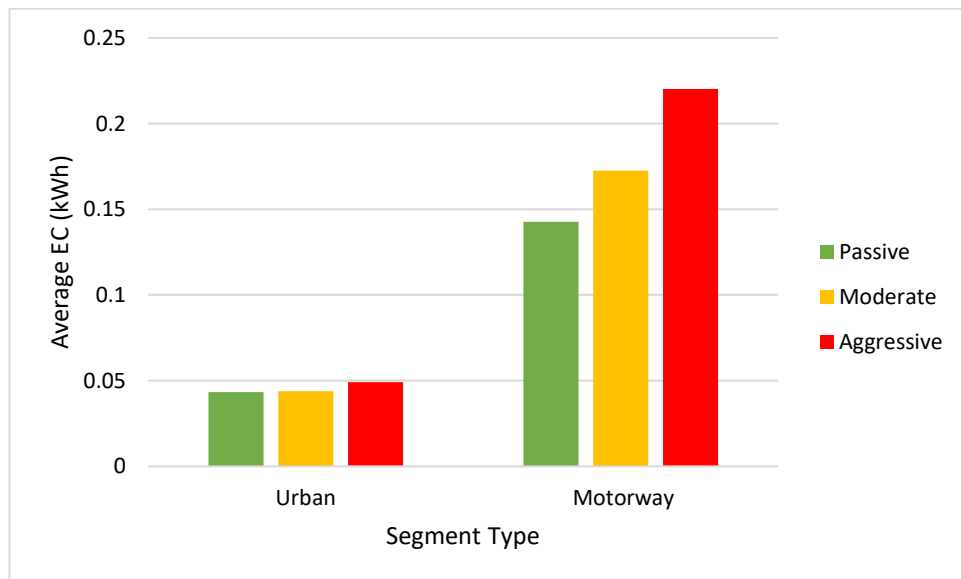


Figure 4.24. Application of Clustering Method to Sample Segments

While Figures 4.23 and 4.24 can visually represent the differences between the two models, a test is needed to determine whether the differences are statistically significant. Again, ANOVA tests are used here as in Sections 4.2, 4.3, and 4.7.2 to determine whether the mean EC is different between clusters and groups on both urban and motorway segments. The results of ANOVA testing are displayed in Table 4.13.

Table 4.13. ANOVA Test Results on Limit-Based and Clustering Methods Applied to Sample Segments

Method and Segment Type	F-Statistic	P Value
Limit-Based Urban	19.91	5.78E-09
Limit-Based Motorway	42.75	1.55E-17
Clustering Urban	1.02	3.64E-01
Clustering Motorway	33.96	5.21E-14

As seen in Table 4.13, both methods are able to classify driving behaviour on motorways in a way that average EC is significantly different for each group. This shows that, in some instances, traditional driving behaviour classification can be beneficial to predicting EC. This result is likely due to the fact that both methods include speed as part of classification. On motorway segments, where average speed was shown to have a significant effect on EC, it follows that driving behaviour classified partly by speed can succeed in showing different levels of EC.

On the other hand, the clustering method fails to show significant differences between EC on urban segments. Although the limit-based method results in significantly different means of EC between groups on urban segments, Figure 4.23 shows that the significant difference is only for passive driving. Aggressive drivers show only a 4% increase in EC compared to moderate driving in the limit-based method, and passive drivers have saved only 1% of EC compared to moderate drivers in the clustering method on urban segments. In the limit-based method, passive drivers had on average 25% higher EC than moderate drivers on urban segments. This is because segments with higher average velocities tended to be classified as passive, and these higher velocities showed a moderate tendency to increase EC even on urban segments. For the clustering method, aggressive drivers had on average 11% higher EC than moderate drivers. Again, this is likely for the same reason although the categories are different. Higher average velocities tended to have

slightly higher EC, and these high average velocities were classified as aggressive in the clustering method.

4.8 Discussion and Conclusion

In this chapter, the relationship between driving behaviour and EC was explored. First, a review of the literature covered how driving behaviour is usually defined by speed and acceleration and classified into three groups: passive, moderate, and aggressive. The methodology for three driving behaviour classification methods from the literature were discussed. After comparing real driving behaviour to EC, it was determined that a SA was needed to discover how driving behaviour affects EC across different scenarios. Finally, synthetic segments were built to compare how different rates of acceleration, rates of deceleration, and cruising velocities affect EC during moderate- and high-speed driving.

A weakness of this section is that auxiliary power consumption, such as use of heating and air conditioning, was not collected. Therefore, the effects of auxiliary consumption on EC were not able to be included in the SA or driving behaviour categorisations. Many studies have shown that auxiliary power consumption can significantly affect EC (Genikomsakis and Mitrentsis, 2017; Mądziel, 2024), but because auxiliary loads are generally optional and highly variable, the results of the SA presented in this section are valid for a base case where auxiliary power consumption is unknown. Future studies can investigate how average auxiliary power consumption affects ECR across speed intervals, or the results of Asamer et al. (2016) can be used.

The results of the SA indicate that no one singular driving behaviour parameter has a significant effect on EC in all possible driving scenarios. Therefore, it was essential to compare the effects of these parameters within similar scenarios. The scale of the ABP model being mesoscopic plays an important role in determining the effects of driving behaviour on EC. The effect of one type of driving behaviour cannot be seen directly unless it is maintained for an entire 500m segment. Instead, the mesoscopic scale of the model means

any change in driving behaviour is measured within the driving context of the total driving segment. For that reason, it was necessary to isolate the effect of driving behaviour on EC through segments that covered the same distance and same total acceleration and deceleration. Microscopic models, such as in Galvin (2017), are able to measure the finite impact of driving behaviour on EC second-by-second but can miss how one driving behaviour impacts the driving segment as a whole.

The predicted range of ECR, generally between 11-25 kWh/100km, matches closely with the range of ECR provided in the literature both through models and real data (Galvin, 2017; Liu et al., 2017; Miri et al., 2021; Xie et al., 2020), reinforcing the accuracy of the ABP model of this thesis. The results presented in Section 4.6 corroborate the findings of Szumska and Jurecki (2021), which noted that increased average speed of a DC leads consistently to higher EC. The prominent increase in ECR at speeds above 100 km/h is similar to the results from real driving data in Wager et al. (2016).

Although previous studies have attributed the fact that EVs perform better in urban areas to many different reasons including an increased number of stops allowing for more regenerated braking (Berjoza et al., 2022; Erdelic et al., 2019), the results of Section 4.6 show that lower velocities are the primary reason that EVs are more efficient in urban contexts. At speeds lower than 20 km/h, the increase in EC due to acceleration is higher than the decrease in EC due to regenerative braking, so more periods of acceleration and deceleration can lead to higher EC compared to higher speeds where the effects of acceleration and deceleration are even. The more minor effect of decelerating observed at low speeds is supported by Genikomsakis and Mitrensis (2017) and Lee et al. (2023), which observed that the maximum energy recovered through regenerative braking occurs at speeds higher than 20 km/h.

The analysis of EC on the synthesised segments showed a complex effect that changing rates of acceleration and deceleration had on EC. This result does not align with

the results of Zhang et al. (2020A), where both acceleration rate and deceleration rate were shown to have a linear effect on EC. The difference is likely due to the type of model used and type of data tested. Zhang et al. (2020A) shows that higher acceleration rates increase instantaneous EC when compared to lower acceleration rates. In Section 4.7, however, the rates of acceleration were applied to driving segments, and the EC of the entire segment was calculated. This reveals the importance of characterising driving behaviour in the context of real driving applications, where the impact of driving behaviour can be seen at a larger scale. The synthetic segments showed that higher deceleration rates may increase EC at both high and low speeds. This result echoes findings from Yang et al. (2023) and Khanra et al. (2018), which highlights the complexity of EV braking systems that is outlined as it relates to regenerative braking in Szumska (2025). Higher acceleration rates were not found to cause higher EC to the expected level, although this is likely due to the range of acceleration rates tested being lower than acceleration rates mentioned in the literature (Galvin, 2017; Sweeting et al., 2011).

Although many studies have used simple levels of aggression to characterise driving behaviour (Bar et al., 2011; Hu et al., 2022; Meseguer et al., 2017), the results of Sections 4.6 and 4.7 have shown these classification techniques inadequately account for the different effects of driving behaviour on EC at different ranges of speed. Zheng et al. (2019) and Iclodean et al. (2025) began the discussion on the relationship between EC, driving behaviour, and speed range, but provided limited information. The results of this chapter provide the first direct results on how different driving behaviour parameters affect EC across all ranges of speed. The results showed that the effects of important parameters such as acceleration rate and deceleration rate have a nonlinear relationship with EC across different speeds. Therefore, future studies should seek to understand the intricacies of these conventional driving behaviour classifications, and more advanced classifications should be developed that directly account for EC of EVs at varying speeds.

The complexity of the results seen throughout this chapter has shown the difficulties in ascribing changes in EC to driving behaviour using methodology typically discussed in the literature. As a result, the importance of analysing driving behaviour within the driving context using empirical EC data has been highlighted in this chapter. Therefore, rather than using blanket classification methods across all driving scenarios, it is essential for drivers to adapt eco-driving behaviours depending on the current circumstances. As this chapter has shown, driving behaviour can significantly affect EC at any speed, but determining exactly how any one behaviour affects speed is less clear. Simple driving behavioural changes, such as changing acceleration rates and cruising velocities, have been shown to have mixed effects on EC across different types of driving segments. Hence, this chapter has underlined the need for advanced eco-driving feedback from the vehicle that considers the current driving scenario and past EC data. Such an idea has been explored in Caulfield et al. (2014) and Sanguinetti et al. (2020), among others. The results of this chapter indicate that advances in such technology could lead to significant changes in EC, and therefore how drivers perceive the effectiveness and usefulness of EVs.

Chapter 5. Estimating Energy Consumption and Journey Travel Time Under Varied Speed Limits Using Realistic Drive Cycles

The final step of this thesis, presented here in Chapter 5, is to apply the EC model to realistic trips. By applying the model to realistic trips, the amount of EC saved by changing driving behaviour and route choice can be quantified. The analysis in Chapter 4 revealed that, above any other parameter, EC is most sensitive to speed (see Section 4.6.2). This result directly implies that the most effective eco-driving behaviour for EVs is to drive at slower speeds. Drivers may voluntarily lower speeds to reduce their energy consumption, or policy can encourage lower speeds by reducing legal speed limits.

In recent years, many countries have proposed and implemented speed limit reduction policies. Ireland is currently reducing speed limits on urban, suburban, and rural roads across the country to improve safety (Department of Transport, 2023A). Wales has also transitioned to lower speed limits in urban areas to disincentivise car travel and increase safety (Welsh Government, 2020). Spain, the Netherlands, and Belgium have also reduced speed limits on some or all urban roads to reduce the severity of crashes (EU Urban Mobility Observatory, 2020).

A review of the literature on speed limit reduction reveals no study has been undertaken to determine the effects of a reduced speed limit on EC to the best of the author's knowledge, leaving a gap in the literature to fill through the work in this chapter. In addition to analysing the effects of reducing speed limits on journey time and EC, the effects of changing driving behaviour and routes are analysed by creating a number of realistic DCs for routes found through a model of Ireland's road network. The building of DCs in this chapter is preceded by a literature review on the construction of DCs. Therefore, the following research questions are considered in this chapter:

1. How does a speed limit reduction policy affect EC in EVs? Is the effect significant enough to warrant consideration in future econometric models?
2. How significant is the effect of changing driving behaviour on EC in realistic trip conditions?
3. Does choosing a route that minimises distance lead to more or less EC compared to a route that minimises journey time?

To answer these questions, this chapter includes several steps. These steps are ordered as follows:

1. Literature Review on Lowering Speed Limits
2. Literature Review on Driving Cycles
3. Methodology
4. Results
5. Discussion and Conclusion

5.1 Literature Review on Effects of Lowered Speed Limits

5.1.1 Speed Limit Reduction Studies

Many studies have been conducted to understand the widespread effects of reducing speed limits (Abohassan et al., 2024; Department of Transport, 2023A; Folgerø et al., 2020; Fondzenyuy et al., 2024; Wooley et al., 2002). An early work on speed limits mentioned that optimal speeds are a compromise between the marginal benefit of decreased travel time and the marginal costs of safety risk and fuel consumption (Jondrow et al., 1983). Typically, modern studies use econometric models to weigh the costs and benefits of reducing speed limits regarding journey time, emissions, fuel consumption, and safety (Hosseini et al., 2015; Mamala et al. 2023; Yang et al., 2021; Yannis and Michelaraki, 2025). This section reviews the state-of-the-art literature on how reducing speed limits affects these parameters.

Implementation of any reduced speed limits requires an understanding of public perception, which is likely to be mixed (Cleland et al., 2021). Cleland et al. (2020)

differentiates between low-speed limits and low-speed zones. Low-speed zones are marked by roadways designed to encourage lower speeds through traffic calming devices like narrow lanes as opposed to simply a lower posted speed limit. Speed limits should be set depending on road type and context to minimise total cost to society (Hosseini et al., 2015). Speed limit reductions have been shown to produce lasting effects of lessening traffic volumes and reducing travel speeds, although the actual speed is generally reduced less than the speed limit due to traffic and travelling at speeds above the speed limit (Alcaraz et al., 2022; Wooley et al., 2002). A study of the USA revealed that lowered speed limits with proper enforcement could decrease travel speeds, fatal crashes, reduce fuel consumption, and limit emissions (Friedman et al., 2009).

The literature is mostly consistent in demonstrating that lowered speed limits on high-speed roads can decrease emissions of ICEVs (Cetin et al., 2018; Fondzenyuy et al., 2024; Sugiyanto et al., 2019). However, Bel et al. (2015) reported higher emissions when reducing speed limits from 100-120 kmh to 80 kmh, while Folgerø et al. (2020) found no air quality impact from reducing speeds in Norway from 80 kmh to 60 kmh. A case study of Dublin showed that lowering speed limits to 30 km/h increases pollution, creating health costs roughly equivalent to the safety benefits of decreased crashes (Tang et al., 2020). Kockelman et al. (2006) suggests that the increased operating costs (such as fuel consumption) of a vehicle driving at a higher speed limit is about half the value gained from reducing travel time.

The literature on urban speed limit reduction often focuses on health and safety aspects including ICEV emissions (Cleland et al., 2020). The EC of EVs is not considered in urban health studies because EVs do not produce local emissions, and the total emissions depend on the energy source used to charge an EV (Sasank et al., 2023). A recent review on the effectiveness of 30 km/h urban speed limits found that studies often neglect the effects of speed limit reduction on ICEV fuel consumption, and no study monitored EC of EVs

(Yannis and Michelaraki, 2025). The lack of literature on fuel and energy consumption was also noted by Lai and Wen (2020).

5.1.2 Effect of Speed on Energy Consumption

As revealed in the previous section, there is no study on how speed limit reduction affects EC. The closest analyses are based on evaluating the most energy efficient speeds through an EC model. Mamala et al. (2023) hypothesised that lowering speed limits from 50 km/h to 30 km/h would increase EC based on a model that showed the lowest ECR at 50 km/h. However, the analysis in Chapter 4 revealed a much wider range of ECR efficiency. Fiori et al. (2019) found that congested conditions consistently decrease EC in EVs, although it was determined that this was due to lower EC generally at lower speeds, a finding reaffirmed in Chapter 4 of this thesis. Fiori et al. (2019) shows that higher speeds generally lead to improved fuel efficiency for ICEVs but higher EC for EVs. Also related, Vrbanić et al. (2022) and Chen et al. (2020A) have shown that variable speed limits, an intelligent transport system that monitors traffic and lowers speeds upstream to minimise congestion, can have a positive effect on EC by limiting acceleration and deceleration events and cruising at lower speeds.

The studies on EC are generally limited to how changing speed on a roadway section can affect EC without accounting for other factors. Wager et al. (2016) showed that higher speeds on highways lead to increased EC and decreased EV range as a result. Krasopolous et al. (2017) demonstrated that variations in speed that occur in real driving conditions can significantly affect EC when compared to constant speed. Desreveaux et al. (2019) found that decreased speed decreased EC by roughly the same percentage that travel time was increased. Madhusudhanan and Na (2020) showed that cruising at a lower speed can significantly reduce EC compared to following other vehicles at higher speeds with more acceleration and deceleration events due to congestion.

Through the literature review, it was determined that no study has evaluated EC of EVs as a variable when evaluating the usefulness of reducing speed limits. However, the analysis in Chapter 4 revealed that reducing speeds can have a major effect on EC. Therefore, this chapter aims to fill this gap in the literature by testing how speed limit reduction scenarios affect EC. Although some common parameters used in econometric models in the literature like emissions and safety are unable to be tested with the available models in this thesis, the effects of reduced speed limits on EC and journey time can be compared for the first time.

5.2 Literature Review on Drive Cycles

As part of the work done in this thesis, DCs will be built to model realistic driving behaviour on the routes extracted from the road network model of Ireland. Driving cycles, also known as drive cycles, are profiles of speed versus time generally used to simulate load on a vehicle. A large part of the work done for the Horizon Europe PowerDrive project that funded this thesis revolves around DCs, so an in-depth literature review on DCs was conducted and included in this section. There is continuous research and improvement into DCs, necessitated by the extensive variety and use of these cycles in different countries and regions (Xiao et al., 2012). One reference contains 256 unique driving cycles in use across the world (Barlow et al., 2009). This is only a small fraction of the various cycles developed and used today.

DCs must be categorised to avoid confusion, but the literature is divided on the criteria that separate cycles. Even still, there exist some major distinctions. The most significant grouping is into legislative and non-legislative cycles, where legislative cycles are enforced by some governing body during vehicle testing (Wang et al., 2021). From there, the DC may be defined based on load, road type, or use case (Berzi et al., 2016). A review found that vehicle type, time of day, speed level, and engine operating mode (e.g., accelerating or cruising) may also be used to classify a cycle (Tong and Hung, 2010). Many

of these are usually combined for a single DC, such as road type, speed level, and operating mode. In some cases, multiple DCs are developed to cover different vehicle or road types.

For ICEVs the main purpose of DCs is emissions testing. With the advent of electric vehicles EVs, new DCs are needed to evaluate EC and electric systems (Oubelaid et al., 2022). All vehicles require DCs for optimising vehicle components such as the drivetrain. Jing et al. (2022) noted that as vehicle manufacturers start to emphasise high-quality over high-speed development, better DCs are necessary to create tests for the ever-increasing modern standards. The same study asserts that standardised DCs over large regions cannot represent local conditions. For this reason, DCs give consistently inaccurate results on crucial statistics like emissions. This is also why such a large number of studies have created localised DCs. Tutuianu et al. (2015) outline the characteristics of a proper DC: “it should be practical (i.e. not too long or too complicated for its execution in the laboratory); it should be repeatable and reproducible; and, most of all, it should provide results representative of the behaviour of the vehicle in real life.”

There exists no standardised method for constructing DCs. Some DCs, like the New European Driving Cycle (NEDC), are called “modal” DCs because they are derived theoretically and contain periods of constant acceleration and cruise (Oubelaid et al., 2022). Most other DCs are “transient” and match the more random nature of actual driving. In general, there are four major steps in the process of creating a DC that will be explored in the following sections:

1. Route Selection
2. Data Collection
3. Cycle Construction
4. Cycle Validation

Following these steps, the final sections of the literature review discuss localised and representative DCs, EV-specific DCs, limitations of DCs, and gaps found in the literature.

5.2.1 Route Selection

When a new DC is needed, the first step is to determine where the data will be collected. It is important that any collected data will be relevant to the region and use case desired for the final DC(s). Proper route selection, which includes determining a route is representative of the average road characteristics of a region, is paramount to the ultimate goals and representativeness of the DC (Wang et al., 2021). There is no standard practice for designating routes. Each study uses different criteria when selecting routes, often with little or no mention of concrete reasons. There are few guidelines for the selection of testing routes. Many DC developers use their best judgement and knowledge of the local area to choose a standard route (Tong and Hung, 2010).

Certain studies such as Mayakuntla and Verma (2018) use quantitative data on traffic characteristics like Annual Average Daily Traffic (AADT) and travel activity patterns to obtain a more reliable and objective representation from the sampled data. Zhao et al. (2020B) similarly developed a quantitative approach to route selection which first involved an in-depth analysis of the region's roads and travellers. Ho et al. (2014) selected routes based on traffic volume, traffic density, and surrounding land use. Yusof et al. (2023) used GIS to identify the most used routes as informed by GPS data. Yu et al. (2017) formed a series of statistical distributions and algorithms to obtain an estimate of the appropriate amount of data needed on certain road types to develop a DC. While this study pushed toward a standardised method for all DCs, the method creates an unnecessary load of heavy mathematics and research onto the front of DC development. Less intensive methods of route selection such as AADT and qualitative assessments are suitable for most purposes.

When data is collected on a large scale, routes are not selected. Instead, vehicles are monitored constantly to gather a more complete collection of a driver's behaviour. In this case, a representative DC may be constructed on data collected over a long period of time from a variety of drivers and vehicles (Ashtari et al., 2014). For the development of the

Worldwide Harmonised Light Vehicle Test Procedure (WLTP) cycle, data was collected across multiple countries across the world. To combine all of the regions into a single database, regions were weighted based on traffic volume (Tutuianu et al., 2015). Huertas et al. (2018) highlights the importance of recording samples across different times of the year to account for variations in driving patterns if data collection is not continuous. Ho et al. (2014) noted that route selection is not appropriate for large-scale data collection when instrumented vehicles are monitored over time. In these cases, the journeys are determined solely by the driver in accordance with their needs. Gong et al. (2018) recruited participants through EV forums and placed no restrictions or requirements onto their driving behaviour to ensure accuracy and fullness of data. All of the sample vehicles in this paper are private passenger electric vehicles.

Driving data may be classified during processing in cases where routes are not selected. An emerging method is using machine learning or predefined thresholds to cluster data into categories. Important features such as road type and level of service may be extracted directly from speed data, but nuances such as speeding would be overlooked (Boulter and McCrae, 2007). For example, there is an obvious difference in average speed, idle time, and cruise time between urban and rural roads that could be identifiable after the data collection period (Guo and Zhang, 2016). If some data is collected under known conditions, the results can be used to validate automatically clustered data (André, 2004). Alternatively, road type may be classified by corresponding coordinates collected by a GPS (Ho et al., 2014). The major ARTEMIS project, which set the standard for DC construction in Europe, filtered data into 15 classes of driving behaviour based on average speed, average positive acceleration, percent of time stopped, and stops per kilometre. Road type and level of service were also determined and categorised from recorded data (Boulter and McCrae, 2007). The updated WLTP cycles (Tutuianu et al., 2015) were unable to geographically determine area type (urban, suburban, or rural), so data was separated by speed.

5.2.2 Data Collection

Researchers have collected driving cycle data through two major methods: chase car and instrumented vehicle (Galgamuwa et al., 2015; Xiao et al., 2012). The chase car method involves driving a set route and tailing a vehicle to record the driving behaviour of the preceding vehicle. In one chase car study, experienced drivers were hired and told to match the speed limit and behaviour of other drivers for the purpose of data collection (de Andrade et al., 2023). With the instrumented vehicle method, on-board data is collected directly in a test vehicle. A third method combines the aforementioned methods (Han et al., 2012; Wang et al., 2021). In the last decade, GPS technology has advanced enough that the chase car method is almost totally outdated and not considered by most state-of-the-art research. Morey et al. (2009) diagnosed two major problems of the chase car method: the driver being followed may notice and change behaviour and the following driver must have sharp acceleration and deceleration events to match the chased car. A review by Tong and Hung (2010) observed that large studies with the available resources usually opt for the instrumented vehicle method. This method ensures data is collected accurately, avoids the issues of the chase car method, and importantly allows for the researchers to collect important data such as emissions and energy consumption simultaneously with speed data (Tong and Hung, 2010). Ashtari et al. (2014) shows that a constructed DC will better match real-world data when more data is collected. In this study's case, it took around 200,000 data points to surpass the representativeness of 8 standardised DCs on the locally collected data. This number is only a loose guide as the standardised DCs were collected from other locations and drivers.

Speed data is collected through GPS or OBD devices (Wang et al., 2021). DC developers choose the collection device based on judgement and seldom report the reasons. Ho et al. (2014) chose OBD over GPS for collecting speed data due to its increased accuracy. The study noted many complications that stem from using GPS such as occasional loss of

signal, resulting in stuck speeds that are recorded but not updated to match real speeds and false zero speeds, and the need to filter and otherwise process the GPS data. Following a search of commercially available tools, Romero et al. (2017) selected the Torque app and ELM327 dongle for data collection, complemented by a Garmin GPS system to obtain more accurate elevation measurements. Low cost and user programmable devices such as ones based on the Raspberry Pi card have been shown to accurately collect EV data on battery parameters such as charge, voltage, and power (Guagalango-Gomez et al., 2023).

The sampling rate of the device is also important when considering DCs. The final DC will be constructed from one speed measurement per second. Therefore, a sampling rate of 1Hz should be the minimum goal of data collection. Corti et al. (2012) shows that energy can be reasonably estimated from data collected at 1Hz, but sampling speeds greater than 1Hz can further reduce errors. Zhao et al. (2020B) opted for a 1Hz sampling rate due to the corresponding combination of high accuracy and low workload. Tsai et al. (2005) filtered speed values by averaging the 10Hz sample rate into 1 second intervals and removing any speed data more than an arbitrarily decided 5% away from the mean. Mayakuntla and Verma (2018) repeated this averaging technique without dropping outliers.

Data processing is also necessary when collecting vehicle speed in large quantities, especially with a GPS device. Bajrami and Palpacelli (2023) asserts that jerks and acceleration discontinuities from noncontinuous data produce power and torque peaks much higher than reality. Lee et al. (2010) smoothed speed data by applying a moving average technique. Wang et al. (2023) compared linear interpolation, step interpolation, and a mix of linear and cubic spline interpolation to see which method is best at filling in sparse data in the context of creating representative DCs. The results showed that cubic spline interpolation provided the most accurate results compared to simulated conditions. The limitations of cubic spline interpolation, including creating negative speed values, can be combatted by mixing the method with linear interpolation at the start and end of MTs. Achour and Olabi

(2016) simply repeated the previous recorded data for missing values. To account for small fluctuations in recorded data, Pan et al. (2015) derived thresholds for acceleration and stop periods by examining the statistics. These flexible limits minimise the effects of inaccuracies on the data. Other studies create thresholds more loosely based on observations (Gong et al., 2018). As well as speed data, GPS elevation data is prone to noise (Shankar et al., 2012). Even with accurate data, boundaries are needed to categorise data into behaviours. One study defines cruise (uniform) time as any period with speed > 0 and the absolute value of acceleration $\leq 0.15 \text{ m/s}^2$ (Jing et al., 2022).

A zero-interface method of data collection may be proposed as an alternative to a method reliant on the driver using a mobile app. Though initially attractive, this option creates several problems. Some of these limitations were discussed in Section 2.1. Vehicles and manufacturers do not use a standard language to communicate parameters through the OBD port, so decoding vehicle data completely is impossible. Yet, some common parameters such as speed could be uniformly captured across vehicle types (Gong et al., 2018). The optimal setup for data collection is vehicle specific in which any parameters of interest can be taken directly from the vehicle (Romero et al., 2017).

5.2.3 Driving Cycle Construction

Similar to the other steps of DC creation, there is no standard method to construct a DC. The two most common methods, microtrip-based and Markov chain-based, account for almost all created DCs. For this paper the distinct methods will be recorded as the microtrip (MT) method and Markov chain (MC) method. Both methods involve piecing together real-world driving data into the finalised DC. Czégé et al. (2020) provides a thorough literature review on driving cycle construction methods that demonstrates the lack of standardisation in current practice. Each study that builds a driving cycle refers to a few papers in the enormous base of driving cycle literature before deciding, mostly arbitrarily, their own methods of choice. Komorska et al. (2021) identified a third method called “segmentation”

that groups DCs according to the characteristics of the road and vehicle, but this is not different enough from the MT or MC method to warrant discussion at length. The DC may be synthetically based on data rather than real-world such as in Settey et al. (2021) and the NEDC, but again most of the literature focuses on representative DCs that are accurate to local traffic conditions. There are other methods used to create DCs that are not as widespread. For instance, Pacheco et al. (2023) proposed a novel method of creating DCs using the machine learning technique Generative Adversarial Networks that creates artificial driving data which mimics the real driving database.

While there are major differences in the methods, they share many steps and details in common. Joumard et al. (2000) strongly suggest categorising driving types and representing them in a weighted manner according to the recorded data so that the DC reflects driving behaviour, a process that has continued in state-of-the-art studies (de Andrade et al., 2023; Tutuianu et al., 2015; Ye et al., 2023). Shankar et al. (2012) classified driving behaviour based on number of stops per kilometre and congestion index. Congestion index is defined as the ratio between standard deviation of speed and average speed. Additionally, the researchers must determine the length of the desired DC regardless of construction method, but there is no standard approach for determining the length of a DC (Tong and Hung, 2010). Günther et al. (2017) developed a non-deterministic DC generator based on a simple MT method and the probability distributions of several characteristic parameters (CPs) in an attempt to automate DC construction.

5.2.3.1 Microtrip Method

The MT method was the first common way of developing DCs from real-world traffic data (Kamble et al., 2009; Xiao et al., 2012). André et al. (1994) proposed the creation of DCs based on a large database of real-world vehicle use divided into “kinematic sequences.” These kinematic sequences were a profile of speed vs. time measured between stops, now the most common definition of a MT. Each MT has a unique combination of

acceleration, cruise, and deceleration modes (Kamble et al., 2009). While this is the popular definition of a MT, many studies divide data differently. Nouri and Morency (2017) studied 8 ways of defining MTs presented here in order of their final ranking: 1. Fixed Distance (250m) 2. Intersection 3. Stop sequence 4. Fixed time (40s) 5. Event bin based on local acceleration 6. Fixed speed ranges 7. Event bin based on MLE 8. Event bin based on k-means method. The database used for constructing these driving cycles was very uniform, so these results must be taken with caution and not as comprehensive.

Amirjamshidi and Roorda (2015) identifies five common steps of the MT method:

1. Define assessment measures or CPs to describe a DC
2. Calculate these assessment measures on the collected data to find the target statistics
3. Develop candidate cycles by stitching together MTs randomly or semi-randomly
4. Calculate the assessment measures on the candidate cycles to find the test statistics
5. Identify the candidate cycle whose test statistics are closest to the target statistics or meet the determined threshold for accuracy

Each step allows freedom for the developer to use their best judgement. Some studies choose to randomly attach MTs together and choose the best candidate cycle, but this process is extremely time- and computation-heavy. Instead, it is better to selectively choose MTs that lead to a final DC that closely resembles the target statistics in a quasi-random selection (Xiao et al., 2012). Yu et al. (2010) limited the quasi-random construction to 50 candidate cycles to further limit computational time. Candidate cycles are generally created iteratively until a predetermined amount are formed (Han et al., 2012; Yu et al., 2010) or until the desired target statistics thresholds are met (Arun et al., 2017; Ho et al., 2014). Chen et al. (2019) made a single optimised DC for the number of MTs ranging from 6 through 30, and the cycle that best represented driving data was made from 21 MTs.

5.2.3.2 Markov Chain Method

Lin (2002) initially proposed a basic MC method to stochastically construct DCs. A sequence with the Markov chain property determines a following state depending solely on the previous state (Lin and Niemeier, 2002). Wang et al. (2019) summarises the process of constructing a driving cycle based on the MC method. The driving cycles are combined by randomly selecting fragments of recorded trip velocities and combining fragments that have a high probability of following the previously chosen fragment. The selection of the subsequent fragment occurs through a transition probability matrix (TPM) that includes the probabilities of all possible next steps. The probability of transitioning from the current to the next driving cycle fragment is an important factor of construction and prevents unrealistic random cycles. As a final precaution, the constructed driving cycles are validated by comparing relevant parameters to the population data as in the MT method. Bishop et al. (2012) discusses in depth the creation of a transition probability matrix and a standard algorithm for the Markov Chain method that will not be which will not be reiterated in this thesis.

Pan et al. (2015) utilised a 7x7 TPM after clustering fragments by average speed into 7 bins. Gong et al. (2018) implemented a two-level TPM scheme with the first determining the next action (cruise, acceleration, deceleration, or idle) then determining the action within that state. In some studies, the start and end of a Markov DC are artificially added in order to set a foundation for the Markov process (Pan et al., 2015). The length of each fragment is up to the researcher; Komorska et al. (2021) used groups of MTs as inputs and outputs of the TPM.

5.2.4 Characteristic Parameters

Before delving into cycle validation, it is necessary to first examine the CPs that DCs are built to replicate. While the final DC is only a measurement of speed versus time, there are many CPs, also called assessment measures, to ensure that the DC represents the intricacies of driving behaviour captured in the complete dataset (Berzi et al. 2016). A

rational set of CPs is necessary in the creation of DCs to define driving behaviour, construct candidate cycles, and determine the representativeness of the chosen cycle (Tong and Hung, 2010). More modern approaches use machine learning algorithms to reduce dimensionality in the CPs or cluster driving data to group behaviour. For example, Jing et al. (2017) used principal component analysis (PCA) to reduce parameters from 15 to 4 principal components while retaining over 80% of explained variance. In a comparison of machine learning clustering methods, K-means and fuzzy C-means were shown to deliver good results while retaining computational efficiency (Abdel-Wahed et al., 2023). Ye et al. (2023) utilised self-organising map and support vector machine algorithms to improve MT clustering beyond the level of K-means.

The most common CPs are kinematics-based and describe the movement of a vehicle over time. An overwhelming amount of the literature looks only at these parameters, and models rarely incorporate the end-goals of fuel consumption or emissions (Huertas et al., 2018). Occasionally additional data is recorded but unused in the construction process (Nouri and Morency, 2017). Gebisa et al. (2021) lists 38 parameters from the literature, all kinematics- or time-based. Komorska et al. (2021) found that the local conditions have been characterised in the literature by CPs such as the average, maximum and minimum values and standard deviations of speed, acceleration and delay. The same study lists 22 common kinematic parameters including one that estimates energy consumption as a function of speed and acceleration. Bishop et al. (2012) classified these CPs into time-related, speed-related, and acceleration-related to draw further distinctions. Kinematics-based CPs may be enough to predict load such as energy consumption. Huertas et al. (2019) suggests that mean speed, idling time percentage, and the speed-acceleration frequency distribution (SAFD) are the most common CPs in the literature. Mieth et al. (2022) identifies the importance of time frame-based load analyses for the optimal design of vehicle parts. Some DCs are created without validating the distribution of kinematic CPs with measurements such as SAFDs.

Acceleration has the most effect on energy consumption of any kinematic measurement (de Andrade et al., 2020). Huertas et al. (2018) developed a Minimum Weighted Difference of Characteristic Parameters methodology to create a representative driving cycle of emissions based only on a speed-time profile. A parameter called road power can be calculated from instantaneous vehicle speed to artificially account for the wheel power requirements of a vehicle when categorising MTs (Ashtari et al., 2014).

External conditions are also key to understanding the load on the vehicle to be simulated by the DC (Ashtari et al., 2014). A common CP is the road's slope or gradient, although few studies include its effects (Bishop et al., 2012). Han et al. (2017) focused distinctly on gradient in addition to speed when developing a cycle for a mountainous military area. Zhang et al. (2019) obtained elevation and road type details in the analysis phase from GPS location data through various third-party software. Salihu et al. (2023) found that emissions increase drastically even at low positive slopes as drivers have higher acceleration and deceleration as the road slopes. Bishop et al. (2012) discovered that including road grade as a CP created candidate cycles that matched the target statistics the same or better than without grade.

Temperature is another important CP when evaluating the energy consumption of a vehicle. (Galgamuwa et al., 2015; Zhang et al., 2023). The ARTEMIS project found that increasing road gradient and vehicle loads increased emissions, while outside temperature, humidity, and other external factors affected emissions output to a lesser degree. There is also a complex interaction between various parameters, such as how speed can influence excess energy use from air conditioning systems (Boulter and McCrae, 2007). External conditions such as temperature can heavily affect an EV's range and may be responsible for miscalculations in the vehicle's displayed range. This effect results in more instances of "range anxiety" where EV drivers display different behaviours than ICE drivers in fear of running out of battery (Pan et al., 2017, Settey et al., 2021). Weldon et al. (2016) found the

same phenomenon in Ireland, concluding that EV drivers prefer short trips and frequent charging. Boundary conditions can be imposed on the DC to ensure CPs are captured other than the final speed-time DC; these conditions may be on gear shifting, atmospheric conditions, tyre pressure, vehicle conditions, cold start conditions, and many other factors that affect vehicle performance (Berzi et al., 2016).

Many parameters have been shown to have a direct link to emissions and energy consumption data that DCs test. Berzi et al. (2016) list driver capabilities, driving context, traffic conditions, and ambient temperature among these factors. However, some of these may only be possible to test in on-road conditions instead of DCs. Xie et al. (2020) showed the extreme difference in energy consumption and range of EVs in congested areas compared to areas of free-flow speeds. Oubelaid et al. (2022) highlighted that load varies heavily due to road slope, wind, and road type. Settey et al. (2021) found that the load of a vehicle has little effect on EV battery usage in a basic, synthetic driving cycle. Liu et al. (2023) used Markov chain techniques to stochastically factor passenger load into the DC of a bus due to the excess energy consumption that occurs with more passengers. A recent PhD thesis studied the emissions effects of various parameters such as altitude, temperature, grade, resistive forces, and driver aggression (Mouzouras, 2022). However, no such in-depth research has studied the effects of such parameters on EV energy use in real-world DCs. Liu et al. (2016) found significant impact of travel distance, speed, time of travel, auxiliary loads, and road grades on energy consumption. Emissions are more sensitive to various parameters depending on driving type. For example, in urban driving emissions are sensitive to stop frequency and duration while in motorway driving emissions are most sensitive to accelerations at high speeds (Boulter and McCrae, 2007).

5.2.5 Driving Cycle Validation

The final step in creating a driving cycle is to validate the chosen candidate cycle meets the criteria determined for the finalised DC. Most studies validate the final cycle using

some relative error term to compare the test statistics to the target statistics of the predetermined CPs such as in Arun et al. (2017). As with each previous step, there is no standard method for validating a DC. An early study by Lin and Niemeier (2002) compared test and target statistics for average speed, maximum and minimum acceleration, root mean square speed and acceleration, and percent of idle time. Chen et al. (2019) contends that differences between DCs and real-world data stem from suboptimal validation methods. The review by Tong and Hung (2010) suggests that validation criteria should minimise differences in CPs related to the goal of the DC, minimise difference in transportation engineering related CPs like modal distributions, and be reported as a relative difference between test and target statistics. Berzi et al. (2016) reported that 0.5-1% of cycles generated in the study fit the 5% threshold of similarity for 11 of 12 CPs. This is due to the inherent variability of the random selected MT method even with clustering, and it proves the necessity of appropriate validation criteria. In addition to a standard parameter comparison, Zhang et al. (2019) employed a statistical Kolmogorov–Smirnov test to compare the distributions of speed, acceleration, and gradient of the constructed DC to the original data. DC developers with unique construction methods also might require unique validation methods such as in Huertas et al. (2019).

5.2.6 Comparing Construction Methods

After reviewing the cycle construction process, it is natural to wonder why there is no standard practice. Different DCs require drastically different energy consumption even from the same vehicle, so creating the right DC should be a scientific process (Singirikonda and Obulesu, 2023 and Suryavanshi and Ghanegaonkar, 2023). Developers of new DCs generally compare the new cycle to legislative cycles rather than other modern DCs. Because of this, there is a lack of standardisation across the deep pool of literature on the subject. Comparisons are difficult across studies due to the fact that each author defines CPs and methods based on their needs (Pereira et al., 2020). However, several studies have compared

methods used on the same dataset with the same guidelines. This section focuses on comparing construction methods, especially the MT and MC methods.

Zhang et al. (2019) found that a DC derived from intelligently selecting MTs (quasi-random MT method) matched CPs better than the MC method, which in turn was more representative than randomly selected MTs after 1000 iterations. However, the MC method better matched the distributions of speed, acceleration, and road slope.

Yang et al. (2023) showed that a scientific approach to the creation of DCs can result in more accurate prediction of fuel consumption than the common MC method. The method in this study combined a genetic algorithm with Metropolis-Hastings sampling to optimise cycle construction, which more closely matched the data's CPs than the MC method.

Chen et al. (2022) compared five construction methods on accuracy to real data, reliance on large amounts of data, and reproducibility of the final cycle. The five methods evaluated were random selected MTs, MT method with PCA, MT method with fuzzy c-means clustering, MC method, and an optimisation-based algorithmic method. The results showed that the methods which included machine learning, PCA and clustering, gave more accurate results than the standard MT and MC methods, but they required a larger database size to be effective. The optimisation-based method provided the most accurate results overall because the method is built to choose the most representative data as the final cycle. In terms of reproducibility, PCA and clustering provided the best results, the optimisation-based third, random selected MTs fourth, and finally MC provided the least reproducible DCs.

Huertas et al. (2019) developed a method of comparing driving cycles to each other and sample driving data. Two tests were developed: one that compares test and target statistics of CPs and another that compares the SAFD. The study did not use the same CPs that validated candidate cycles in their evaluation because by design they are within a set threshold. So, a high degree of importance was placed on the speed-acceleration probability

distribution relative difference. Huertas et al. (2019) showed that constructing 100 DCs is insufficient for obtaining average statistics; 500 DCs provide a better evaluation. Ultimately, MT and the study's unique fuel-based construction methods significantly outperformed the MC method in representing local driving patterns. Some DCs such as the fuel-based method of this study simply select an entire test run as the final DC. Generally, the selected test run is the most representative of the averages of all runs. While being statistically representative of the dataset, a DC from just one test run may miss out on significant driving behaviour captured in other drives (Tong and Hung, 2010).

Nouri and Morency (2017) supposed the Markov chain method is regarded as superior to microtrip-based methods in the literature, but that claim was not substantiated directly. Yu et al. (2010) developed a DC replicating the dataset's average vehicle speed profile (defined as the instantaneous power per unit mass of a vehicle) which predicted emissions much better than a DC based on common CPs such as idle time and average speed. The ARTEMIS project similarly found that DCs based on SAFD matched emissions closer than a model based on various kinematic CPs (Boulter and McCrae, 2007). Within the MC method, candidate cycles can be improved by increasing the number of Markov repetitions at the cost of increased computational time. Accuracy was also improved by increasing the number of speed modes in the TPM, but longer cycles were less accurate than shorter ones (Bishop et al., 2012).

Gebisa et al. (2021) serves as a kind of meta-analysis of the current state of DC construction with regard to emissions testing, but no conclusions are made when comparing methods. Even when the average statistics of DCs are similar, a simple visual comparison can display the difference between the MT and Markov Chain methods. Nouri and Morency (2017) emphasise that although the construction method does not affect the actual driving behaviour but rather how it is represented and communicated in the DC. For example, the Markov Chain method can produce much longer MTs than those stitched together in the MT

method from the same data (Pereira et al., 2020). This is not a generalised result. Zhao et al. (2020A) found that the MC method produced DCs closer to test data than DCs constructed from various clustering algorithms. Chen et al. (2019) developed an optimisation-based EC method that matched the dataset better than a traditional MC method, but the MC DC created a reasonable match with much less computation time. Jing et al. (2017) recommends linear discriminant analysis over data clustering for small samples to reduce time consumption and ensure the data is processed in the desired manner. A DC constructed with a discriminant function also resembled real-world data closer than the clustering method. Ali and Asfoor (2023) found their novel method of construction, piecing together the average speed of collected data every 10 meters to form a cycle, delivered similar relative errors of average and standard deviation of speed to the MC method. Pan et al. (2017) used fuzzy C clustering analysis and kernel principal component analysis to find parameters that enable the prediction of EV range through predictive driving cycles.

5.2.7 Localised and Representative Driving Cycles

As previously stated, researchers who develop their own localised driving cycles generally compare the new cycle to the legislative standard and find significant differences (Amirjamshidi and Roorda, 2015; Ashtari et al., 2014; Berzi et al., 2016; Han et al., 2017; Ho et al., 2014; Günther et al., 2017; Guo and Zhang, 2016; Kamble et al., 2009; Mayakuntla and Verma, 2018; Tsai et al., 2005;). Lin (2002) concluded that regional differences in driving behaviour are great enough to have distinct influences on emissions, and numerous local DCs have emerged since to confirm that fact. Despite being used as legislative cycles in some Asian countries, European DCs generally have higher average speeds with low acceleration, resulting in smoother DCs than the rest of the world (Tong and Hung, 2010). de Andrade et al. (2020) found that the average localised DC has lower average speed and acceleration compared to the legislative FTP-75 DC. de Andrade et al. (2023) found that

motorcycles and cars require distinct DCs. Different cities can be represented by the same DC for urban driving, but highway driving is distinct enough to warrant separate DCs.

A review of localised DCs found that the biggest problem of most local DCs is that they are developed without a substantial database to draw from. Additionally, not enough planning is put into the proper execution of each step of the cycle construction process (Gebisa et al., 2021). Broadly, the use of real-world data to develop DCs is seen as the most important way of ensuring DC representativeness and ensuring the resulting DC is useful (Arun et al., 2017). Even a simulated DC developed from a micro-simulator may be concerned about representativeness. In this case, the simulation may be run multiple times to gain variety in the data and therefore enhance the confidence that it is representative of real-world conditions (Amirjamshidi and Roorda, 2015).

5.2.8 EV Specific Driving Cycles

While DCs were initially developed to standardise emissions testing, research in the last decade has focused on creating DCs that evaluate the energy consumption and manufacturing of EVs. Generally, literature on EV DCs still focuses on time series data without looking at EV-specific parameters. EVs introduce new parameters such as range limitation, relevance of auxiliary battery usage, regenerative braking, and perception of performance (Berzi et al., 2016). Energy economy and overall efficiency of EVs are also commonly used as evaluation metrics (Jing et al., 2022). Both road infrastructure and driving behaviour have major effects on the energy consumption of an EV (Suryavanshi and Ghanegaonkar, 2023). Waldowski et al. (2011) proposed generating individual driving cycles to help new EV consumers to find the correct drivetrain for purchase based on the potential use cases.

Zhao et al. (2020A) found that internationally regulated cycles like the NEDC may overestimate EV driving range and underestimate energy consumption. A comparison with ICE DCs showed that EV DCs have higher maximum and average accelerations, but longer

time spent accelerating. The regenerative braking capabilities of EVs can also be seen in the DC differences, which shows that EV cycle data should be collected from drivers familiar with EVs. This study sets a good framework for future localised EC DC development (Ye et al., 2023). A study on the WLTP cycle found that a drop in an EV battery's state of health from 100% to 80% can decrease the EV's range by 23%. Additionally, the number of passengers and use of air conditioning will have a major effect on energy consumption (Mansour and Raeesi, 2023).

Chen et al. (2020B) noted that just like ICEVs, there is no standardisation in the practice of developing DCs for EVs. This study developed a DC based on average energy consumption, but the kinematic parameters of the finalised DC were remarkably similar to the dataset's average. In this study, the energy consumption was strongly based on kinematic parameters derived from vehicle power equations, so the similarity is not surprising. Brady and O'Mahony (2016) delved deep into the specific needs of an EV driving cycle, but the measurements of energy use were derived from a simulation rather than real-world vehicle data. Zhang et al. (2019) merged energy parameters such as voltage and current data with speed data in order to create an EV-specific DC. Shankar et al. (2012) asserts that it is possible to predict power usage from obtaining distributions of CPs such as speed, acceleration, and power, but there needs to be an understanding of environmental factors such as road type too. That study recorded battery power (power measured at battery terminal), motor bus power (motor current x motor voltage), motor output power (motor speed x motor torque), and wheel power (derived from resistance due to wind, gradient, tyre friction, inertia resistance, vehicle mass, and speed) to measure drivetrain efficiency. Lee et al. (2011) developed an EV DC based on kinematic CPs, and the resultant DCs matched average energy consumption of the real-world data well. In a simple test of several DCs, Xie et al. (2020) found that the DC's average running speed, length, and deceleration frequency have the largest effect on EV energy consumption. Chen et al. (2019) constructed an EV DC

based on CPs that dictate energy consumption. Jing et al. (2017) used power consumption per 100 km to verify the constructed cycle represented real-world EV expected loads.

The driving behaviour of an EV driver can be significantly different than of a driver of an ICEV. Waldowski et al. (2011) found that driving behaviour has a significant impact on the efficiency of EVs. Berzi et al. (2016) suggests that EVs are designed to communicate with the drivers on factors such as regenerative braking so drivers can change their behaviour advantageously. Regenerative braking has extreme potential to improve energy economy in everyday DCs (Ruan et al., 2016). Regenerative braking has been shown to have major effects on an EV's energy consumption, with the greatest benefits occurring in urban areas with heavy traffic (Tuan et al., 2023). Zhao et al. (2019) shows an enormous gap between international DCs and real-world driving in terms of driving range and energy consumption. It should be noted that these DCs were constructed for ICE vehicles, however. Factoring auxiliary energy consumption and regenerative braking are essential to understanding the performance of an EV. Auxiliary energy consumption such as air conditioning can account for over 30% of the well-to-tank emissions of an EV (Rangaraju et al., 2015). The method of utilising DCs to predict energy use and driving range of EVs has become an emerging topic in the field of DCs. K-means clustering has been used to predict driving behaviour in real time by comparing the current driving behaviours to clustered previous driving data (Hao et al., 2023).

5.2.9 Limitations of Driving Cycles

There has been significant criticism levied at DCs over time, especially at the shortcomings of legislative cycles. Amirjamshidi and Roorda (2015) believe there is a need for more specific DCs to analyse emissions than the legislative cycles used because of the number of factors that affect emissions. Different cycles and procedures will result in different emissions outcomes even from the same vehicle. Manufacturers may optimise parts for the specific driving cycle used in testing rather than the range of real-world conditions a

vehicle part will experience in its lifetime (Tutuianu et al., 2015). Methods such as that described in Nyberg et al. (2014) introduce multiple equivalent driving cycles to be tested together. These methods add a layer of realism to driving cycles by ensuring parts are not optimised for one standard trip. It has long been understood and acknowledged in the literature that standardised driving cycles, especially modular ones like the NEDC, routinely underestimate emissions (Joumard et al., 2000; Wang et al., 2015). Developing better standardised DCs for both ICEVs and EVs can highlight the distinct advantages of EVs.

Lin and Niemeier (2002) criticise the antiquated construction methods of a supplemental driving cycle that is still used during emissions tests in California. By quasi-randomly piecing together MTs based on matching the speed-acceleration frequency distribution of the real-world, the resultant driving cycle does not necessarily match the “frequency, duration and intensity” of recorded vehicle actions. MT DCs have “been criticized that the parameters and criteria being used for cycle constructions are not directly related to emissions” (Tong and Hung, 2010). Zhang et al. (2011) heavily criticised the random MT selection method because it only looks at a fraction of the data, it is not computationally efficient, and it does not generally scientifically evaluate the representativeness of real-world driving behaviour. Older criticisms such as these have already been addressed by improved methodology such as the construction method proposed by Ashtari et al. (2014) and general improvements as the field grows.

Mayakuntla and Verma (2015) contend that, with the accepted definition of a MT (vehicle speed between idle times), simple CPs that most studies use fail to characterise driving behaviour cohesively and instead suggest combining trip segments. Even when cycles are validated and evaluated, there is no scientific measure of effectiveness determining how well a DC mirrors real-world driving (Xiao et al., 2012). Xiao et al. (2012) criticised the computational ineffectiveness of DC creation, especially the random MT method.

The discourse surrounding DCs criticising legislative cycles calls into question the purpose of DCs. The purpose of legislative cycles is to create a standard by which to test all vehicles against. The results of the ARTEMIS project showed that using a common test instead of a vehicle-specific driving cycle “could lead to very different emissions estimates” (Boulter and McCrae, 2007) especially because drivers will adapt driving behaviour to their vehicle (André et al., 2006). Even among legislative cycles, the driving statistics are quite different. When the NEDC was updated and largely replaced by the WLTP, the standard behaviour shifted dramatically and increased the energy required by an ICEV by 45%; the WLTP cycle has higher average speed, maximum speed, distance, average driving power, and maximum acceleration and deceleration (Karmangil and Tekin, 2022). While the legislative benchmarks can be shifted due to inherent differences in cycles like these, this result shows how much each cycle differs. Standard limits are applied through the DC, and vehicles or parts can be directly compared by assessing their performance through the DC (Berzi et al., 2016). Inherently a standard will not be perfectly applicable to all situations. Therefore, it might be unnecessary to create a novel DC for each situation and instead adjust thresholds based on the standard. Berzi et al. (2016) concluded that a single DC cannot possibly represent all aspects of driving behaviour for a region and a compromise is necessary to accept when considering DCs. To help with this problem, Berzi et al. (2016) developed a driving cycle generator based on input parameters and data with the goal of creating purpose-built cycles on command from the same database.

Pettersson al. (2018) proposed an “operating cycle” to replace the relatively trivial DC. An operating cycle takes into account the external factors that have large effects on vehicle performance, whereas a DC solely consists of a speed vs time profile. An operating cycle considers parameters of road, weather, traffic, and mission. Romano et al. (2023) introduced stochastic models of generating operating cycles to address three issues of creating DCs: 1. A single DC must be representative of an entire database 2. Variability due

to randomness in external parameters may not be captured by a speed time profile 3. The many parameters that affect operating behaviour are lost when classified only by a few CPs.

5.2.10 Conclusion and Gaps in the Literature of Drive Cycles

This section outlines any gaps in the literature that have been found in the previous sections. The most apparent gap is the lack of standardisation in the DC construction procedure. At each step of construction, there exists a range of choices left to the developer's judgement. There is no definitive state-of-the-art practice to develop the best DC. Comparison studies have been conducted, but they provide contradictory results. Generally, studies that introduce a novel method claim to be an improvement on the common MT and MC methods, but the findings are only supported within the context of that study's database and methodology. A scientific approach to DC creation is necessary to ensure that the demand for representative local DCs is met with accurate DCs.

Another issue is the lack of information that results from creating a DC that is simply a speed-time profile. To expand on this, there are two major, related issues that stem from simplifying data into a DC based on CPs. First, driving behaviour and patterns can be lost either by generating artificial speed data as in the MC method or only selecting a fraction of the total data for the DC as in the MT method. It is difficult to define driving behaviour in ways other than CPs, but the data can be analysed beforehand through methods such as clustering to ensure all types of driving are represented. Second, the effects of various external CPs can be lost by generalising the data into a DC. For example, following the same DC in differing weather and road grade conditions may lead to vastly different emissions or energy consumption. The concept of an operating cycle has been suggested as a solution, but this option lacks the simplicity to be used as a standardised test. It remains to be seen how well these external factors can be accounted for through a speed-time profile.

Through the literature review, some gaps were identified for future studies to fill. A future study can scientifically analyse DC creation across different locations with a mixture

of EV and location-based parameters along with speed data to identify the best DC construction practices to simulate realistic EV loads. It is apparent that a large-scale effort such as the ARTEMIS project is needed to clearly define the state of affairs of EV DCs and what is needed for the future of legislative DCs built to test EVs instead of ICEVs as the default vehicle.

5.2.11 Summary of Previous Drive Cycles

As part of this literature review, each study that developed its own DC was compiled into an informative table on the methodology used. A cycle was recorded in this table when the cycle was new (i.e. did not build upon previous research), unique (i.e. only one cycle was built in the research), and its methods were well documented and clearly stated in the paper. This table that summarises previous DC development is included at the end of this thesis as Appendix A. Appendix A may be informative to future studies as it provides an outline of state-of-the-art DC development.

This summary presented in Appendix A, along with the literature review as a whole, establishes best practices which can be used to construct the DCs in this chapter. First, it was found that with large scale data collection efforts, such as what was described in Chapter 2, prior route selection is not an option. Instead, it is important to understand the context of driving data in terms of road type or area type classified through GPS or speed data. Second, the instrumented vehicle method of data collection used for this thesis is the most appropriate data collection method for modern studies. Two methods to construct DC are prevalent in the state-of-the-art literature, the MC and MT methods. The MT method preserves real driving data directly but can be limited in its application. The MC method creates DCs to fit any scenario using the patterns of recorded driving data. The created DC must closely match the average statistics of driving data. This is usually enforced through an acceptable threshold of error comparing the DC's statistics to the statistics of the database.

5.3 Methodology

This section covers the methodology of all steps performed in this chapter, including the setup of the network model, the implementation of speed limit scenarios, the route search function used to identify trips, and the construction of DCs.

5.3.1 Visum Model Overview

The network used for this model is the National Transport Model (NTpM), developed by AECOM for the National Transport Authority of Ireland (AECOM, 2019A). The model was developed in the software PTV VISUM (PTV Group, 2025), referred to simply as Visum here for clarity. Visum is a macrosimulator commonly used for demand modelling and transport planning scenarios (PTV Group, 2025). A macrosimulator is generally more applicable to larger models, such as state- or city-wide network models, than microsimulators, where each vehicle is modelled individually (Krivda et al, 2021). In a macrosimulator, vehicle flow is treated as a continuous stream and flow is typically assigned to a network using static traffic assignment (Flötteröd and Rohde, 2011). The specific traffic assignment algorithm used in this thesis is discussed in the following sections, along with other specifics of the NTpM. The roadway network of Ireland within the NTpM can be seen in Figure 5.1 below. In Figure 5.1, motorways are blue, national primary roads are green, national secondary roads are red, and regional roads are yellow.

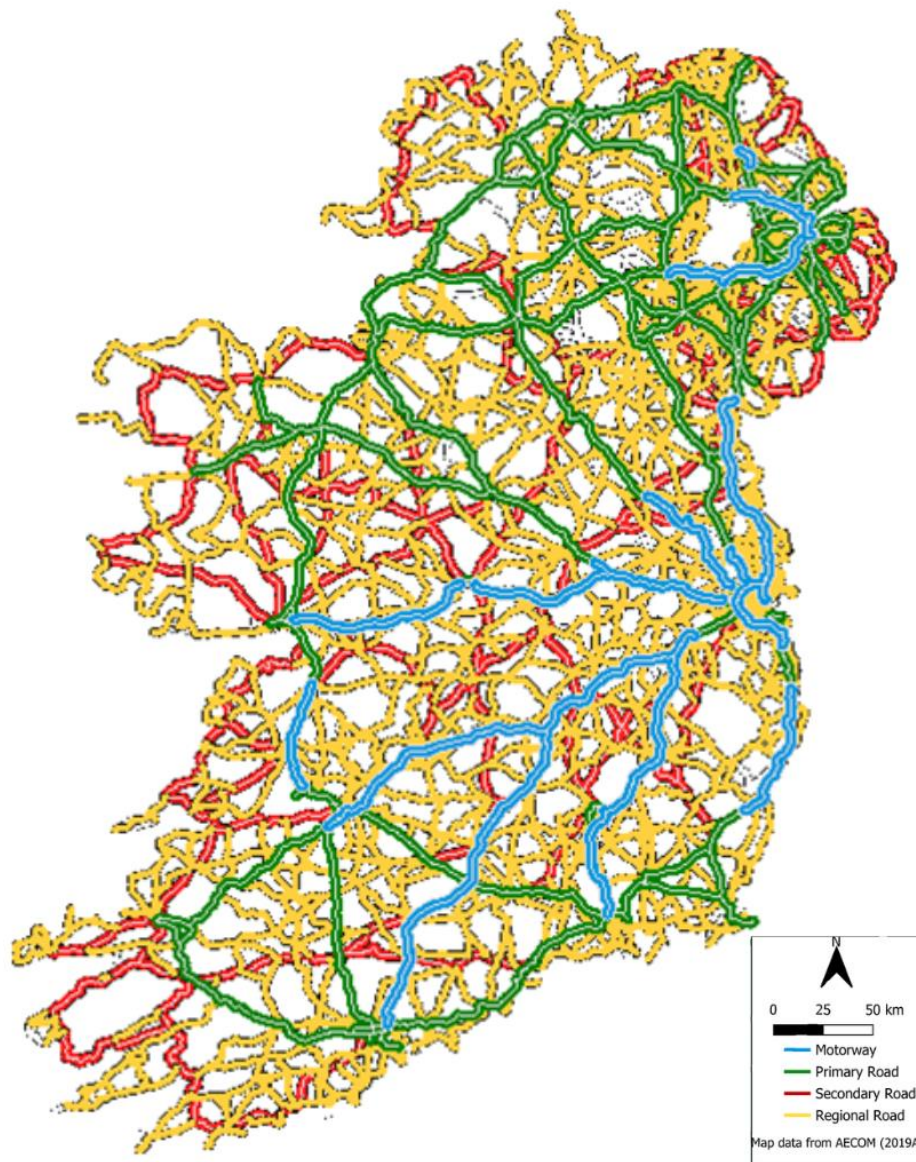


Figure 5.1. NTpM Roadway Network (AECOM, 2019A)

The NTpM is used for strategic transport planning for Ireland (AECOM, 2019A). The NTpM is a high-level model, meaning not every road in Ireland is included as a link in the model, particularly in urban centres, but it is well-suited for interurban demand modelling (AECOM, 2019A). The NTpM covers 70,482 km of roads in Ireland, which accounts for about 70% of all road cover in Ireland (Central Statistics Office, 2021). The modelled volumes were validated using Transport Infrastructure Ireland's Traffic Monitoring Units, and the modelled journey times were validated using Google Maps (AECOM, 2019B). Because the model was built to be accurate to 2016 data, the traffic volumes from 2016 are used in this thesis rather than 2030 predicted volumes, which may

be inaccurate. For the same reason, the methodology used in the development of the NTpM, such as the traffic assignment algorithm, are unchanged in this thesis to ensure results are as accurate as possible. The methodology used to predict traffic flows in the NTpM are described in the following paragraphs. The NTpM can model flows for the AM peak period, 7:00-9:00, and the interpeak period, 12:00-14:00. To avoid an overabundance of results, only the AM peak period is considered in this thesis. Future studies could investigate any differences in the AM peak and interpeak periods. Additionally, only light vehicle trips within the Republic of Ireland are considered in this thesis.

The model generally follows the common four-step demand modelling process, where trips are generated and assigned to a network (McNally, 2007). For the first step, trip generation, the NTpM uses a car ownership, demographic, and freight model as inputs (AECOM, 2019A). Trips are then distributed between 1,129 zones in the model using a trip distribution model (AECOM, 2019A). The trip generation and distribution models in the NTpM were informed by the Place of Work, School or College Census of Anonymised Records (POWSCAR) conducted by the Central Statistics Office during the 2016 Irish Census (AECOM, 2019B). The origins and destinations (real addresses) from the POWSCAR survey were converted into the zonal system used in the NTpM. An additional adjustment was performed to convert the 3-hour AM peak period of the POWSCAR into the 2-hour AM peak period of the NTpM by applying a reduction factor of 0.826 to the 3-hour volumes (AECOM, 2019A). Non-commuting trips were added to the NTpM's origin-destination matrices from the NTA Regional Modelling System, which was informed by the 2012 National Household Travel Survey.

The next step, modal split (McNally, 2007), is not used in the NTpM directly. Instead, modal split analysis is considered using an additional python script called the variable demand model to compare modal usefulness across test scenarios (AECOM, 2021). That means modal split is not a factor for the traffic assignment performed in Visum and will not

affect the journey times or volume in the scenarios modelled in this chapter. Additionally, bicycle and walking modes, which are particularly important in urban demand shift modelling, are not considered in the NTpM. For these reasons, modal split was not considered as part of this thesis. Instead, demand was assumed to be not affected by any reductions in speed limit. It is also assumed that the 4.1% decrease in vehicle kilometres travelled by private vehicles from 2016 to 2023 does not affect the results of this chapter (Central Statistics Office, 2024).

The final step of the four-step model, traffic assignment (McNally, 2007), is traffic assignment. Traffic assignment is the process of assigning traffic to links of the roadway network based on the origin-destination matrices developed in the previous steps, and it is a well-studied branch of transport engineering, with entire books devoted to the subject (Boyles et al., 2020). There are two major types of traffic assignment, static and dynamic. Dynamic traffic assignment can more accurately model realistic traffic flows, but it requires more input data and computational time (Boyles et al., 2020). The NTpM uses static traffic assignment, so the same process is used in this thesis. The specific static traffic assignment methodology is described below.

Static traffic assignment is the process of assigning traffic volume to links on the roadway network, updating travel times based on these assigned flows, and rerouting a portion of traffic onto faster routes. This process is repeated iteratively until an equilibrium is reached within a preset threshold (Boyles et al., 2020). The NTpM uses the Lohse Equilibrium algorithm (Schnabel and Lohse, 2007) for traffic assignment. The first step of a Lohse traffic assignment is an all-or-nothing assignment, where each road user is assigned the route with the least cost based on free-flow speeds (Boyles et al., 2020; Schnabel and Lohse, 2007). The free-flow speed of a road link is the average speed of travel of vehicles without any slowdown due to congestion. In the NTpM, the free-flow speeds are set to the legal speed limits (AECOM, 2019A).

Once the all-or-nothing traffic assignment has been performed, each link in the network updates current travel time based on the volume of traffic assigned to it. The updated travel time is determined by the Bureau of Public Roads 3 (BPR) function, shown below in Equation 5.1 (AECOM, 2019A):

$$t_{cur} = t_0 * \left(1 + \left(a * \left(\frac{V}{C * c} \right)^b \right) \right) \quad (5.1)$$

Where t_{cur} is the updated travel time of a congested link in km/h, t_0 is the free-flow speed of that link in km/h, V is the assigned volume of the link in the current iteration in number of vehicles, C is the capacity of the link in number of vehicles, and a , b , and c are user-defined parameters. The capacity of each link and the user-defined parameters were determined by road type during model development through fitting models to real data. An example of this process for one road type, taken from the NtPM's data collection report, is seen in Figure 5.2. There are 100 different link types, and 39 sets of user-defined parameters used in the NTpM (AECOM, 2019A).

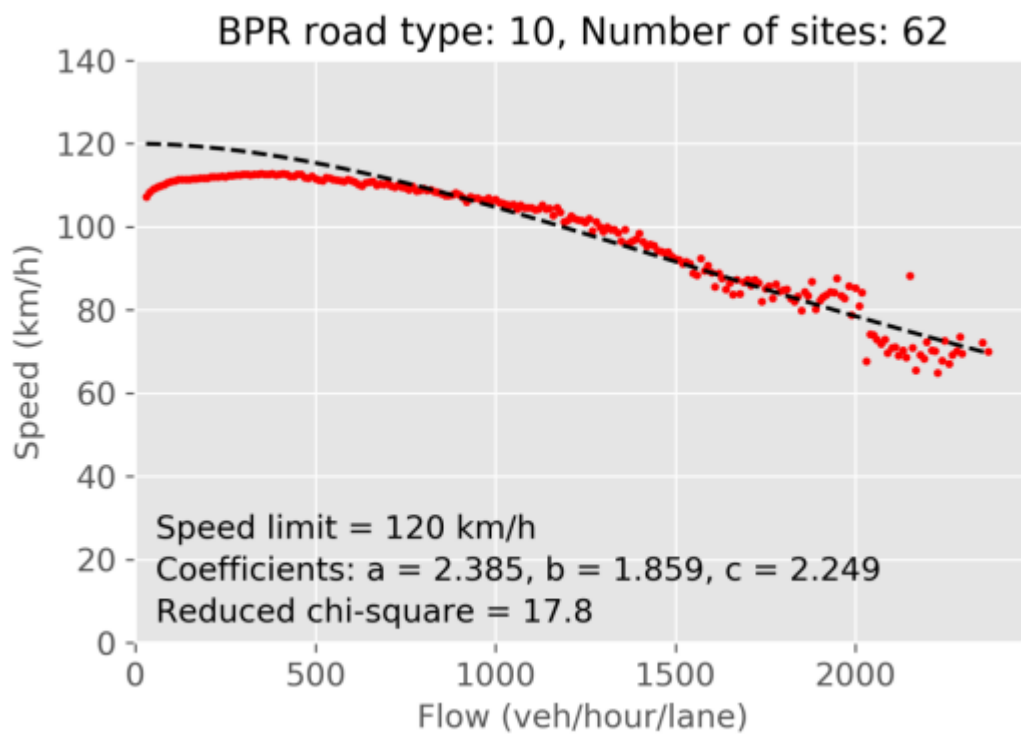


Figure 5.2. Application of BPR Function (AECOM, 2019B)

Once the travel time of each link has been updated using Equation 5.1, the costs of each route are recalculated based on the congested speed of each link along each route. The cost of a link is determined by the distance and current travel time of each link using Equation 5.2:

$$Cost = Value\ of\ Travel\ Time * t_{cur} + Vehicle\ Operating\ Cost * Distance + Tolls \quad (5.2)$$

where t_{cur} is the congested travel time of a link in the current iteration (s), value of travel time is the monetary value of a driver's time (cents/s), vehicle operating costs is the cost of fuel and general vehicle maintenance (c/m), and distance is the distance of a link (m). For the NTpM, the value of travel time is set to 0.7402 cents/s and operating cost is set to 0.0169 cents/m.

The goal of traffic assignment is to ensure that user equilibrium is reached, meaning that no driver can change their route to a route with less cost (Boyles et al., 2020). For each iteration after the all-or-nothing assignment, a proportion of trips will be assigned to routes with less cost until equilibrium is reached (Schnabel and Lohse, 2007). The number of trips reassigned to shorter routes is $1/n$, where n is the current iteration (Schnabel and Lohse, 2007). This means that less and less traffic is reassigned in later iterations as equilibrium is reached. The iterations continue until 300 iterations are completed, or the equilibrium threshold is reached. Equilibrium is considered achieved and traffic assignment is ended if either 98% of network links are assigned less than 1% of a change in volume across four consecutive iterations or the difference in cost between routes in the network is less than 0.1% (AECOM, 2019A). At the completion of traffic assignment, each link in the network contains a realistic travel time, updated by Equation 5.1 based on the final assigned volumes, for the AM peak period that is used for the analysis in this chapter.

5.3.2 Speed Limit Scenarios

Scenarios were developed to analyse the effects of reducing speed limits on EC and journey time. In general, these scenarios were informed by actual speed limit reductions that are currently being implemented in Ireland. Based on the results of modelling and analysis of reduced speed limits in Ireland, including some work using the NTpM, the Speed Limit Review recommended lower speed limits across several road types in Ireland (Department of Transport, 2023A). Speed limits on rural Irish roads were reduced from 80 km/h to 60 km/h in early 2025 (Citizens Information, 2025). Additional speed limit reductions will be implemented later in 2025, as speed limits on national secondary roads will be reduced from 100 km/h to 80 km/h and speed limits in urban cores to 30 km/h (Citizens Information, 2025).

The key performance indicators analysed during modelled scenarios in the Speed Limit Review were safety, emissions, and travel time (Department of Transport, 2023B). Additionally, modal demand shift away from cars was analysed only in urban scenarios (Department of Transport, 2023B). It should be noted that neither fuel consumption of ICEVs nor EC of EVs was investigated in the Speed Limit Review (Department of Transport, 2023B). As mentioned in Section 5.1.2, consideration of EC in speed limit reduction modelling is a novel concept of this thesis.

For the Speed Limit Review of Ireland, modelling of urban speed limit reduction was conducted separately from suburban and rural modelling (Department of Transport, 2023B). For this thesis, a similar approach is continued, where a scenario is developed for urban speed limit changes, and a different scenario is developed for speed limit reductions on high-speed interurban and rural roads. A third test scenario is also developed with all speed limit reductions jointly modelled. The tested reduced speed limits were based on the actual speed limit reductions being implemented in 2025 (Citizens Information, 2025). Because the analysis conducted here is not made to recommend actual policy decisions but instead to

study the interaction between reduced speed limits, EC, and journey time, more speed limit reductions were added to ensure clear outcomes. Additional speed limit reductions were placed on motorways because they make up the bulk of roads travelled during interurban trips.

In the Speed Limit Review, the speed limit was reduced using a blanket approach, meaning the speed limits were reduced regardless of road type (Department of Transport, 2023A). To follow the trend of reduced rural speed limits by 20 km/h, the motorway speed limits were reduced from 120 km/h to 100 km/h. Speed limits in heavily used urban roads were also reduced slightly from 30 km/h to 25 km/h on the basis that low-speed zones would need to remain distinct. The following scenarios were developed specifically for this thesis: Scenario A for reduced speed limits on suburban and rural roads, Scenario B for reduced speed limits on urban roads, and Scenario C with the speed limit reductions from both Scenarios A and B. These scenarios are compared to a base case scenario with no speed limit reductions. The base case speed limits and reduced speed limits are compared below in Table 5.1.

Table 5.1. Base Case and Reduced Speed Limits

Road Types Affected	Base Speed Limit (km/h)	Reduced Speed Limit (km/h)
Regional and Secondary	80	60
Motorway and Primary	100	90
Motorway	120	100
Urban Road	30	25
Urban Road	40	30

Reducing speed limits within Visum is a simple and straightforward process. As mentioned in Section 5.3.1, free-flow speed of links within the NTpM are set equal to the speed limit on that road. Therefore, speed limits can be reduced by modifying the free-flow speed of a link. The list of all 269,700 links in the NTpM was gathered, and any link with a base speed limit equal to a speed limit included in Table 5.1 was reduced to the changed speed limit noted in Table 5.1 in the appropriate scenarios. Then, traffic assignment, as

described in Section 5.3.1, was performed again on the network with modified speed limits, and a new congested speed was found for each link in the network. This process was repeated for each scenario.

5.3.3 Visum Route Selection

In order to investigate the effects of reduced speed limits on EC and travel time, a set of routes is needed. As both interurban and urban speed limit reductions are tested separately, it is essential that both urban and interurban routes are selected for analysis. Three routes of each type, urban and interurban, were selected after a thorough examination of the NTpM.

As mentioned in Section 5.3.1, the NTpM network is coded with only major roads, meaning many real-world roads are not included in the NTpM. This is particularly true about urban centres, where only main roads are coded as links in the network. See for instance the modelled roads in Dublin in Figure 5.2, where only links with coded speeds are shown. Even in the already sparsely coded NTpM, not every link is given a real speed. Only links with a coded speed value are used for traffic assignment modelling, and these links in the Dublin area are shown in Figure 5.3. The yellow links in Figure 5.3 are regional roads, while the white links are local roads (AECOM, 2019A).

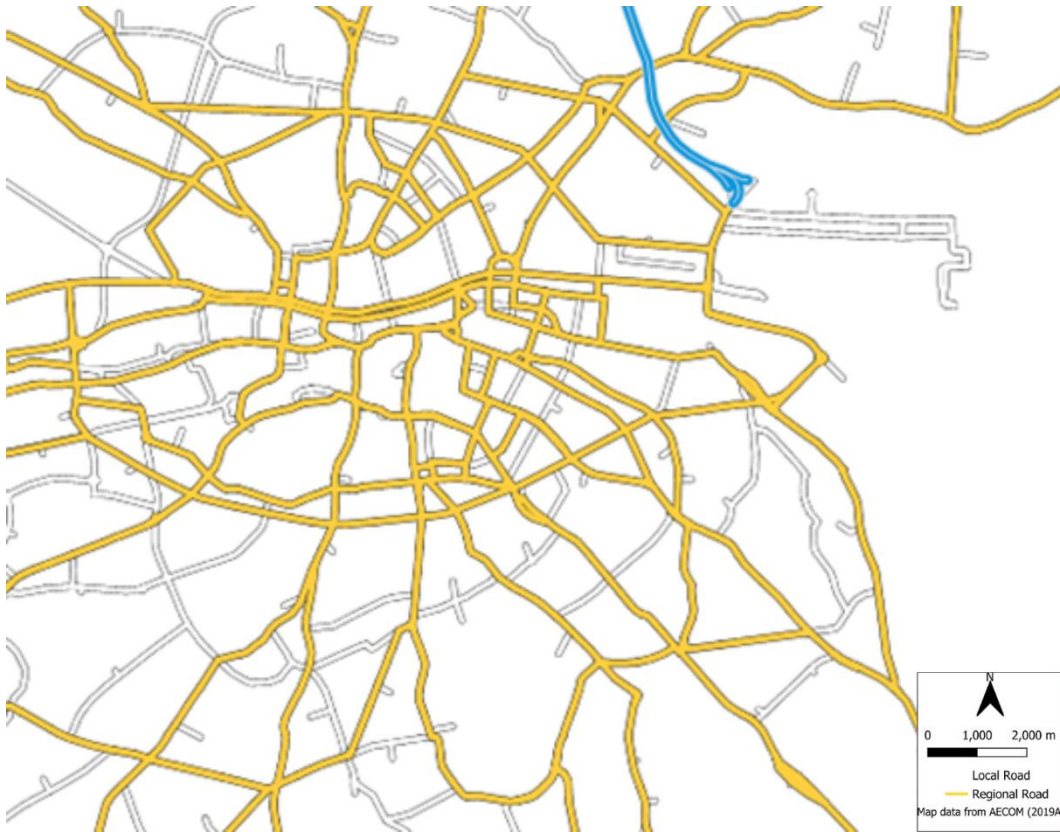


Figure 5.3. Modelled Links of the NTpM in Dublin Area

Figure 5.3 shows why the Speed Limit Review of Ireland did not use the NTpM to investigate the effects of reduced speed limits in urban contexts, as there are limited road links to fully investigate changes like modal and route shifts due to speed limit reductions (Department of Transport, 2023B). Still, the Dublin road network in the NTpM is detailed enough to allow for an analysis on the comparison of EC and travel time under the assumption that demand is generally unchanged in the speed limit reduction scenarios. In the speed limit reduction scenarios, routes may change due to the BPR function of Equation 5.1, but not enough is known to change the overall demand on the network. The network of the Dublin urban area is more detailed than any other urban area in Ireland, so all urban routes are set in Dublin. The origins and destinations of the urban routes were chosen to incorporate a wide range of links in the Dublin network and allow the possibility of different routes. These urban routes are shown in Figure 5.3 along with their endpoint nodes. Unlike

Figure 5.3, Figure 5.4 shows every link in the Dublin area, even those that traffic cannot be assigned to.

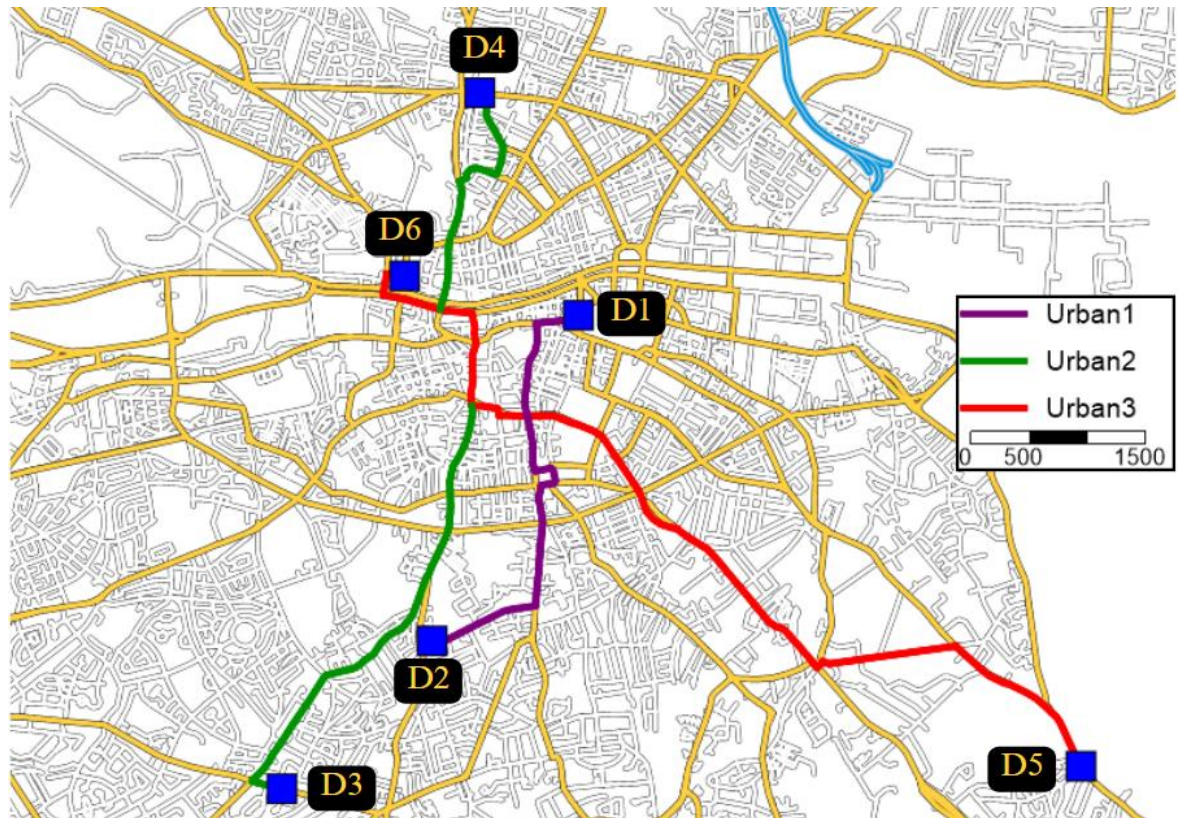


Figure 5.4. Urban Routes

The origins and destination nodes of the urban routes are noted with D1-D6 on Figure 5.4, and their roles within each trip are noted in Table 5.2. The urban routes cover a north-south trip, a south-north trip, and an east-west trip, each through the city centre of Dublin.

Table 5.2. Urban Route Origins and Destinations

Route Name	Origin	Destination	Distance (km)
Urban1	D1	D2	3.76
Urban2	D3	D4	7.04
Urban3	D5	D6	8.03

For the interurban trips, Dublin again serves as a major hub due to its prominence in the road network of Ireland. All interurban trips start in Dublin from point D1, which can be seen on Figure 5.4, directly outside Trinity College Dublin. It was desirable to have interurban trips that would be commonly travelled and push the limits of modern EV range.

Therefore, three destination cities were found across the corners of Ireland: Cork, Galway, and Donegal. These cities are presented on the NTpM road network in Figure 5.5 along with the interurban routes that connect them.

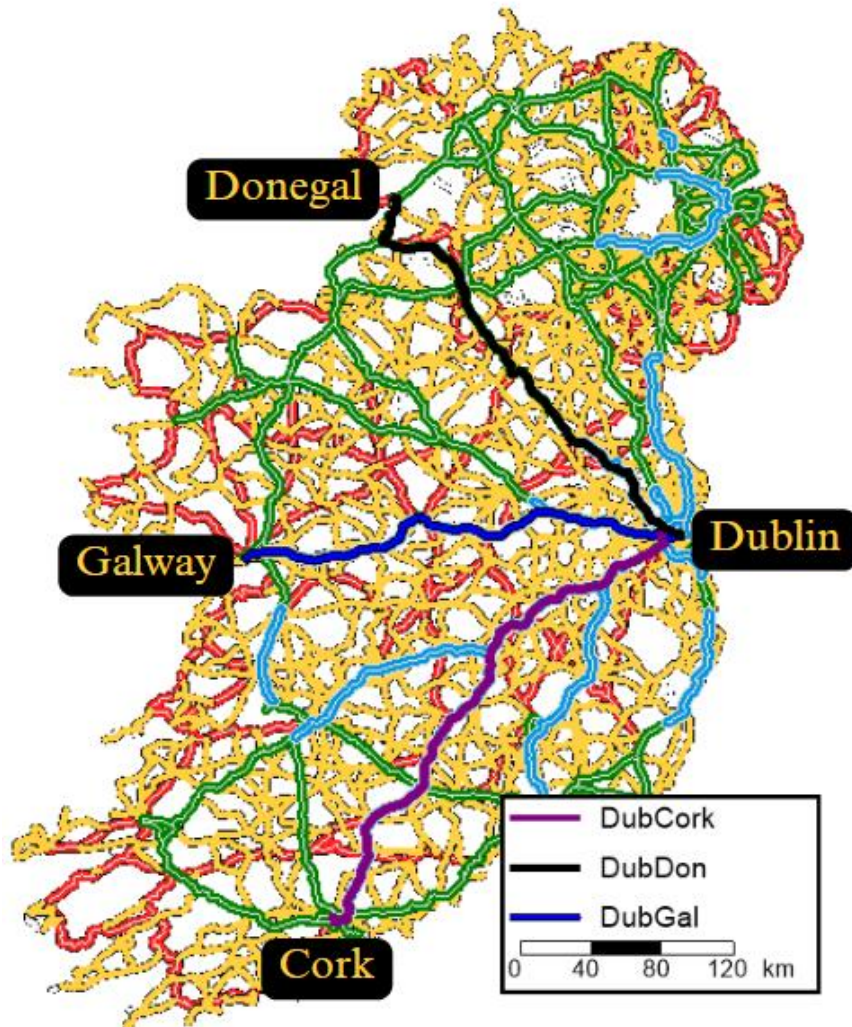


Figure 5.5. Interurban Routes

As mentioned in Section 5.3.1, the colours of roads in Figure 5.5 represent the following road types: blue roads are motorways, green roads are national primary roads, red roads are national secondary roads, and yellow roads are regional or local roads. The routes to Galway and Cork are almost entirely traversed on motorways, which is why motorway speed limits have been reduced in the scenarios of Section 5.3.2. Only a fraction of the route from Dublin to Donegal is connected by a motorway, and the rest of the journey can be

traversed on primary or regional roads. Many alternate routes are possible through all road types. A summary of the interurban trips can be found in Table 5.3.

Table 5.3. Interurban Route Origins and Destinations

Route Name	Origin	Destination	Distance (km)
Dub-Cork	D1	Cork	258.37
Dub-Don	D1	Donegal	229.68
Dub-Gal	D1	Galway	208.40

5.3.4 Visum Route Search Function

An essential part of Visum’s traffic assignment process, which was described in Section 5.3.1, is searching for the best routes between zones. During traffic assignment, routes are searched between network zones in the origin-demand matrix. Those routes are found through a shortest path algorithm that minimises cost as calculated by Equation 5.2. The same shortest path algorithm can be used to connect nodes in the network within Visum. The shortest path algorithm is used to gather route data in this chapter.

The exact algorithm to determine shortest paths in Visum is unclear. Schneck and Nökel (2020) suggests the shortest path algorithm used by Visum is an efficient implementation of the A* algorithm, like what is presented in Dibbelt et al. (2016). The A* algorithm is particularly well-suited to find the shortest path between a single origin-destination pair, compared with all origins and destinations of the network (Boyles et al., 2020). An overview of the A* method is given here to explain the process within Visum, although the process within Visum might vary slightly. To help explain the A* algorithm, a sample network was developed as shown in Figure 5.6.

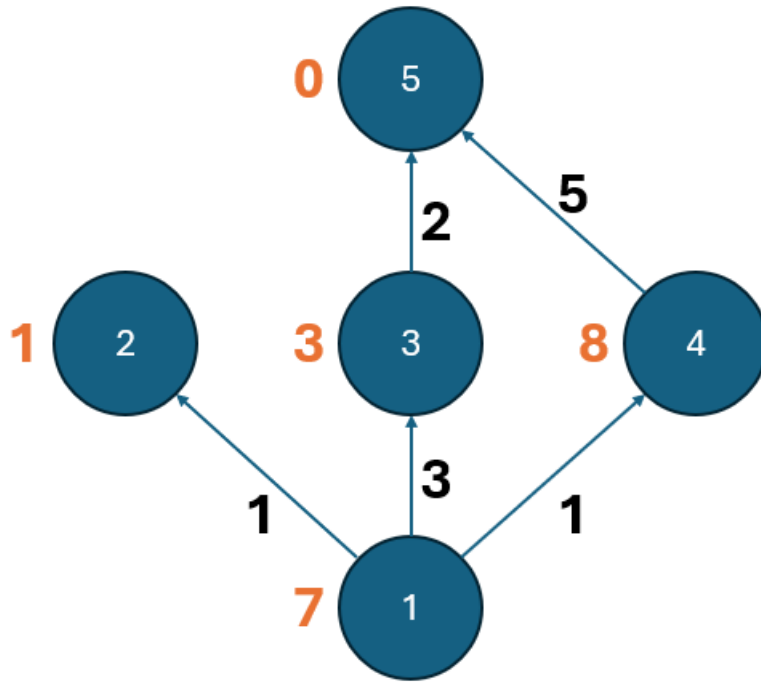


Figure 5.6. Sample Road Network

In the sample network in Figure 5.6, five nodes shown in blue are connected by links, and the cost of each link is denoted by the black number abutting the link. To save on computational time when running the route search algorithm on a large network like the NTpM, the A* algorithm introduces a value given to each node called the heuristic (Boyles et al., 2020). The heuristic estimates a minimum possible cost from each node to the destination node. The heuristic can be set using a variety of methods, including using the geometric distance between each node and the destination node without accounting for available links (Boyles et al., 2020). The heuristics in the sample network are shown in orange to the left of each node.

As an example, consider a trip with an origin at node 1 and a destination at node 5 through the sample network in Figure 5.6. The shortest path will be determined using the A* algorithm. The steps of the algorithm are shown in Figure 5.7 for reference. The A* algorithm progresses iteratively from the origin node. As a first step, all nodes connected to the origin are added to a scan eligible list, a list of all nodes that can be reached but have not

yet been explored (Boyles et al., 2020). The cost from the origin node to each node in the scan eligible list is added to the heuristic of that node to obtain a sum that is used to determine which node is explored in the next step (Boyles et al., 2020). As node 1 is the origin, the path costs in step 1 are simply the link cost between node 1 and the nodes in the scan eligible list. The node with the smallest sum is chosen, which in this example is node 2. As can be seen in Figure 5.6, there are no nodes departing from node 2 so no information is gained in step 2. Since node 2 has been removed from the scan eligible list, the node with the smallest cost sum is node 3. Node 3 is scanned in step 3, and node 5 is added to the scan eligible list as a result of the link connecting nodes 3 and 5. The path cost to node 5 is calculated as the link cost from the origin to node 3 added to the link cost from node 3 to node 5. The A* algorithm terminates when the destination node is added to the scan eligible list, so step 3 is the final step, and the shortest path is chosen as node 1 to node 3 to node 5.

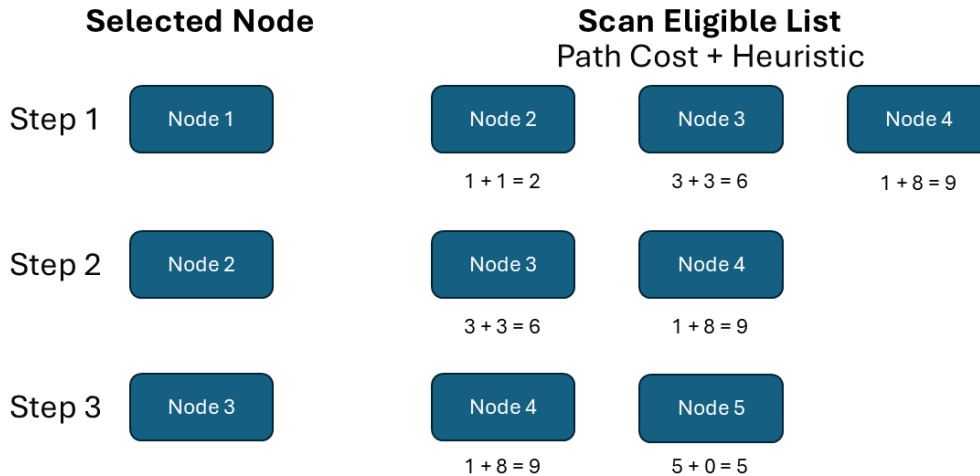


Figure 5.7. A* Algorithm Example Steps

The shortest path search is achieved in Visum using the built-in function “Route Search PrT”, where PrT denotes private transport. The shortest path search was automated to search for all routes in Python through the Visum COM system, which enables the coding of Visum processes outside of the Visum software. Visum gives the option to search for a

shortest path that minimises general cost, travel time, distance, or free-flow travel time. General cost, as described in Equation 5.2, is useful for traffic assignment but does not directly relate to real-world route choices. Drivers generally choose routes that minimise travel time (Wang et al., 2020). The routes in Figures 5.4 and 5.5, as well as the distances presented in Tables 5.2 and 5.3, are from the shortest paths found by minimising travel time. As part of eco-driving behaviours, drivers can also opt for routes that minimise distance (Wang et al., 2020). Chapter 3 of this thesis showed the strong relationship between distance and EC. Additionally, choosing routes that minimise free-flow time are modelled here to represent a driver who knows the fastest route generally but does not know current traffic conditions. Therefore, for each route described in Tables 5.2 and 5.3, a path is found that minimises travel time, distance and free-flow travel time. That means three paths are generated for each route, and the results of each path will be compared to analyse the effect of route choice on EC and travel time. To simplify results, route choice comparisons are only made in the base case, where no speed limit changes are made.

5.3.5 Drive Cycle Construction

Once the paths have been found in Visum for each route in each scenario, realistic drive cycles are built to represent real driving behaviour on the routes. The research that funded this thesis heavily relates to drive cycles, so the literature review on this subject was in-depth and related to all types of DCs. As mentioned in section 5.2 above, most DCs are built to represent the average trip of an area for use in standardised vehicle testing. Though the DCs built in this chapter represent the average trip of a predetermined route, much of the same processes are used to build these DCs as to build standardised DCs, from data collection to cycle construction. All steps of the methodology used to develop realistic DCs are explained in this section.

Starting with data collection, many studies on DCs point to route selection as an important step (Wang et al., 2021). However, Ho et al. (2014), noted that route selection is

not applicable to large scale data collection efforts like what is described in Chapter 2 of this thesis. Therefore, the DCs in this chapter are built on driving data from all around Ireland, the same data used to build the EC model in Chapter 3.

In section 5.2.3, two primary methods to construct DCs were found in the literature: the MT method and the MC method. The MT method pieces together MTs, generally defined as the driving period between consecutive stops, to form a complete cycle. However, as the routes in this thesis are made of links with average speeds and distances already determined by Visum, it is unlikely or impossible to have MTs in the database that match the exact speed profiles needed. The MC method, on the other hand, is particularly useful for the application here. The MC method, as described in Section 5.2.3.2, uses a TPM to probabilistically select the speed value of the next second based only on the speed value of the current second, employing the Markov property. This MC method is explained in more detail below.

In the MC method, a DC is built second-by-second using historical driving data. A TPM is constructed to determine the speed at the next second probabilistically. To build the TPM, first every second of real driving data was compiled into a matrix. The matrix contains an equal number of rows and columns, which are integers from 0 km/h to the maximum recorded speed 128 km/h. As seen in Table 5.4, each row represents a possible value of the current speed. The current speed is either the preselected first speed value of a DC or the last speed value selected by the TPM. Each column represents a possible speed value to be selected for the next second of the DC. The matrix was filled by row, with each column in a row filled with the number of times that speed value immediately followed the row's speed value in the database. Every cell in the matrix is then divided by the sum of its row, transforming the counts into probabilities.

This new matrix, which gives the probability of the next speed value given a current speed value, is the TPM. A small section of the entire TPM is included in Table 5.4 to better visualise and understand the function of the TPM. Table 5.4 contains only speed values from

1-10 km/h, and each cell shows the probability of a next speed being selected given a current speed value. Redder cells in Table 5.4 indicate higher probabilities of selection.

Table 5.4. Transition Probability Matrix Snippet

		Next Speed (km/h)									
		1	2	3	4	5	6	7	8	9	10
Current Speed (km/h)	1	30%	13%	5%	4%	3%	3%	1%	2%	1%	0%
	2	17%	23%	12%	9%	5%	4%	3%	2%	1%	1%
	3	10%	21%	19%	13%	6%	7%	3%	3%	2%	3%
	4	7%	13%	15%	23%	10%	7%	4%	5%	3%	3%
	5	5%	11%	10%	17%	15%	13%	6%	7%	4%	5%
	6	3%	5%	7%	14%	13%	22%	10%	7%	5%	4%
	7	1%	2%	4%	8%	8%	14%	31%	10%	5%	5%
	8	0%	2%	2%	7%	6%	12%	13%	21%	10%	9%
	9	0%	1%	1%	3%	4%	7%	7%	18%	18%	12%
	10	0%	0%	1%	2%	3%	5%	6%	11%	14%	20%

As an example, if the current speed while building the DC is 1 km/h, the probability that the next second in the DC will be 1 km/h is 30%, and the probability the next second will be 2 km/h is 13%. Once the next speed value is chosen, it becomes the current speed value and the TPM determines another speed value probabilistically until the desired length of the DC is reached. A trend can be seen in Table 5.4 across all speed values that the speed value of the next second is most likely to be equal to the current speed. Then, probabilities gradually decrease for more distant speed values. In the full TPM, the probabilities for each row add to 100% as each row contains all possible next speed values for a current speed.

The final step of DC construction is to validate the built DC. As mentioned in Section 5.2.5, hundreds or thousands of candidate DCs are built until a DC matches the desired CP within a certain error threshold. That is to say, the final DC must match the desired statistics, generally the average statistics of CPs, like average speed and acceleration, from a driving database. In this chapter, the DC must be built to match the route characteristics taken from Visum. The outputs of the route search in Visum are a distance and average speed for each link travelled along the route. Within this context, the validation of DCs necessitates a slightly varied methodology than the literature generally utilises.

For this chapter, DCs are built for each individual link along a route, and each link-level DC is pieced together to form a DC for the entire trip. Link-level DCs are necessary to ensure accurate velocities are preserved along the entire route. For each link-level DC, the TPM is used as described above to add speed values second by second. The link-level DC is finalised when the link distance is reached or exceeded. Distance is tracked during the DC construction process by multiplying the speed value of each second in m/s by 1s to find distance in metres. The result is a candidate cycle, and the average speed of this candidate cycle is compared to the average speed on the link. If the average speed of the candidate cycle is within 5% of the average speed of the link given by Visum, the candidate cycle is selected to be included in the route-level DC. A low threshold of 5% was chosen to ensure speed in the DC closely matched speeds modelled by Visum, and it is a commonly used threshold in the literature (Berzi et al., 2016; Huertas et al., 2019; Yang et al., 2023). If no candidate cycle reaches this threshold after 1,000 cycles are constructed, the candidate cycle with the average speed closest to the average speed from Visum is selected. To ensure continuity in the route-level DC, the speed at the end of the link-level DC is carried over to the next link as the input to the TPM.

To capture a variety in driving behaviours and ensure robust results, 100 DCs were built for each route taken from Visum. From Chapter 4, it can be seen that common categories of driving behaviour aggressiveness are unable to consistently impact results of EC. Therefore, the 100 DCs are all created using the TPM, and any variations in driving behaviour are caused only by the probabilistic, stochastic nature of the MC method used to create DCs. After the 100 DCs are created, the mean value of EC is used for analysis on the effects of speed limit reduction, and the empirical method from Chapter 4 is used to investigate what driving behaviours most affected EC within the 100 DCs.

5.3.6 Preparing Drive Cycles for Energy Consumption Modelling

The route-level DCs are segmented into 500m segments and processed to be input to the model just as real-world trips were in Chapter 3. The statistics for each 500m segment, such as average speed, acceleration, and deceleration were calculated using the same automated processes in Python used for Chapters 3 and 4. However, environmental, driver, and vehicle parameters cannot be extracted from the Visum model. Climate parameters, such as temperature, rain, and wind speed were set to the mean value of the real data used to train the EC model. The same was done for the driver and vehicle parameters. These mean values can be found in the descriptive statistics tables from Chapter 3.

The only remaining parameters to allow the DCs to be modelled are average slope, percent time spent in urban areas, and percent time spent on the various road types. Although these parameters cannot be gathered directly from the Visum links, the NTpM network can be used as a base from which this information can be gathered. Although only link information is gathered during the route search function of Visum, each link contains information about which nodes the link starts and ends at. The NTpM has coded each node with a Boolean variable that labels the node either urban or not urban. This variable was used to determine the percentage of the DC in urban areas. If both the start and end node of a link were labelled as urban in the NTpM, each second of that link-level DC was marked as urban. If one or both of the end nodes of a link were marked as not urban, each second of that link-level DC was marked as not urban.

Unlike the links, the nodes exist at a singular point in space, and Visum provides the coordinates of each node in the network. To gather slope data, each of these nodes was given elevation data. This process was done by extracting the shapefile of the nodes to the GIS software QGIS (QGIS.org, 2024), which was previously used in Chapter 3 to gather road type information. An elevation data source, an EPA 5m resolution DEM, was obtained to provide elevation data on each node. This DEM provides one elevation reading for each 5

square metres in Ireland. In QGIS, each node was tagged with the elevation data of the square that the node's coordinates corresponded to. To get information on slope, a simple two-dimensional geometric problem was solved to get the slope angle using the link's start node's elevation, the end node's elevation, and the link distance. This slope angle was assumed constant along the link and assigned to each second of the link-level DC.

Finally, road type data was needed for each link of the NTpM. The methodology for classifying Visum road types is similar to that used in Chapter 3 to classify road types of GPS coordinates. When building the NTpM, OSM road types were used to classify links (AECOM, 2019B). However, tertiary, residential, and unclassified roads were all grouped into one category called regional roads in the NTpM. To differentiate these road types, QGIS was used to tag links from the NTpM to the nearest road in OSM (OpenStreetMaps, 2024), and the OSM road type was given to the NTpM link. The road network of the NTpM and OSM matched closely in QGIS as can be seen in Figure 5.8, allowing for accurate road type classification. In Figure 5.8, green lines are OSM roads and red lines are NTpM roads.



Figure 5.8. Comparison of NTpM and OpenStreetMaps Roads

5.3.7 Drive Cycles with Road Type TPMs

During initial tests of the DC methodology, it was found that a single TPM was unsuitable to make DCs for interurban routes as presented in Section 5.3.5. When all roads are counted together, the average speed is lower than motorway speeds. This results in the TPM pulling the DC toward more moderate speeds, as the speed values with higher counts have a higher probability of being selected. The speed counts from the entire database can be seen in Figure 5.9 below. The histogram counts the number of seconds each speed value was recorded in the database.

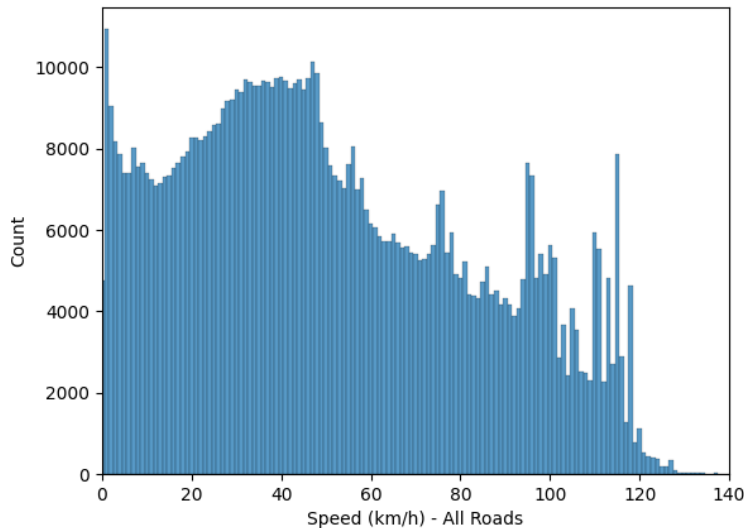


Figure 5.9. Speed Count Histogram on All Roads

Fig 5.9 shows the driving behaviour of the participants as a whole during this study. Most of the driving was conducted at relatively lower speeds, around 30-50 km/h. As speed increases, speed counts tend to decrease. There are notable exceptions, seen in several spikes at distinct speeds. These spikes are likely due to participants cruising at constant speeds for long durations during long distance trips.

The Dub-Cork route with the minimum distance criteria was used to test the methodology. The 100 resulting DCs had an average speed of 76.25 km/h and an average

duration of 3:20. All trip durations in this chapter are rounded to the nearest minute and presented in the form hours:minutes. This is exceptionally slow for an interurban trip, especially when the route gathered from Visum that the DCs are supposed to simulate had an average speed of 94.26 km/h and a duration of 2:40. After an investigation, it was found that the MC method was unable to consistently achieve high speeds as the TPM would tend to select low speeds.

As a solution, the TPM was split into multiple TPMs. One TPM was built for each of the six road types. Separating TPMs is not a novel idea in the literature. Gong et al. (2018) and Wang et al. (2019) developed multiple TPMs at various speed and acceleration states. Mouzouras (2022) documented a similar problem, which they labelled a “dead end” probability in the TPM, where the TPM would get stuck predicting the same speed values. Mouzouras (2022) developed three distinct TPMs, one for each road section: urban, rural, and motorway. A similar idea is employed in this chapter, the only difference being the TPMs are categorised by road type. Zhang et al. (2021) among others showed how road type can influence speed distributions. The speed counts for each road type can be seen in Figure 5.10 below. It should be noted that the y-axes of the charts in Figure 5.10 all differ based on the volume of counts for each road type. The number of counts by road type are found in Table 5.5.

Table 5.5. Speed Counts by Road Type

Road Type	Count
Motorway	83463
Primary	35422
Secondary	361590
Tertiary	154891
Residential	152028
Unclassified	119703

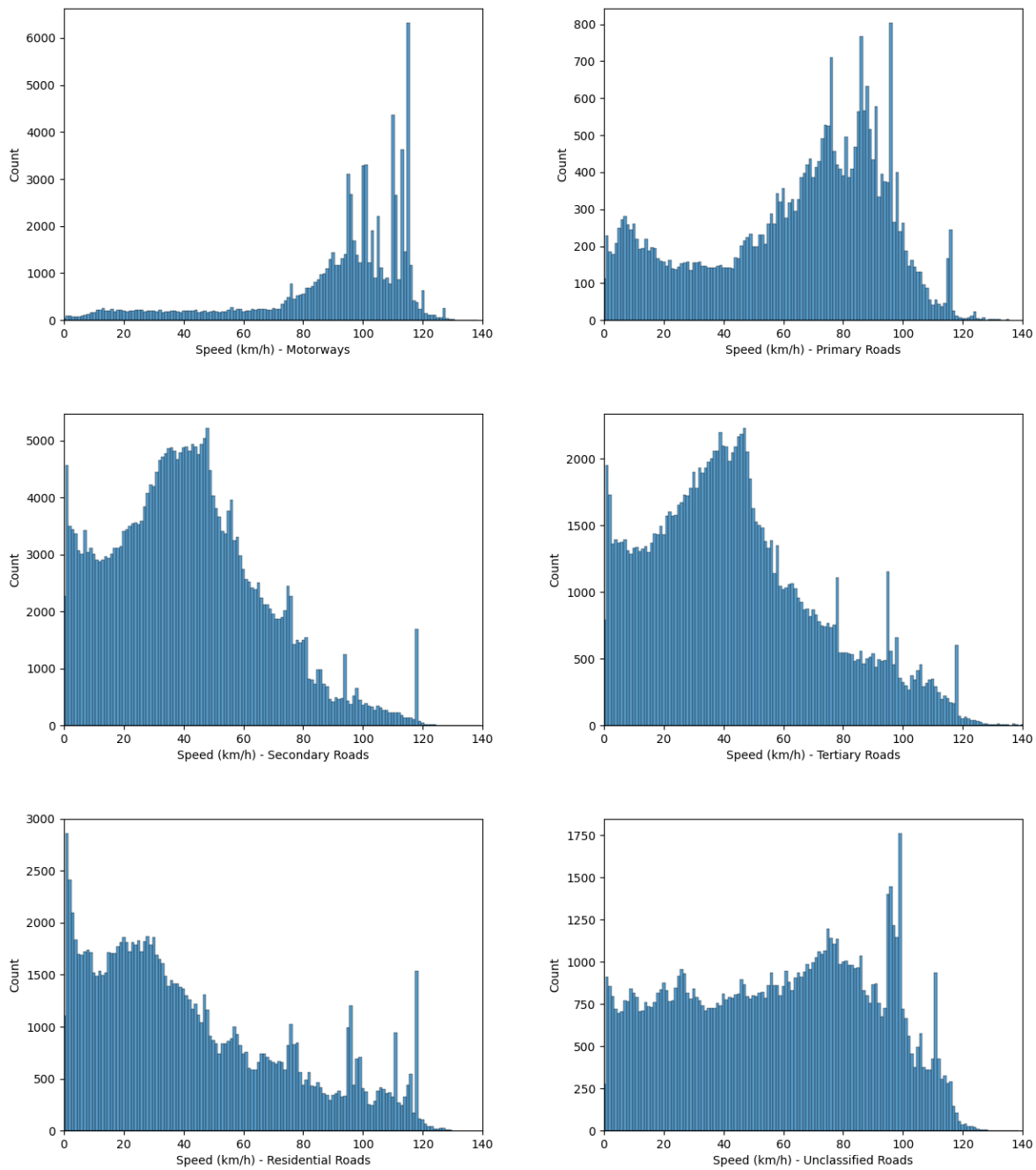


Figure 5.10. Speed Count Histogram by Road Type

Figure 5.10 shows that speed counts vary significantly by road type, making road type a viable classifier on which to split the TPMs. The speed counts generally reflect road type well, with motorways having the highest speeds on average, followed by primary roads and dropping to residential roads with the lowest speeds. The speed on unclassified roads shows a relatively uniform distribution, meaning a DC for an unclassified link can achieve any speed required by a Visum route.

Forming six TPMs, defined by road type, improved the DC creation process in all metrics. The 100 DCs built for the Dub-Cork test route with six TPMs much more closely matched the Visum route's average speed of 94.26 km/h and duration of 2:40; the average speed of the DCs was 92.57 km/h, and the average trip duration was 2:42. An additional improvement came in the form of decreased computational time. The previous methodology struggled to form link-level DCs that matched the desired average speeds, so in many cases 1,000 candidate cycles were built. With the differentiated TPMs, the MC method could keep high speeds on motorways without dipping to low speeds. This resulted in the computational time to build 100 DCs dropping by more than half from 3,152 seconds to 1,036 seconds.

5.3.8 Drive Cycles with Combined Links

To speed up the DC construction process, it was theorised that links with similar speeds could be combined. Since up to 1,000 candidate cycles may be created for each link, computational time could be significantly shortened by limiting the number of links along a route. A method was developed in the Python code to combine any and all consecutive links with an average speed within 5 km/h of the first link. The average speed of the combined link is a combination of the average speed of all independent links weighted by link distance. For example, if a 1 km link with a speed of 20 km/h is followed directly by a 4 km link with a speed of 25 km/h, they would be combined into a 5 km link with a speed of 24 km/h.

Once again, the Dub-Cork route found with the shortest distance was used to test this methodology. Combining similar links reduced the number of links on this route, and thus the number of DCs required, from 325 to 25 links. The majority of the combined links were found on the motorway section of the route, which was combined into one 199 km-long link. The result on computational time was found to be opposite of what was expected: computational time to build 100 DCs rose from 1,036 seconds to 1,525 seconds. This is because with smaller links, there is less time to deviate significantly from the starting speed. On the other hand, this means the methodology with combined links also can leave more

room for differences due to driving behaviours. This can lead to higher variation in EC with combined links. The statistics for EC of the Dub-Cork route with both standard and combined links are displayed below in Table 5.6.

Table 5.6. Energy Consumption with Standard and Combined Links

	Mean EC (kWh)	Min EC (kWh)	Max EC (kWh)	Std EC (kWh)
Standard Links	46.70	46.01	47.16	0.23
Combined Links	46.31	45.22	47.58	0.49

The method with combined links did allow for higher variation in EC, with a standard deviation more than double that of the DCs built with standard links. Although combining links does not change the average speed or distance of the route, it affects the DC creation. Longer links in the combined methodology means speed can have higher dips and peaks along the DC, which causes greater shifts in EC. Because of the computational time advantage as well as keeping more aligned with speeds from the Visum network, links were not combined when building DCs to gather the following results.

5.4 Results

With the methodology established, the results will be presented in the following subsections. The first subsection gives a brief analysis of how different driving behaviours captured in 100 different DCs can affect EC. Next, the effect of route choice on EC and journey time is assessed. The final subsection details how the speed limit reduction scenarios affect EC and journey time.

5.4.1 Driving Behaviour and EC

To build on the relationship between driving behaviour and EC, the empirical method discussed in Section 4.4 is employed here on the realistic DCs built for the Visum routes. For each route, 100 DCs were built to ensure robust measurements of EC as well as to analyse how any differences in driving behaviour within these DCs affects EC. The empirical method sorts these 100 DCs into thirds by EC so that the average statistics of each

EC group can be compared. To keep results brief, only the Dub-Cork and Dub1 routes with minimum distance are examined through the empirical method so that one urban and one interurban route are studied. The empirical method applied to the 100 DCs for the four remaining routes can be found in Appendix B.

The first route to analyse is the Dub1 urban route. The methodology for the empirical method was established in Section 4.4, and the results of its implementation on the Dub1 route is seen below in Table 5.7.

Table 5.7. Average Statistics by EC Group: Dub1 Route

Parameter	Energy Consumption		
	Low	Moderate	High
Duration (Sec)	557	556	558
Average Speed (km/h)	25.57	25.59	25.53
Std Dev Speed (km/h)	4.21	4.21	4.41
Average Acceleration (m/s ²)	0.71	0.72	0.74
Maximum Acceleration (m/s ²)	2.95	2.99	3.37
Std Dev Acceleration (m/s ²)	0.45	0.48	0.51
Average Deceleration (m/s ²)	-0.73	-0.77	-0.78
Maximum Deceleration (m/s ²)	-3.06	-3.02	-2.74
Std Dev Deceleration (m/s ²)	0.51	0.51	0.51
Cruise Time (%)	22.17	22.55	23.27
Acceleration Time (%)	39.84	40.27	39.96
Deceleration Time (%)	37.98	37.18	36.78
Energy Consumption (kWh)	0.53	0.57	0.61

The results of the empirical method show statistical trends in most acceleration and deceleration parameters between the 100 DCs when grouped by EC. Although the differences across the groups may seem small, these differences equate on average to an 8% increase or decrease in EC compared to the moderate consumption group. The largest statistical differences between the groups are in maximum acceleration and deceleration. This is because all other parameters are already determined by the methodology, and the short links do not allow for much variation in speed. The high EC group has a 13% higher maximum acceleration and a 9% lower maximum deceleration compared to the moderate group. Higher standard deviation of acceleration, acceleration rates, and deceleration rates

also tend to increase EC. These results show that lower and steadier acceleration rates tend to be effective eco-driving behaviours in real-world urban driving scenarios. Deceleration rates have a more complicated relationship due to the intricacies of regenerative braking discussed in Chapter 4. The results of Table 5.7 show that slower, more moderate decelerations lead to lower EC on average. However, stronger maximum decelerations also can lower EC in some low-speed driving scenarios.

The Dub-Cork route is used to compare EC and driving behaviours for interurban trips. The empirical method was applied to the 100 DCs for the Dub-Cork route, and the results are shown in Table 5.8.

Table 5.8. Average Statistics by EC Group: Dub-Cork Route

Parameter	Energy Consumption		
	Low	Moderate	High
Duration (Sec)	9892	9861	9847
Average Speed (km/h)	92.74	93.02	93.17
Std Dev Speed (km/h)	1.62	1.61	1.61
Average Acceleration (m/s ²)	0.32	0.32	0.32
Maximum Acceleration (m/s ²)	3.65	3.70	3.78
Std Dev Acceleration (m/s ²)	0.11	0.11	0.12
Average Deceleration (m/s ²)	-0.33	-0.33	-0.34
Maximum Deceleration (m/s ²)	-3.85	-3.83	-3.68
Std Dev Deceleration (m/s ²)	0.12	0.12	0.12
Cruise Time (%)	70.00	70.29	69.82
Acceleration Time (%)	15.48	15.34	15.55
Deceleration Time (%)	14.52	14.37	14.63
Energy Consumption (kWh)	46.44	46.72	46.93

For the interurban trip, both driving behaviours and EC are more uniform across all groups. On average, the low EC group has 45 second longer DCs and 0.49 kWh lower EC. No parameter shows more than a 4% difference between the moderate group and the high or low consumers. Maximum acceleration and deceleration show some trends across the groups, but their values are less meaningful on interurban trips since it only captures one second of driving. The lack of trends across EC groups is due to the fact that driving behaviours were not predetermined, so acceleration and deceleration rates can average out

over long trips to similar values. This can be seen as the maximum average acceleration value of the 100 DCs is only 4% higher than the minimum value. The highest average speed is only 1% higher than the lowest average speed. Stronger trends between driving behaviour and EC on interurban trips may be seen if driving behaviour is categorised beforehand and uniform throughout the trip. But that would mean moving away from realistic driving behaviours that are essential to obtain realistic results of EC.

5.4.2 Route Choice and EC

As mentioned in Section 5.3.4, Visum can search for a path between two points that minimises distance, travel time, or free-flow travel time. The resulting routes are labelled here as least distance, fastest time, and free-flow time, respectively. In this section, a comparison is made between journey time and EC for these different route choice criteria.

The shortest paths found with the different criteria generally followed similar links. As previously discussed, the NTpM does not include every road in Ireland, so generally few routes with close to minimum costs exist between two points. An example is the Dub-Don route, which was particularly theorised to have strong alternative routes due to the lack of a motorway connection. On this 229 km route, 91.5% of links are shared between the shortest time and shortest distance paths. This results in a route change of links that total only 2.94 km of distance.

Few alternate routes on interurban sections were found through the different criteria. Most of the links that differ between route choices are within the Dublin area. An example of how the different criteria affect paths in Visum can be seen in Figures 5.11 and 5.12 below. These figures show the portion of the Dub-Cork route within Dublin for with the least distance and fastest time criteria, respectively. It can be seen that the fastest time route varies slightly to limit travel time at the expense of increasing travel distance. In Figures 5.11 and 5.12 the path links are shown in purple.



Figure 5.11. First Links of Least Distance Dub-Cork Route



Figure 5.12. First Links of Fastest Time Dub-Cork Route

The first comparison between route choices tabulates how EC is affected. For the three route choice criteria, 100 DCs were built for each route using the base case speed limits. The average EC of each route choice for each route can be seen in Figures 5.13 and 5.14 below.

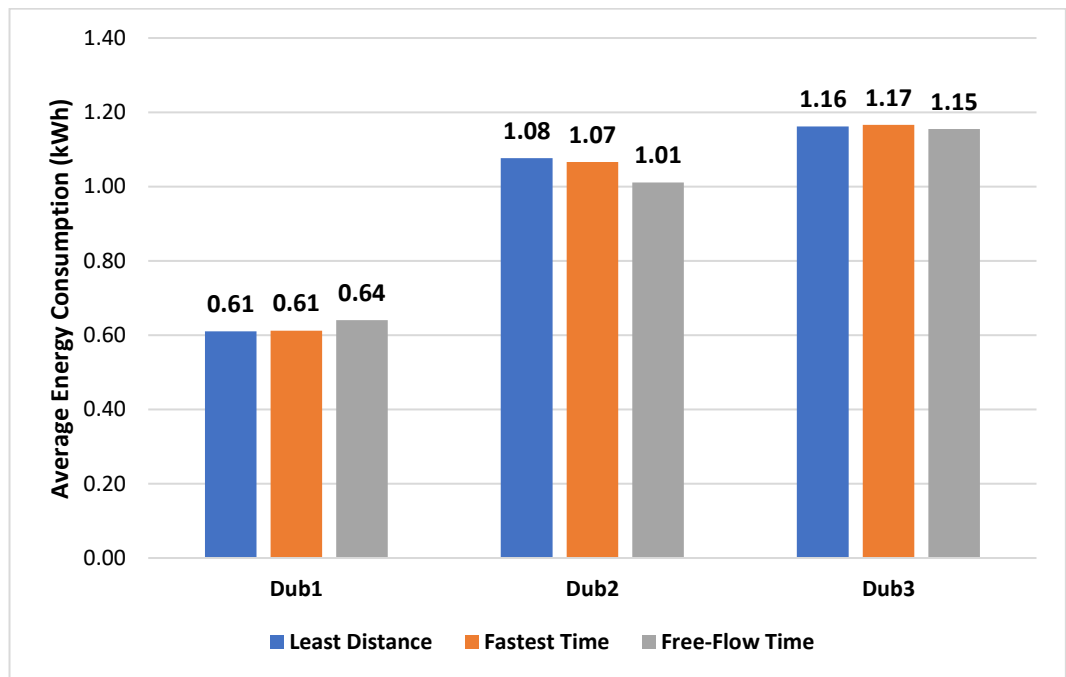


Figure 5.13. EC by Route Choice on Urban Routes

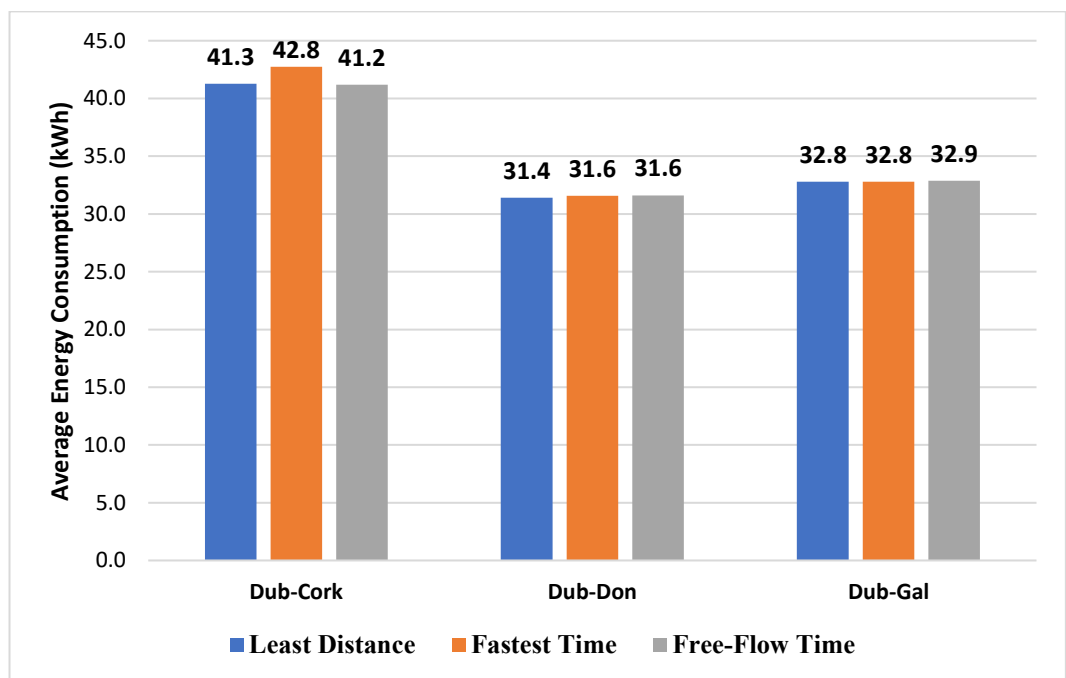


Figure 5.14. EC by Route Choice on Interurban Routes

Figures 5.13 and 5.14 show that route choice can have an effect on EC, although the difference depends on the route. The amount of EC saved is determined by the alternate routes available for the shortest path function. There is no clear correlation between route choice and EC on the urban routes. Minimising free-flow time results in the most EC on the Dub1 route, but the least EC on the Dub2 and Dub3 routes. For the interurban routes,

minimising distance consistently reduces the EC of a route, but the amount of that reduction varies significantly.

Because 100 DCs were built for each route choice and scenario, ANOVA testing can be used to compare differences across groups as performed in Section 4.7.2 (Kim et al., 2017; Virtanen et al., 2020). ANOVA is a statistical test that helps determine whether the mean of points within one group is statistically different to the mean of points of a different group. In this application, a one-way ANOVA test is implemented to determine whether the mean of 100 DCs developed for a route choice criterion is statistically different from the mean of the 100 DCs created for either of the other two route choice criteria (Virtanen et al., 2020). It was decided to not include pairwise t-tests in this analysis, which could be used to determine which specific route choices are statistically different to other route choices, as this information can be directly inferred from the figures. The results of the ANOVA test for determining whether EC is statistically different for different route choices are displayed in Table 5.9. In these ANOVA tests, the null hypothesis is that there is no difference in EC between differing route choice criteria.

Table 5.9 ANOVA Test Results for EC by Route Choice

Route	F-Statistic	P-Value
Dub1	0.46	6.32E-01
Dub2	10.97	6.96E-05
Dub3	2.22	1.16E-01
Dub-Cork	24.52	7.56E-09
Dub-Don	0.18	8.36E-01
Dub-Gal	0.50	6.10E-01

Table 5.9 shows that the statistical significance of the difference in EC varies greatly depending on the route. The f-statistic gives information on how different the variance is between route choice criteria compared to the variance between DCs within the same route choice criterion. The p-value, then, determines the confidence with which the null hypothesis can be rejected. If the null hypothesis is rejected, it means that route choice

creates a significant difference in EC. Typically, statisticians declare statistical significance at the 95% confidence interval, meaning the p-value is below 5.00E-02. In the case of the routes used for this study, Table 5.9 shows that the mean of EC by route choice was only statistically significant on the Dub2 and Dub-Cork routes. It can be assumed from Figure 5.13 that Free-Flow time led to significantly lower EC on the Dub2 route and from Figure 5.14 that the Fastest Time criteria led to significantly higher EC. Another piece of information to evaluate route choice is the journey time, which are compared below in Figures 5.15 and 5.16.

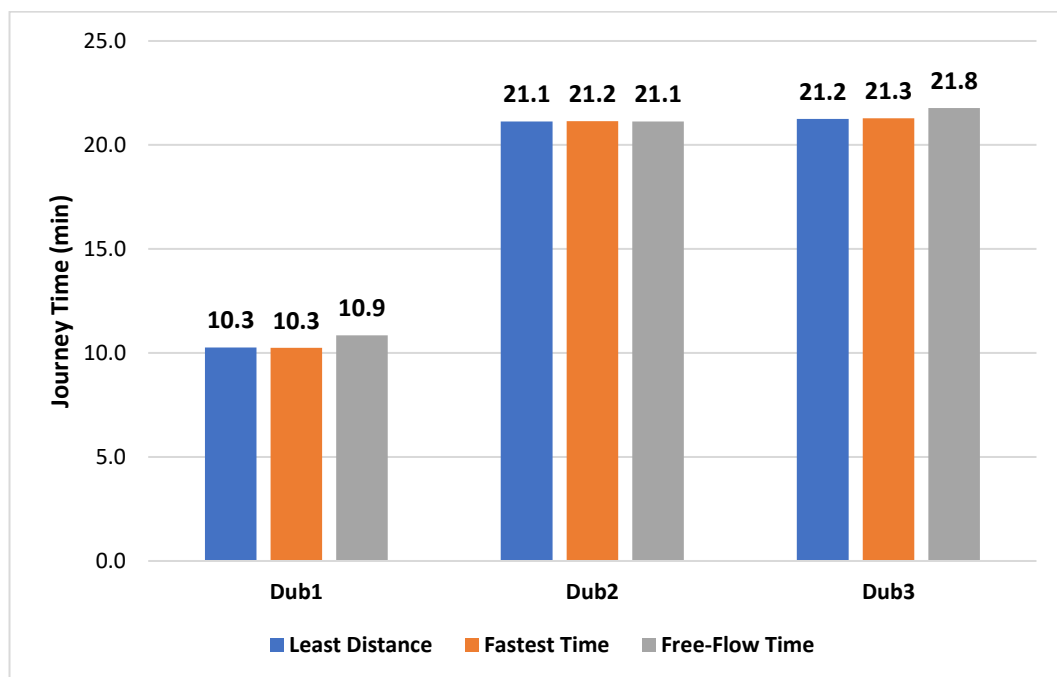


Figure 5.15. Travel Time by Route Choice on Urban Routes

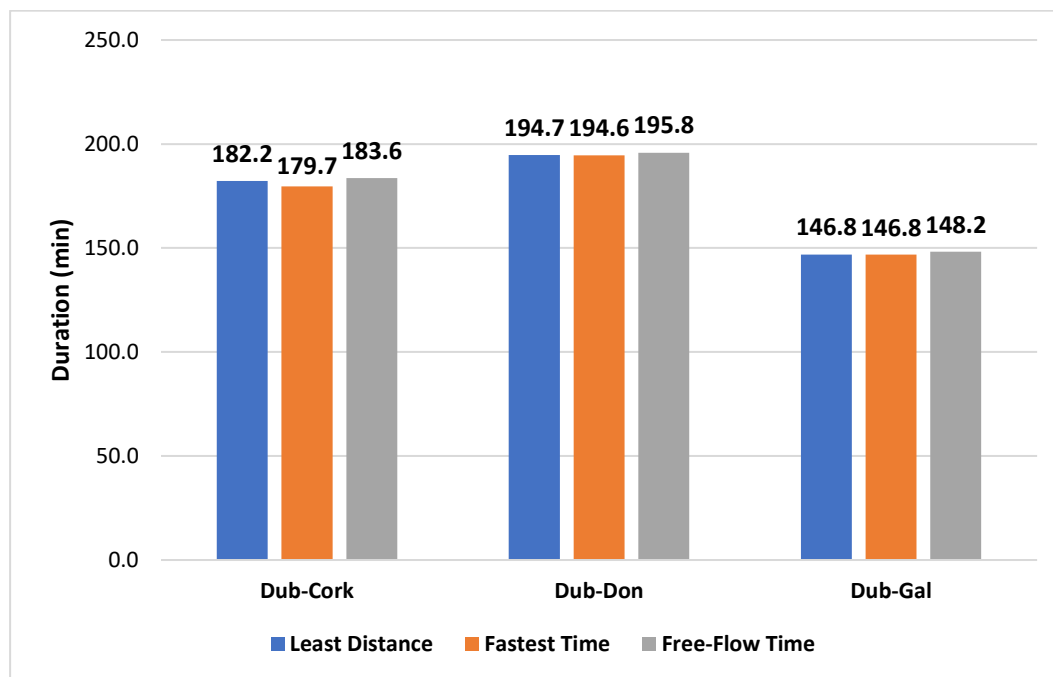


Figure 5.16. Travel Time by Route Choice on Interurban Routes

Figures 5.15 and 5.16 demonstrate that travel times were only minorly affected by route choice in this study. For all routes, the path that minimised current travel time does have the least trip duration, but again the significance of the difference highly depends on the route. Routes that minimise free-flow time without accounting for speeds updated by traffic consistently have the slowest (or equally slowest) travel times, even longer than minimum distance routes. To fully understand how paths differ between the different criteria, information on a third parameter, distance, is needed. If trip distances do not differ much between routes, that means there was a lack of suitable alternate routes. The distances by route choice for each route are shown in Figures 5.17 and 5.18.

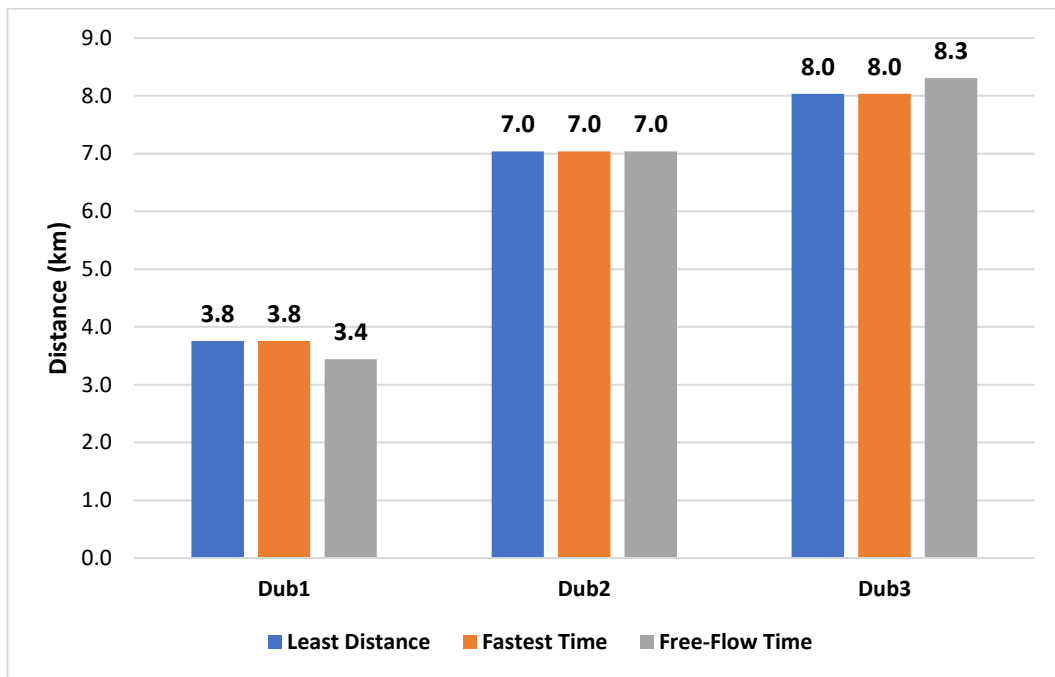


Figure 5.17. Distance Travelled by Route Choice on Urban Routes

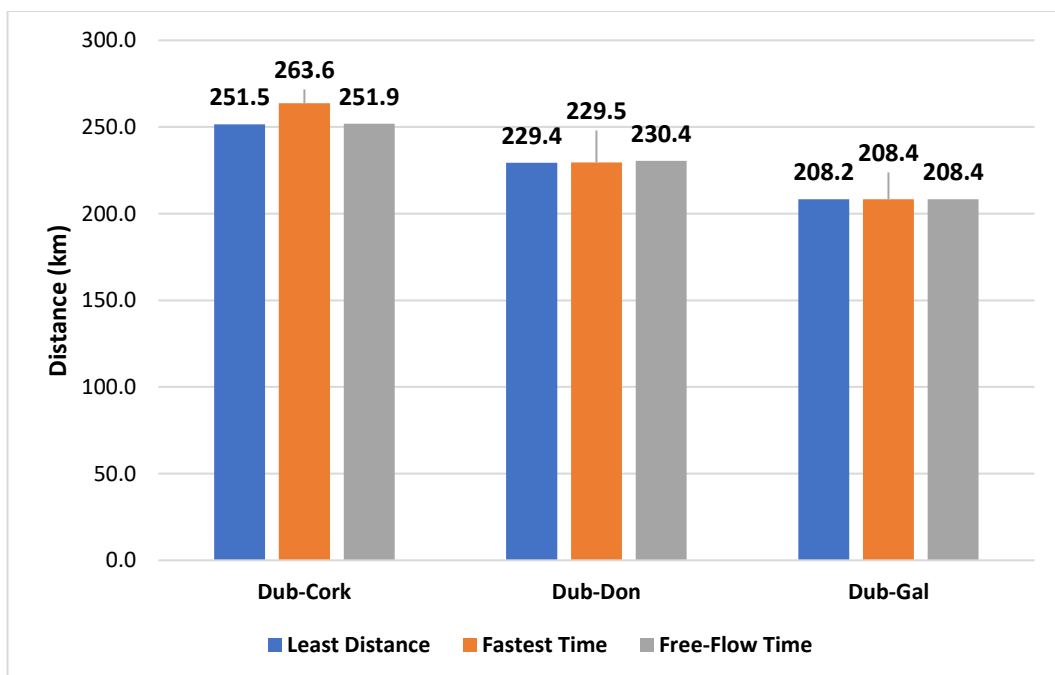


Figure 5.18. Distance Travelled by Route Choice on Urban Routes

Figures 5.17 and 5.18 show that, for many routes, the paths varied only slightly or not at all. For all three urban routes, the path chosen by least distance and fastest time are identical. The free-flow time path varies slightly on the Dub1 and Dub3 routes. This confirms that the NTpM is not the most detailed road network, so analysis of route choice

is hindered by the lack of available alternate routes. It could be seen in Figures 5.13-5.16 that the effect of route choice on EC and travel time was variable between routes. Figures 5.17 and 5.18 show that this is due to some routes having no alternatives and therefore providing the same path for multiple path search criteria. The differences for the Dub1 route in Figure 5.17 even indicate a failure in the shortest path search by least distance, as the free-flow time route has an even lower distance than the path found with the least distance criteria.

The route with the largest difference between the least distance and fastest time paths in Figure 5.17 or 5.18 is the Dub-Cork route. It is no coincidence, then, that the Dub-Cork route also shows the largest shifts in both EC and travel time between these path criteria. For the Dub-Cork route, it was shown that choosing the route which minimises travel time saves 1.48 kWh, 3.5% of the total EC of the route. By minimising distance, only 2.53 minutes is added in travel time, and the route is shorter by 12.17 km. The results of this section show that choosing a route that minimises distance has the potential to significantly reduce EC even without major changes to travel time, but the impact is heavily dependent on the alternative routes available.

5.4.3 Effects of Speed Limit Reduction on Energy Consumption and Travel Time

As the final investigation of this thesis, the effect of reducing speed limit on EC and travel time is analysed here. Section 5.3.2 outlines the speed limit reduction scenarios and how each speed limit was changed, but a reminder is included here. The speed limit changes are presented in Table 5.1. The base case provides an evaluation of the state of speed limits in Ireland before 2025. Speed limits are reduced on high-speed interurban links to produce Scenario A. Scenario B studies how lowering urban speed limits can affect EC and travel time. Scenario C combines all speed limit reductions on urban and interurban roads.

First, it is useful to view how the DCs change when speed limits are reduced. All analysis in this chapter is from DCs that are built to represent realistic driving styles on

routes taken from the NTpM. Reduced speed limits should cause visible changes in the DCs by shifting speeds downward and adding more time to the DC. Figures 5.19 and 5.20 below compare the DCs in the base case and Scenario C for the Dub-Cork route. These figures show the results of the DC creation methodology and allow a comparison in the driving between these two scenarios. The first of 100 DCs built for each scenario are shown in Figures 5.19 and 5.20, and only minor changes would be expected in the other 99 DCs.

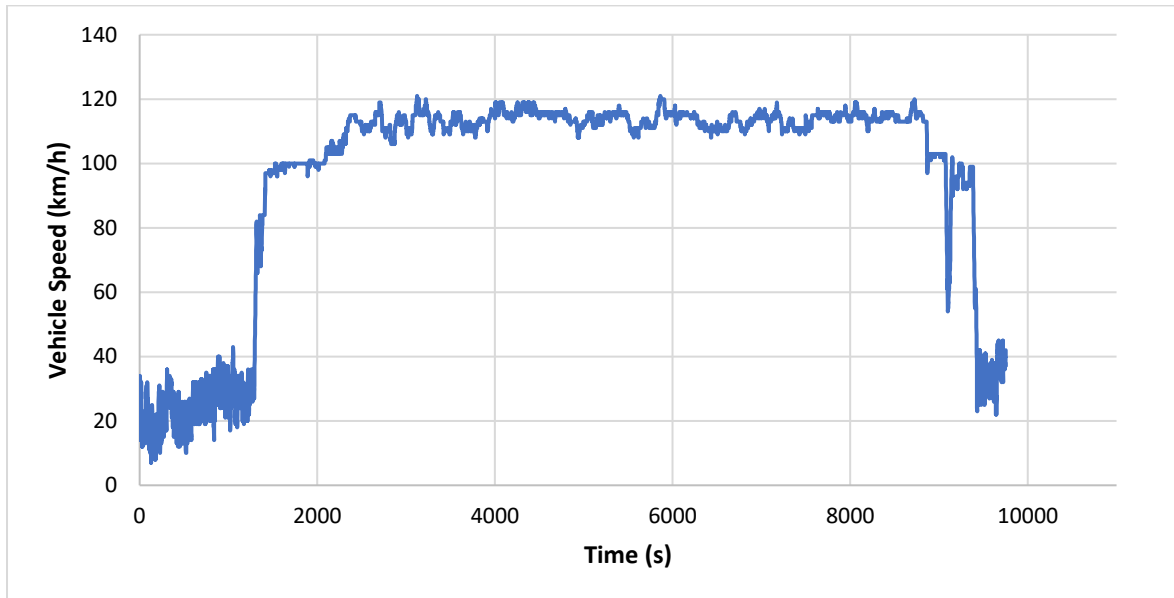


Figure 5.19. Dub-Cork Base Case Drive Cycle

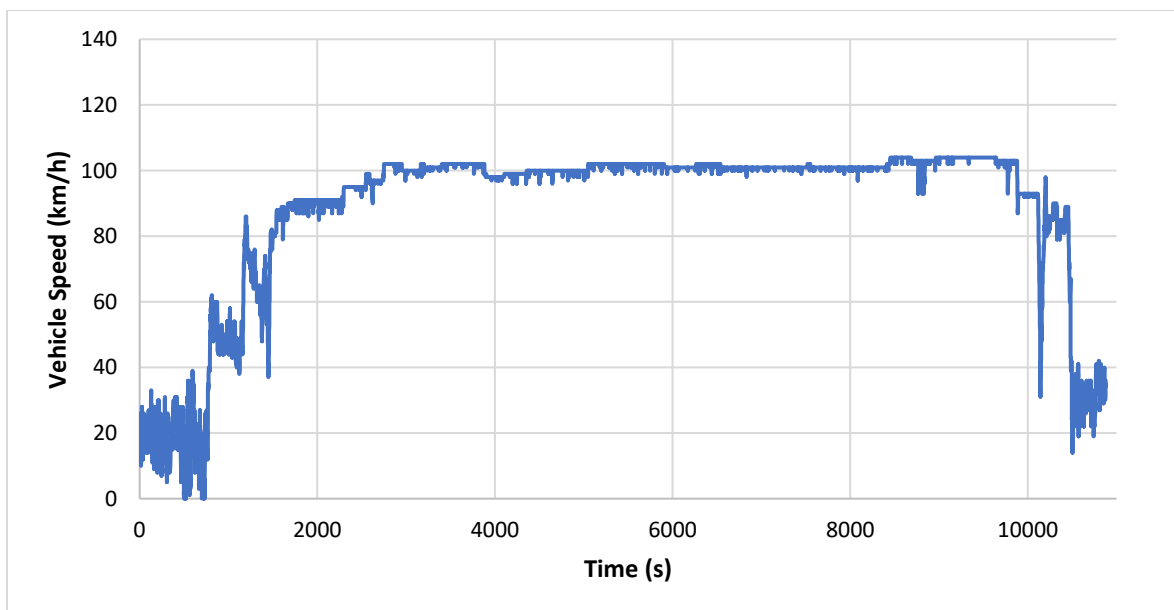


Figure 5.20. Dub-Cork Scenario C Drive Cycle

The differences between Figures 5.19 and 5.20 show the expected results; Scenario C in Figure 5.13 has lower speeds across the entire DC and is therefore longer by over 1,000 seconds. Both DCs exhibit steep acceleration and deceleration peaks in the urban areas at the start and end of the trip with more cruising behaviour on the motorway section. The DC of Figure 5.19 has noticeably less noise during cruising on the motorway section, likely due to differences in real world data collected at these speeds influencing the motorway TPM as there was a higher probability of remaining around 100 km/h. To see how the DCs created using the methodology in this thesis compare to real-world driving, the database of real-world trips was consulted. The first two trips found in the database that reached motorway speeds were taken to be shown in Figures 5.21 and 5.22.

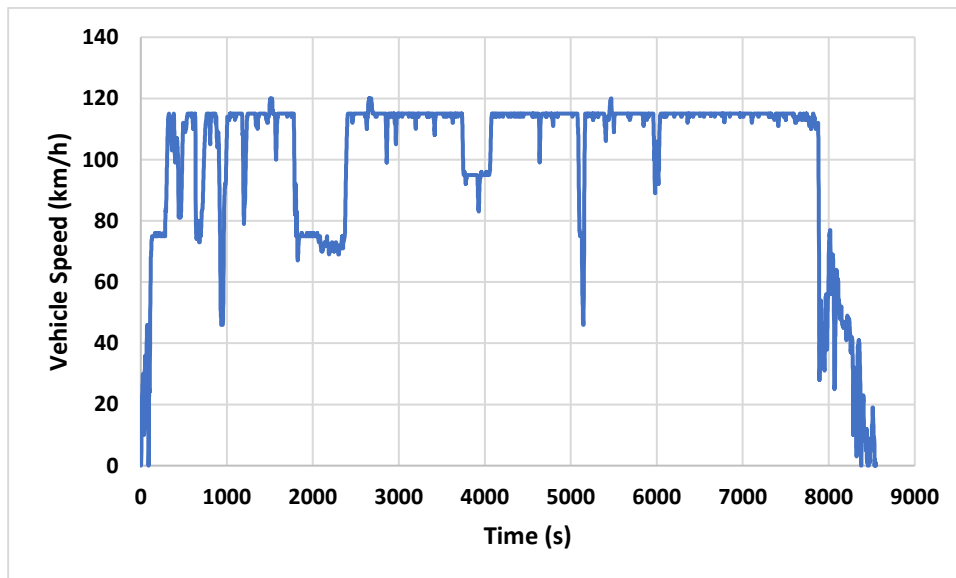


Figure 5.21. Real-World Drive Cycle Example 1

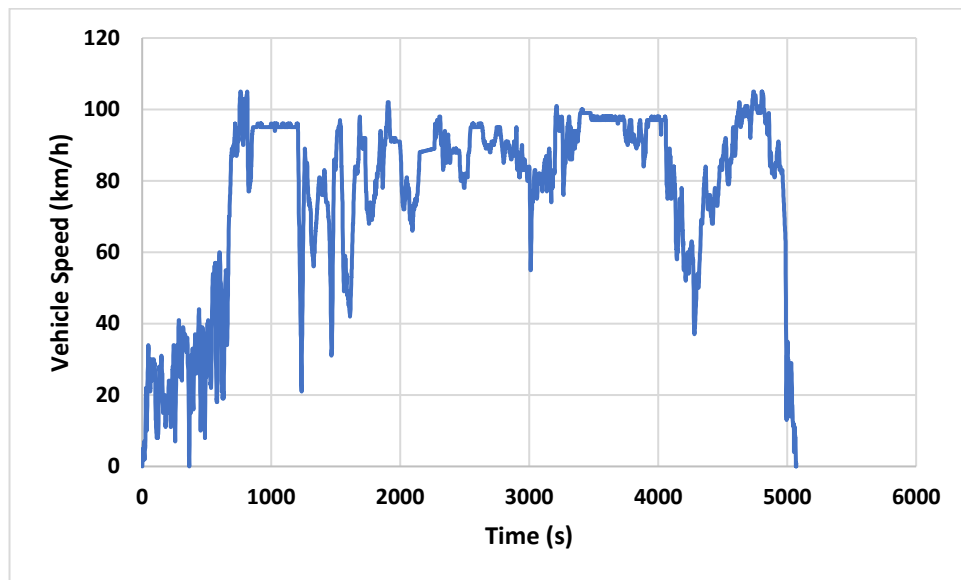


Figure 5.22. Real-World Drive Cycle Example 2

Figures 5.21 and 5.22 are included here to show that the DCs created for use in this chapter are formed of realistic driving styles. Although the DCs of Figures 5.21 and 5.22 are quite different than Figures 5.19 and 5.20 because they do not follow the same routes, they show that the individual slices of the created DCs mirror real-world driving behaviours. For instance, the lower-speed driving sections at the end of Figure 5.21 and the beginning of Figure 5.22 show similar acceleration and deceleration spikes to the urban sections that bookend the DCs in Figures 5.19 and 5.20. The high-speed driving section of Figure 5.21 closely resembles the motorway driving of Figure 5.19, likely showing automatic cruise control with minor speed adjustments. The high-speed sections of Figure 5.22 show more noise, and more closely resembles the motorway driving of Figure 5.19. In all, Figures 5.19-5.22 show that the DCs created in this chapter can represent realistic driving styles. As a contrast, Figure 5.23 shows the average speed of the links taken directly from the NTpM in Visum.

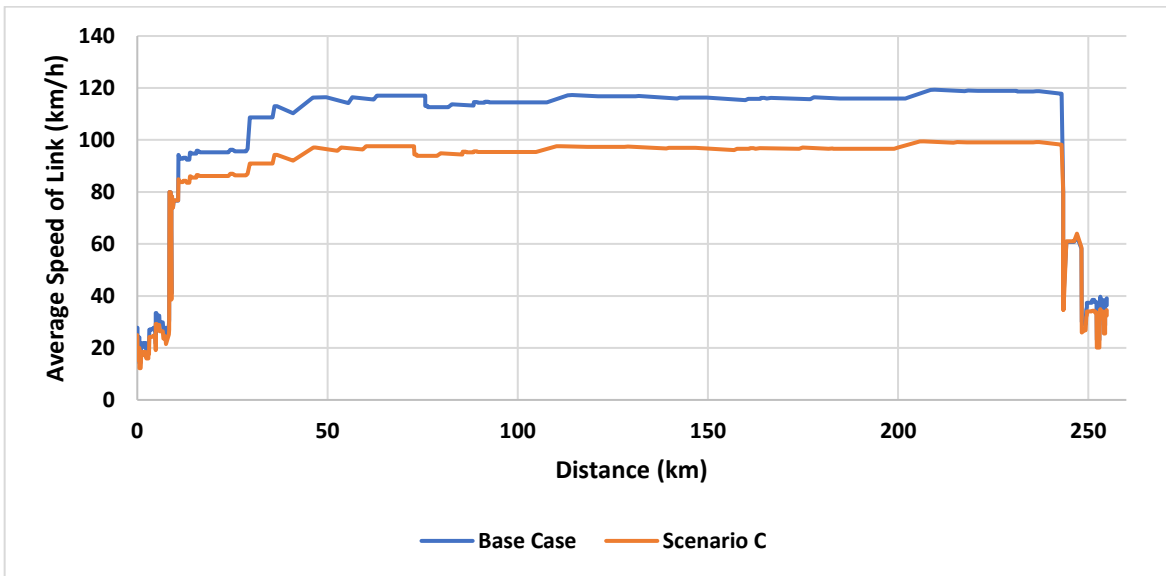


Figure 5.23. Average Link Speed of Dub-Cork Route: Base Case & Scenario C

Figure 5.23 shows how DCs make a significant improvement in representing realistic driving over the average speed information provided by Visum. The average link speeds result in long, flat cruising periods with minimal acceleration and deceleration between links. By using more realistic speed variations, more realistic EC estimations can be made using the methodology used here than state-of-the-art studies which use constant speeds, like what is seen in Figure 5.23, such as Basso et al. (2021). Figure 5.23 also provides a visual difference of how the speed limit reduction scenarios impact link speeds.

Now that it is established that the DC construction methodology creates realistic DCs for a given route, all that is left is to explore the results of the speed limit reduction scenarios. As stated previously, the parameters impacted by speed limit reduction included in this study are journey time and EC. First, journey times are compared across the speed limit reduction scenarios on all routes in Figures 5.24 and 5.25.

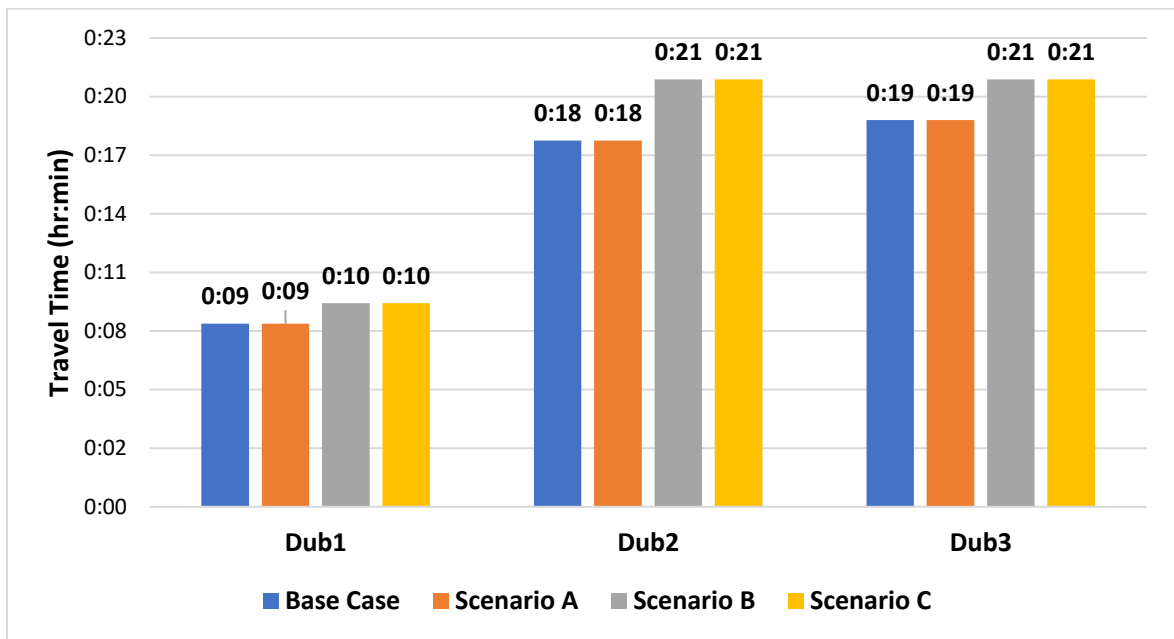


Figure 5.24. Journey Time by Speed Limit Scenario on Urban Routes

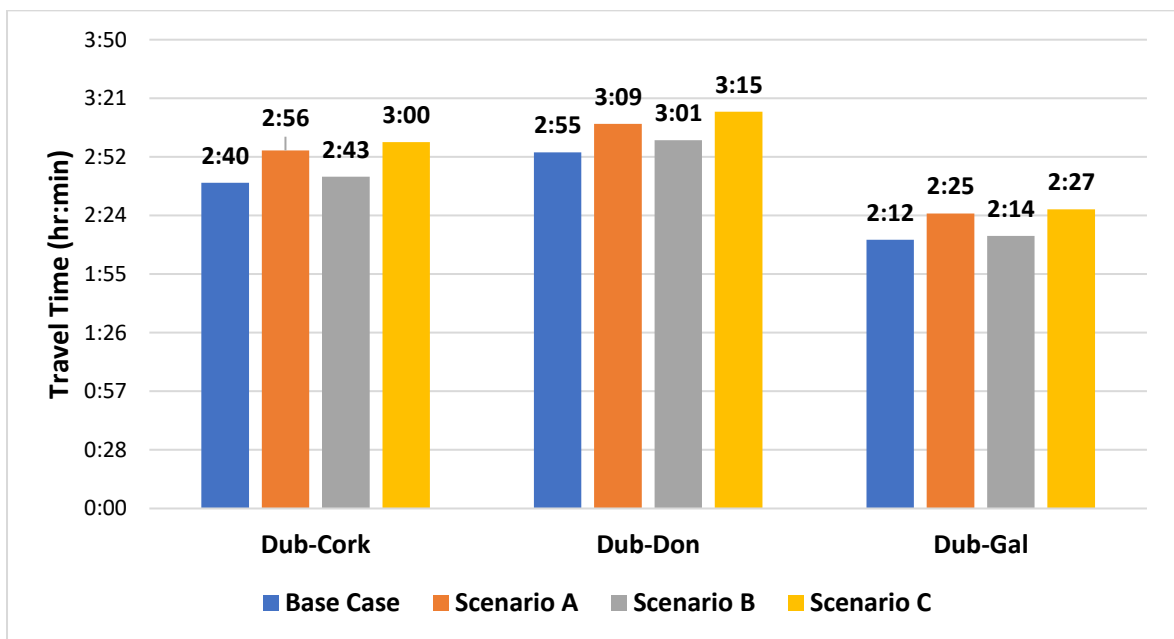


Figure 5.25. Journey Time by Speed Limit Scenario on Interurban Routes

Figures 5.24 and 5.25 demonstrate how, in most cases, when speed limits decrease, travel times increase. The exception in Figure 5.24 is for the urban routes in Scenario A, where none of the urban links that make up the route are affected. For the urban routes, lower speed limits result in a journey time increase of one to three minutes. This journey time increase is minimal, but it is likely to cause frustration to commuters and heavy traffic without demand reduction. The difference in travel time is more pronounced for the

interurban trips, where the reductions in Scenario C lead to journey time increases of up to twenty minutes. On average, Scenario A increases interurban journey time by 9%. Scenario B, which only impacts urban links, even increases the journey time of long, interurban trips by over 2%. The combination of both sets of speed limit reductions in Scenario C increases journey time by 11%.

Although reducing speed limits increases journey time, they can provide benefits of lowered emissions, increased safety, and increased fuel efficiency. The following analysis determines whether an additional benefit can be seen in the form of decreased EC in EVs. The effect of each speed limit scenario on EC can be found in Figures 5.26 and 5.27 below.

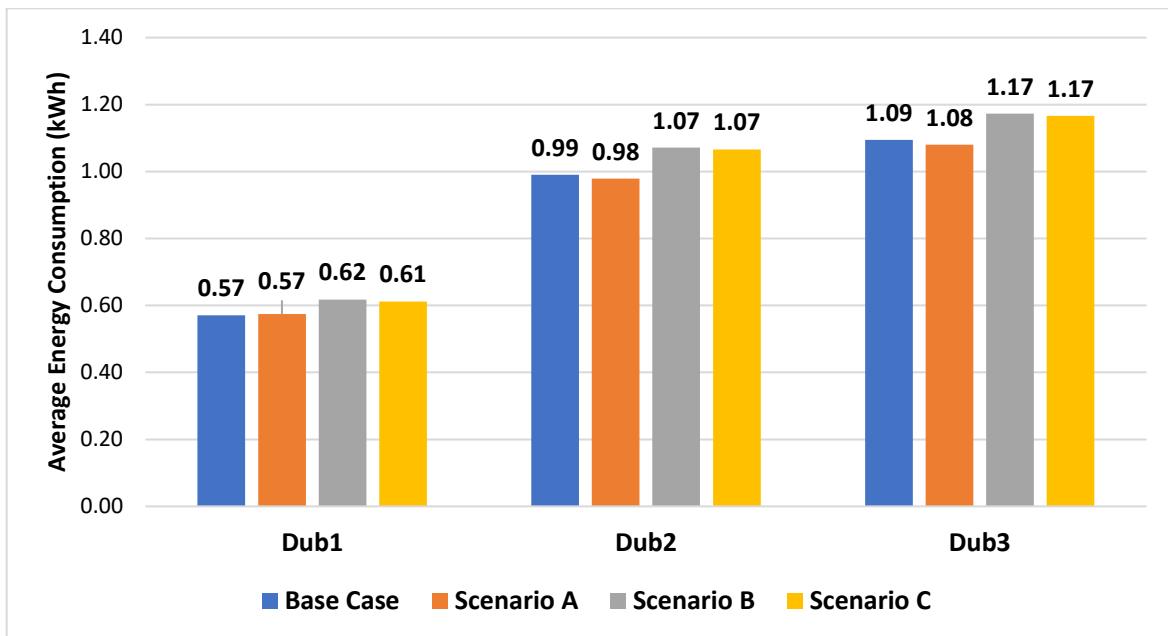


Figure 5.26. Energy Consumption by Speed Limit Scenario on Urban Routes

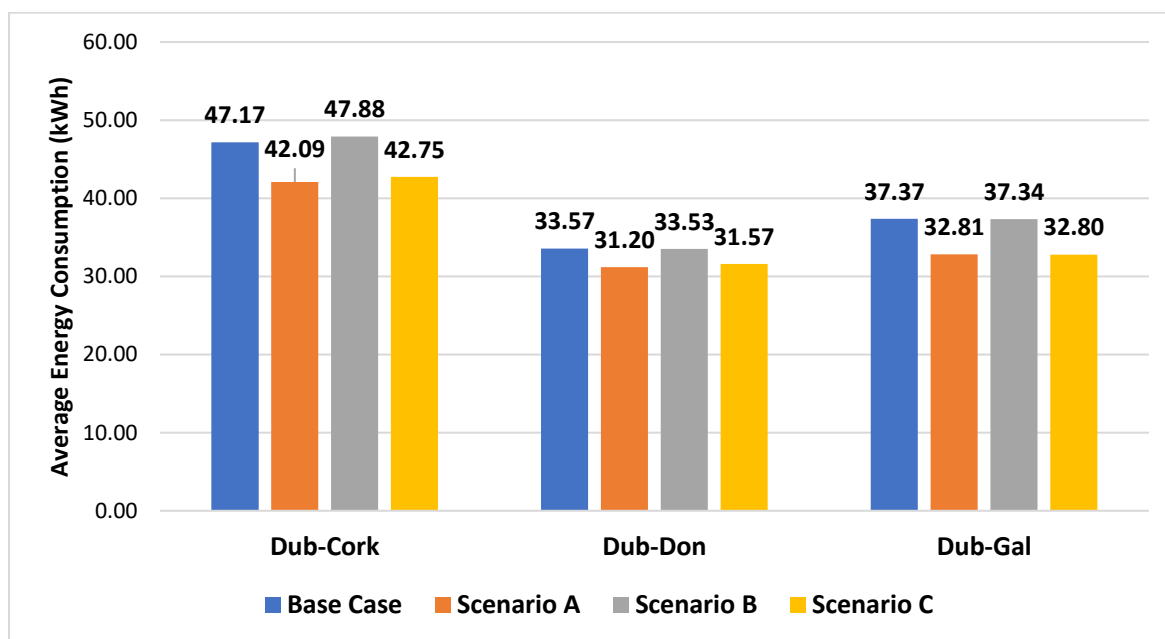


Figure 5.27. Energy Consumption by Speed Limit Scenario on Interurban Routes

For the urban trips, a reduction in speed limit actually results in a negative impact on EC as seen in Figure 5.26. In Scenarios B and C, where the urban speed limits were reduced, the EC on urban routes rose by an average of 7.8%. This is in addition to a 12.8% increase in travel time on the urban routes. Lowering speed limits in urban areas, which were already moderate at 30-40 km/h, has a negative impact on journey time and EC for urban trips. This is due to speeds dropping below the most efficient range of EC rate. For interurban trips, however, a significant positive effect on EC can be observed in Figure 5.27, especially for Scenarios A and C. In Scenario A, the EC on interurban trips decreased by an average of 10%. In Scenario C, this number decreased slightly to 9.2% due to the slower urban sections.

Again, ANOVA testing is used here to determine the statistical significance of the difference in EC as in Section 5.4.2. In this section, the ANOVA tests will determine whether the changing speed limits resulted in significant differences to EC. The results can be seen in Table 5.10.

Table 5.10. ANOVA Test Results for EC by Speed Limit Scenario

Route	F-Statistic	P-Value
Dub1	13.46	2.10E-07
Dub2	22.99	2.68E-11
Dub3	29.27	1.59E-13
DubCork	5790.08	2.62E-108
DubDon	546.74	3.41E-60
DubGal	3849.30	7.43E-100

Table 5.10 shows that, for each route, speed limit reductions lead to changes in EC with more than 99% confidence. The differences on longer interurban trips are even more pronounced. The comparison between longer journey times and EC can be visualised through Figure 5.28, which shows the results for a DC on the Dub-Cork route in the base case and Scenario C.

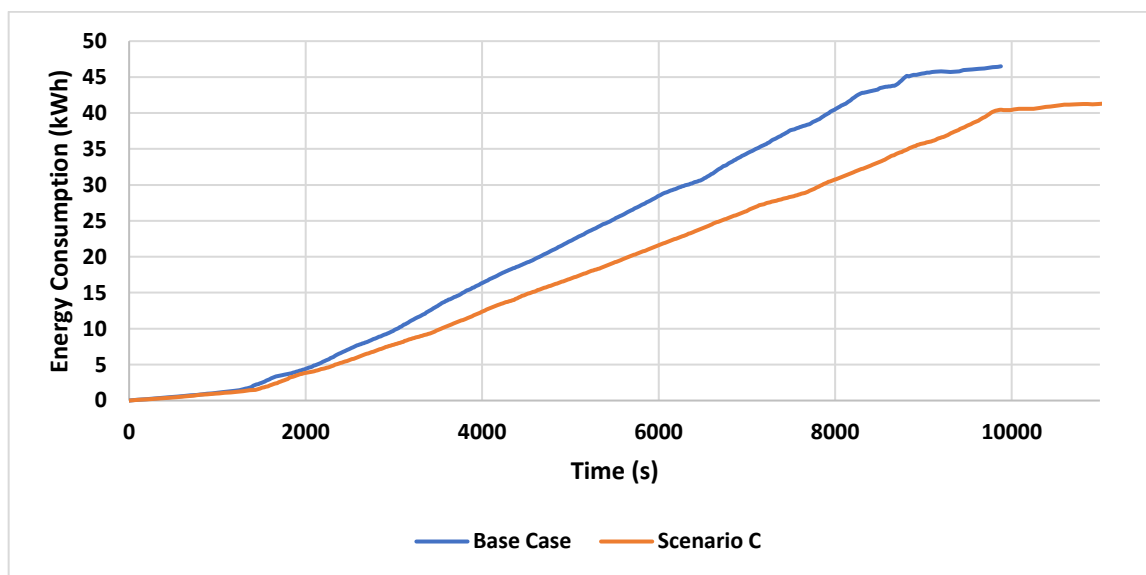


Figure 5.28. Energy Consumption for Dub-Cork Route – Base Case and Scenario C

Figure 5.28 highlights the trade-off between a longer journey time and a reduction in EC. In the urban sections, the EC is similar, and the slope is less steep. On the motorway sections, EC rises more drastically, but the slope is less steep in Scenario C. This result ties in directly with the EC rate figures in Section 4.6.3. When comparing 100 km/h speeds to 120 km/h speeds, the 120 km/h speeds were shown to have a higher EC per unit distance in

Chapter 4. The results here in Chapter 5 show how much this change can affect real-world trips. Slower driving may lead to longer trips, but it can also significantly decrease EC during interurban travel.

5.5 Discussion and Conclusion

The work done in this chapter focuses on how reducing speed limits affects EC and journey times for EVs. So far, no econometric model of a speed limit reduction policy has included EC of EVs as a parameter (Yannis and Michelaraki, 2025). The results of this chapter show that EC deserves to be included in these models by policymakers. The results showed that the modelled speed limit reductions can decrease EC by 10% while increasing travel time by 9%.

The effect of lowering speed limits on EC is significant. However, including EC as a parameter in an econometric model may be tricky. Take for instance Scenario A on the Dub-Cork route, where EC is lowered by 5.08 kWh compared to the base case. In the current economy in Ireland, this is equivalent to €3.00 using public fast chargers, €2.75 with public standard chargers, or €1.21 with at home charging (Electric Ireland, 2025; ESB, 2025). These differences, especially with the market share of EVs, could provide a significant economic benefit to consumers to help outweigh the economic losses of increased travel time. Still, the largest benefits could be nonmonetary.

The current discourse around EV adoption revolves heavily around EV range and range anxiety fears. As technology improves, this issue may be ameliorated. However, in the current EV market, 5.08 kWh is a significant portion of the total usable battery power in most EVs. Within the vehicles monitored in this thesis, 5.08 kWh is 9.28% of the average vehicle's total battery power. Banking an extra 5.08 kWh for “free” could be a highly significant factor to allay range fears. However, the effect of lowered speed limits on range anxiety is almost impossible to quantify. So, although EC is an important measurement for reduced speed limit studies, policymakers would struggle to monetarily quantify its effects.

Much of the discussion around enacting reduced speed limit policy would revolve around driver compliance and driver enthusiasm about such a change. In this study, strict compliance with the new speed limits is assumed as speed limits are assumed as free flow travel time, or the desired speed, in the NTpM. This means results here likely overestimate the differences if the speed limits were enacted, as a study in Italy found when the speed limit on rural roads dropped from 90 km/h to 80 km/h, observed speeds fell by just 3.3 km/h on average (Cerema, 2020). In the Speed Limit Review of Ireland, it is reported that total travel time on the network increased by 2.2% with lowered speed limits, so it is difficult to say how much the lowered speeds impacted individual journeys (Department of Transport, 2023B). With strong enforcement of lowered speed limits and encouragement by sharing the positive outcomes, similar results to this study may be observed. Even if drivers were less compliant with reduced speed limits, similar results could be expected to a lesser degree.

It is worth noting that the same results could be accomplished by drivers voluntarily reducing their speeds without needing a change in speed limit, but Filipi et al. (2022) showed drivers are generally unwilling to sacrifice speed voluntarily for minor improvements in energy efficiency. A survey by Wang et al. (2020), on the other hand, showed that many drivers are willing to trade journey time for energy efficiency at a reasonable rate. The results of this chapter show significant improvements to EC can be made with a moderate sacrifice of journey time. It is the job of future studies and policymakers to determine how much drivers are willing to sacrifice in journey time to make substantial gains in energy efficiency.

As part of the work conducted in this chapter, the effect of eco-driving behaviour on EC for realistic trips was examined. Because driving behaviour was not categorised and predetermined for different DCs, different driving behaviours tended to equalise on interurban trips, resulting in no significant differences. On urban trips, however, some conclusions could be drawn. High maximum accelerations and decelerations tended to result in more extreme EC results. Low and steady acceleration rates tended to reduce EC on

average. The results mimic those reported in Lee et al. (2024). If driving behaviour for a DC was predetermined by levels of aggressiveness, as in Iclodean et al. (2025) and Donkers et al. (2020), the effects of driving behaviour may be more pronounced. However, classifying driving behaviour risks losing the realistic driving styles analysed in this chapter. So far, no study has managed to classify driving behaviour in EVs that meaningfully results in consistent changes of EC.

The analysis on the effect of route choice on EC showed the possibility that routes which minimise distance can significantly improve EC compared with routes that minimise time. The analysis was limited in robustness due to the lack of alternative links included in the NTpM, which only includes major roads. An exciting, related topic is green vehicle routing or eco-routing, where a navigation system can direct a driver to the most energy efficient routes using a combination of historical and real-time data. Such a concept has been suggested in Zhou et al. (2016A), Wang et al. (2018), and Straub et al. (2022), among others. Still, the field is young, especially considering EVs compared to ICEVs. The results of this chapter show that an eco-routing software can significantly improve EC, especially on longer routes where range anxiety is most common.

Chapter 6. Discussion

6.1 Contributions

The work done in this thesis sits in a large pool of literature covering EC in EVs. As an emerging technology, new research is published every day in the field of EVs. There are many areas where additional research is needed to add to a scientific understanding and reach a consensus. Additionally, some gaps in knowledge were discovered through the several literature reviews of this thesis that the work in Chapters 3-5 helped to fill. This chapter adds to the discussion included in Chapters 3-5 by framing the work of this thesis in the broader context of the literature and summarising how the results can be used in practice as well as some weaknesses that could be improved on in future studies.

A major contribution of this thesis to the field of EC modelling is the comparison of data-driven EC models at the macroscopic and mesoscopic scales. Mesoscopic models appear to be underutilised in the literature, especially when considering that mesoscopic models are a direct substitute for macroscopic models with less data aggregation. Mesoscopic models have gained more attention recently, although the literature has generally used MTs as a measure of segmentation (Pan et al., 2023). It was demonstrated by Petersen and Sax (2022) that other segmentation methods can improve EC predictions, so this study used 500m segments instead. A clear and necessary extension of this thesis and the work of Petersen and Sax (2022) would be to apply the methodology to other segmentation methods to see how segmentation affects the prominence of different parameters within the model.

With the mesoscopic and macroscopic models tested in this thesis, it was shown that macroscopic models can tend to rely on trip distance for EC predictions. This can lead to underestimation of other parameters that affect EC at a lower level when trips of all distances are modelled together. An example is in Ullah et al. (2022), which likely underestimates the effects of factors such as speed and gradient due to the high level of correlation between

distance and EC. As a result, future studies should focus on exploring the potential of mesoscopic models. Macroscopic models will still have a place in the literature, especially when data is extremely infrequent such as in Corti et al. (2012). This result also gives a general warning on using black-box ML methods in that the appropriate model for each study must be selected.

The work done on driving behaviour in Chapter 4 provides some novel insight on the connection between driving behaviour and EC. In the literature, there is little reflection on how the three commonly ascribed categories of driving behaviour (passive, moderate, and aggressive) can affect EC differently at different speeds. Zhang et al. (2022) was one of the first to evaluate the influence of driving behaviour on fuel economy accounting for the dependence on speed. This thesis extends the analysis to EVs, where regenerative braking provides a new dimension for consideration. The major result shows that common driving behaviour classifications do not directly correlate with changes in EC across speed levels, which broadens the recent findings of Iclodean et al. (2025), which were limited to a few DCs. This encourages a reflection on how driving behaviour is categorised and studied. Future work would look into how driving behaviours differ along a constant route to determine proportional divisions of driving behaviour seen in real driving. Such a study could utilise the empirical method described in Section 4.4 to divide driving behaviours by EC, with the difference being all data is taken from an identical route whereas in this thesis routes taken during data collection were not predetermined.

Finally, Chapter 5 applied the EC model to realistic DCs built on routes from the NTpM. This chapter provides an innovation to the field by including real driving behaviour in estimations of EC taken from a road network model. Addressing this gap was the driving idea behind the entire scope of this thesis. The gap in the literature was mentioned specifically in guiding works such as De Cauwer et al. (2017), which states “individual driving style can modify the speed profile... As this information was not present in the

available datasets, it presents some limitations of the model in its current state.” Similarly, Straub et al. (2022) states “prospective work will feature a driver model based on distributions of the presented driving profile parameters sourced from a fleet.” After a thorough literature review, no study was found which filled this gap before this thesis. These works, together with the follow-up work of De Cauwer et al. (2020), are in the specific field of green vehicle routing, which involves selecting the most energy efficient route.

Although not explored in the scope of this thesis, the work completed here can have real applications in the green vehicle routing problem. The green vehicle routing problem involves calculating the route for one vehicle or a fleet that minimises fuel consumption, energy consumption, or emissions. The field began by seeking to lower emissions of ICEVs but naturally extended to EC of EVs (Asghari and Al-e-hashem, 2021). Historically, many implementations of green vehicle routing have assumed speed to be constant (Abousleiman and Rawashdeh, 2015; Karbowski et al., 2012). Karbowski et al. (2012) even noted that future studies could improve on their work by introducing variable speed patterns to simulate real-world driving behaviour more accurately. Some models include stochastic speed estimations based on slight variations from a link’s mean speed (Thorgeirsson, 2024).

Green vehicle routing studies is the most prominent area in the literature that combines EC modelling and road network modelling like what was done in this thesis. Martins et al. (2021) shows that publications in the field of green routing of electric vehicles increased dramatically from 2012 to 2021, so there is growing interest in the subject. By building DCs from realistic data that match a route from a road network model, this thesis has provided a methodology to the literature that can be replicated and used in future studies. As a result, this thesis was able to calculate for the first time how much changes in driving behaviour can affect EC on real-world trips using realistic driving data. The methodology created here shows how future studies can build DCs to match a road network model to estimate how differences in realistic driving behaviours can influence EC along a route in a

network. An exciting application could incorporate federated learning (Saputra et al., 2019; Thorgeirsson, 2024), where an ML model is trained using each individual's data instead of the generic vehicle EC model developed in Chapter 3.

In summary, the work in this thesis adds a thorough, scientific analysis of eco-driving behaviours to the literature. There are very few scientific papers that analyse EC in EVs in a way that is helpful for EV drivers. This thesis was undertaken to analyse how factors either in control of drivers or easily understood by drivers, such as usable battery power and weather parameters, influence EC. Parameters such as battery voltage and current, which have been used to predict EC in studies such as Sasank et al. (2023), were not included in the EC model of this thesis as drivers generally have no direct knowledge or control of them. The results of Chapters 4 and 5 can be used to educate drivers on how acceleration, deceleration, and speed can affect EC and remaining range on real-world trips. EV drivers do not have a complete understanding of the effectiveness of eco-driving behaviours, and the literature does not appear to be aimed toward the education of drivers on this subject (Wang et al., 2020). Specifically, the graphs of modelled and actual average ECR versus speed of Figures 4.15 and 4.17 can aid in the education of eco-driving behaviours by showing how the most energy efficient driving speeds are around 20-60 km/h, although the difference is most significant at speeds above 100 km/h. These figures can be supplemented by the results of Chapter 5 demonstrating how adding just a few minutes of travel time can significantly decrease the EC of a long-distance trip and accordingly soothe any fears from range anxiety. Figures 5.25 and 5.27 specifically provide a strong argument in favour of slowing down on high-speed roads when EV range appears to be a limiting factor of a trip. Quantitative information on the trade between journey time and EC is rare if present at all in the literature, with similar information needing to be inferred from studies such as Wager et al. (2016), which may give outdated information.

Other groups that can benefit from the work in this thesis are policymakers and road designers. With knowledge on just how much speed can affect EC in EVs, the most energy efficient speeds for EVs can be prioritised more on roadways. Specifically, development on urban and suburban roads can be encouraged to maintain speeds between 20-60 km/h. On high-speed roads, an argument here is presented to lower the desired speeds to be closer to 100 km/h in order to encourage EV use and minimise range anxiety. Lower speeds on high-speed roads could ease the urgent burden on policymakers to rapidly expand charging infrastructure by increasing energy efficiency on interurban trips through lowered speeds. On another note, the results of Chapter 5 encourage the inclusion of EC as a factor in econometric models when evaluating speed limit reduction policies. EC of EVs and fuel consumption of ICEVs is rarely if ever included in such studies, although the impact of reducing speed limits on these parameters was shown to be significant.

6.2 Limitations

Although this work provides strong results gathered through robust methodologies, there are some limitations that could be improved upon in future studies. First, the EC model was built using elevation data gathered through GPS. While GPS provides reasonable measurements of elevation, improvements could be made to accuracy and consistency of gradient measurements by using a higher-fidelity elevation map, like the one utilised in Chapter 6 (Jacko et al., 2024). This improvement could lead to a more accurate assessment of gradient's impact on EC. Furthermore, measurements such as load of the vehicle, auxiliary load, and tyre pressure, which have been shown to influence EC, were unknown for each trip and therefore not included in the EC model here (Agrawal et al., 2023; Liu et al., 2018; Settey et al., 2021). Future studies should include these variables to gain a better understanding of their effect on EC in comparison with the parameters studied in this thesis.

The analysis of eco-driving behaviours was limited to just three of the many classification methods used in the literature. Other methods exist in the literature, such as

can be seen in Hong et al. (2021), Lai et al. (2023), and Zhang et al. (2020A), among others, that may provide different results to the ones seen in Chapter 4. Time constraints of implementing new methodologies are the main reason only three methods of classification were tested in this thesis, so future studies should be able to expand the processes used here with little difficulty. In addition, gas and brake pedal position has been used to help categorise aggressiveness in previous studies, but that parameter was unavailable in this study (Cheng et al., 2018; Nan et al., 2022; Wei et al., 2023). It is understandable how driver aggression can be revealed further by the inclusion of these parameters, so future studies should incorporate pedal position when possible.

Due to the lack of clear outcomes in the SA of Chapter 4, additional analysis was conducted using synthetic segments to evaluate how different speeds and rates of acceleration and deceleration can affect EC in practice. In order to produce clear results, these segments were limited to a few discrete acceleration rates and speeds. In addition, the segmentation used to build the EC model meant that driving behaviour could not be analysed at a smaller scope than a 500m segment. Small differences in the steadiness of acceleration, for instance, could not be analysed directly using this model. A future study could investigate a wider range of driving scenarios than the segments produced in Chapter 4. Additionally, a different model could provide more insight of driving behaviour at a smaller scale.

There were several limitations of Chapter 5 due to the use of the NTpM. Foremost, the NTpM models the road network by factoring idle times into average speeds. The junctions are not modelled, so it was not known which links were connected by stop signs, traffic signals, or nothing. Because of this, idle periods were not included in the created DCs because it could not be known when an idle time was likely to occur. Some previous studies have probabilistically modelled traffic signals (Hu et al., 2018). For this thesis, it was decided to not account for the possibility of traffic signals in favour of maintaining the average speeds of the NTpM that were used to validate modelled journey times (AECOM,

2019A). Optimising signal times after reducing speed limits could possibly mitigate the increase in EC on urban roads, but this possibility was unable to be studied in this thesis.

Also, it was mentioned in Chapter 5 that the road network of the NTpM does not include enough roads to present many viable alternate routes. This limited the analysis on how route choice can affect EC. Accordingly, both of these limitations could be amended in future studies by using a more detailed road network model. Another limitation was the lack of a specific driving behaviour categorisation method used to build the DCs in Chapter 5. Future studies could build DCs that match different driving styles, although this was not done in this thesis because no classification method was found that correlated with EC in Chapter 4.

Chapter 7. Conclusion

EVs provide the potential to be a significant part of the global solution to combat climate change. For years, policies have been enacted worldwide to encourage more drivers to transition to EVs with zero local emissions. However, these policies have yet to encourage the substantial portion of drivers to adopt EVs in order to reach the emissions goals set by the legislative bodies. EVs have not yet reached the level of social acceptance that was theorised. Although many reasons have been suggested for the limited adoption of EVs, range anxiety stands out as a constant source of deterrence. A better understanding of the real-world causes of variations in EC can help to alleviate the stresses due to range anxiety and help encourage more social trust in EVs.

Chapter 2 described the methodology used to collect the data used within this thesis. As part of the European Union's POWERDRIVE project, speed, GPS, and battery information was tracked during participants' everyday trips. In this study, the data from 36 different EVs and drivers was essential to be able to complete the work done here. Drivers used the Car Scanner app and OBDLink MX+ dongle to provide trip data.

In Chapter 3, these data were used to form a data-driven EC model. A literature review established the norms for state-of-the-art EC modelling. For the first time, an analysis was conducted in detail to show the differences in parameters included in a mesoscopic model versus a macroscopic model. The macroscopic model was shown to be overly reliant on distance as an input parameter, meaning differences in driving behaviour along a route may not be captured. In a comparison of ML methods, ANNs were shown to provide the best results on both the macroscopic and mesoscopic models. The results of Chapter 3 show that future research should primarily utilise ANNs to build data-driven models. Additionally, the limitations of macroscopic modelling revealed in this chapter shed light on the subject for future researchers. As a result, future data-driven models should discretise trips into segments to form mesoscopic models, which may result in some additional computational

time but make up for that with an improved ability to adjust predicted EC to different driving behaviours.

Driving behaviour was analysed in Chapter 4, including several common methods of classifying driving behaviour. A review of the literature was used to understand the current knowledge around eco-driving behaviours and to identify common driving behaviour classification methods. A second literature review investigated the field of sensitivity analyses. The goal of Chapter 4 was to find classifications of driving behaviour that could reliably result in differences in EC for use in Chapter 5. The SA showed that, even when accounting for interaction, most parameters have little effect on EC relative to the major parameters of average speed, time spent accelerating, and time spent decelerating. The SA showed that, under all realistic driving scenarios, EVs are most efficient in the 20-50 km/h range. At speeds above 100 km/h, EC per unit distance increases dramatically.

To complete the work of Chapter 4, synthetic driving segments were built, and the EC model was used to determine how different speeds and rates of acceleration and deceleration affect EC at low and high speeds. The results showed that variations in speed between 20-60 km/h have a relatively minor effect on EC. The results showed that moderate rates of acceleration and deceleration are effective eco-driving behaviours. The clustering and limit-based methods of classifying driving behaviour were shown to be limited in their correlation to EC. Therefore, it was decided that the empirical method would be used to classify driving behaviour of the DCs in Chapter 5.

Chapter 5 merged the EC model with the NTpM road network model of Ireland. First, separate literature reviews established state-of-the-art practices in DC construction and reduced speed limit policy modelling. In Chapter 5, routes were extracted from the NTpM, and DCs were built to simulate realistic driving behaviours along those routes. This methodology provides a basis for future studies to incorporate in green vehicle routing applications. The results showed how driving behaviours can cause significant differences

in EC at a trip level. In addition, it was demonstrated that, when alternative routes are available, minimising distance instead of travel time can lead to consuming less energy on a route. The developed methodology was used to test reduced speed limit scenarios on the network. The results showed that reduced speed limit policies can greatly increase energy efficiency of interurban trips. At the urban level, however, negative effects were seen in both travel times and EC. The effects on EC were great enough to suggest that future studies of reduced speed limit policies should include EC and fuel consumption when comparing the benefits and drawbacks of such policies. It was also shown that generally drivers can sacrifice modest amounts of travel time for a significant decrease in EC, meaning that lowering speed is an effective eco-driving behaviour that could significantly reduce range anxiety.

References

- Abburu, S. K., O'Reilly, C. J., & Casanueva, C. (2024). Network Theory and Global Sensitivity Analysis Framework for Navigating Insights from Complex Multidisciplinary Models. *IEEE Access*, 12(October), 157201–157217. <https://doi.org/10.1109/ACCESS.2024.3486358>
- Abdel-Wahed, A. T., Ullah, Z., Abdel-Khalik, A. S., Hamad, M. S., Ahmed, S., & Elmalhy, N. (2023). Drive Cycle-Based Design with the Aid of Data Mining Methods: A Review on Clustering Techniques of Electric Vehicle Motor Design with a Case Study. *IEEE Access*, 11(October), 115775–115797. <https://doi.org/10.1109/ACCESS.2023.3325400>
- Abediasl, H., Balazadeh Meresht, N., Alizadeh, H., Shahbakhti, M., Koch, C. R., & Hosseini, V. (2023). Road transportation emissions and energy consumption in cold climate cities. *Urban Climate*, 52(August), 101697. <https://doi.org/10.1016/j.uclim.2023.101697>
- Abohassan, A., Contini, L., Elmasry, H., & El-Basyouny, K. (2024). Assessing the effectiveness of speed limit reduction in Edmonton: A case study analysis. *Accident Analysis & Prevention*, 195, 107379.
- Abousleiman, R., & Rawashdeh, O. (2015). A bellman-ford approach to energy efficient routing of electric vehicles. *2015 IEEE Transportation Electrification Conference and Expo, ITEC 2015*, 1–4. <https://doi.org/10.1109/ITEC.2015.7165772>
- Abukhalil, T., Almahafzah, H., Alksasbeh, M., & Alqaralleh, B. A. Y. (2020). Fuel Consumption Using OBD-II and Support Vector Machine Model. *Journal of Robotics*, 2020. <https://doi.org/10.1155/2020/9450178>
- Achour, H., & Olabi, A. G. (2016). Driving cycle developments and their impacts on energy consumption of transportation. *Journal of Cleaner Production*, 112, 1778–1788. <https://doi.org/10.1016/j.jclepro.2015.08.007>
- Adnane, M., Khoumsi, A., & Trovão, J. P. F. (2023). Efficient Management of Energy Consumption of Electric Vehicles Using Machine Learning—A Systematic and Comprehensive Survey. *Energies*, 16(13). <https://doi.org/10.3390/en16134897>
- AECOM. (2019A). National Transport Model Update – Model Development Report. <https://www.tii.ie/media/eqbmenfj/ntpm-vol1-model-development-report.pdf>
- AECOM. (2019B). National Transport Model Update – Data Collection Report. <https://www.tii.ie/media/gvmhwyrw/ntpm-vol2-data-collection-report.pdf>
- AECOM. (2021). National Transport Model Update – Variable Demand Model Report. <https://www.tii.ie/media/h0odc1do/ntpm-vol4-variable-demand-model-report.pdf>
- Agrawal, A., Singh, R., Kumar, N., Singh, V. P., Alotaibi, M. A., Malik, H., Marquez, F. P. G., & Hossaini, M. A. (2023). Mathematical Modeling of Driving Forces of an Electric Vehicle for Sustainable Operation. *IEEE Access*, 11(August), 95278–95294. <https://doi.org/10.1109/ACCESS.2023.3309728>

- Ahmed, M., Mao, Z., Zheng, Y., Chen, T., & Chen, Z. (2022). Electric Vehicle Range Estimation Using Regression Techniques. *World Electric Vehicle Journal*, 13(6). <https://doi.org/10.3390/wevj13060105>
- Alcaraz Carrillo de Albornoz, V., Molina Millán, J., Lara Galera, A., & Muñoz Medina, B. (2022). Road speed limit matters – Are politicians doing the right thing? *Socio-Economic Planning Sciences*, 79(October 2020). <https://doi.org/10.1016/j.seps.2021.101106>
- Alhanouti, M., & Gauterin, F. (2024). A Generic Model for Accurate Energy Estimation of Electric Vehicles. *Energies*, 17(2). <https://doi.org/10.3390/en17020434>
- Ali, A. M., & Asfoor, M. S. (2023). Optimal Battery Sizing and Stops' Allocation for Electrified Fleets Using Data-Driven Driving Cycles: A Case Study for the City of Cairo. *IEEE Transactions on Transportation Electrification*, 9(1), 896–911. <https://doi.org/10.1109/TTE.2022.3160615>
- Ali, M. U., Khan, H. F., Masood, H., Kallu, K. D., Ibrahim, M. M., Zafar, A., Oh, S., & Kim, S. (2022). An adaptive state of charge estimator for lithium-ion batteries. *Energy Science and Engineering*, 10(7), 2333–2347. <https://doi.org/10.1002/ese3.1141>
- Alkinani, M. H., Khan, W. Z., & Arshad, Q. (2020). Detecting Human Driver Inattentive and Aggressive Driving Behavior Using Deep Learning: Recent Advances, Requirements and Open Challenges. *IEEE Access*, 8, 105008–105030. <https://doi.org/10.1109/ACCESS.2020.2999829>
- Al-Wreikat, Y., Serrano, C., & Sodré, J. R. (2021). Driving behaviour and trip condition effects on the energy consumption of an electric vehicle under real-world driving. *Applied Energy*, 297(January). <https://doi.org/10.1016/j.apenergy.2021.117096>
- Al-Wreikat, Y., Serrano, C., & Sodré, J. R. (2022). Effects of ambient temperature and trip characteristics on the energy consumption of an electric vehicle. *Energy*, 238. <https://doi.org/10.1016/j.energy.2021.122028>
- Amirjamshidi, G., & Roorda, M. J. (2015). Development of simulated driving cycles for light, medium, and heavy duty trucks: Case of the Toronto Waterfront Area. *Transportation Research Part D: Transport and Environment*, 34, 255–266. <https://doi.org/10.1016/j.trd.2014.11.010>
- André, M. (2004). The ARTEMIS European driving cycles for measuring car pollutant emissions. *Science of the Total Environment*, 334–335, 73–84. <https://doi.org/10.1016/j.scitotenv.2004.04.070>
- André, M., Joumard, R., Hickman, A. J., & Hassel, D. (1994). Actual car use and operating conditions as emission parameters: derived urban driving cycles. *Science of the Total Environment*, The, 146–147(C), 225–233. [https://doi.org/10.1016/0048-9697\(94\)90241-0](https://doi.org/10.1016/0048-9697(94)90241-0)

- André, M., Joumard, R., Vidon, R., Tassel, P., & Perret, P. (2006). Real-world European driving cycles, for measuring pollutant emissions from high- and low-powered cars. *Atmospheric Environment*, 40(31), 5944–5953. <https://doi.org/10.1016/j.atmosenv.2005.12.057>
- Apple. (2025). *B On Track*. <https://apps.apple.com/us/app/b-on-track/id916723519>
- Arthur, D., & Vassilvitskii, S. (2006). k-means++: The advantages of careful seeding. *Stanford*.
- Arun, N. H., Mahesh, S., Ramadurai, G., & Shiva Nagendra, S. M. (2017). Development of driving cycles for passenger cars and motorcycles in Chennai, India. *Sustainable Cities and Society*, 32(April), 508–512. <https://doi.org/10.1016/j.scs.2017.05.001>
- Asamer, J., Graser, A., Heilmann, B., & Ruthmair, M. (2016). Sensitivity analysis for energy demand estimation of electric vehicles. *Transportation Research Part D: Transport and Environment*, 46, 182–199. <https://doi.org/10.1016/j.trd.2016.03.017>
- Asef, P., & Laphorn, A. (2021). Overview of Sensitivity Analysis Methods Capabilities for Traction AC Machines in Electrified Vehicles. *IEEE Access*, 9, 23454–23471.
- Asghari, M., & Al-e-hashem, S. M. J. M. (2021). Green vehicle routing problem: A state-of-the-art review. *International Journal of Production Economics*, 231, 107899.
- Ashtari, A., Bibeau, E., & Shahidinejad, S. (2014). Using large driving record samples and a stochastic approach for real-world driving cycle construction: Winnipeg driving cycle. *Transportation Science*, 48(2), 170–183.
- AT&T. (2025). *HARMAN Spark*. <https://www.att.com/buy/connected-devices-and-more/harman-spark.html>
- Ayman, A., Sivagnanam, A., Wilbur, M., Pugliese, P., Dubey, A., & Laszka, A. (2022). Data-Driven Prediction and Optimization of Energy Use for Transit Fleets of Electric and ICE Vehicles. *ACM Transactions on Internet Technology*, 22(1). <https://doi.org/10.1145/3433992>
- Baek, D., Chen, Y., Bocca, A., Bottaccioli, L., Cataldo, S. Di, Gatteschi, V., Pagliari, D. J., Patti, E., Urgese, G., Chang, N., Macii, A., Macii, E., Montuschi, P., & Poncino, M. (2019). Battery-Aware operation range estimation for terrestrial and aerial electric vehicles. *IEEE Transactions on Vehicular Technology*, 68(6), 5471–5482. <https://doi.org/10.1109/TVT.2019.2910452>
- Bajrami, A., & Palpacelli, M. C. (2023). A Proposal for a Simplified Systematic Procedure for the Selection of Electric Motors for Land Vehicles with an Emphasis on Fuel Economy. *Machines*, 11(4). <https://doi.org/10.3390/machines11040420>
- Bar, T., Nienhüser, D., Kohlhaas, R., & Zöllner, J. M. (2011). Probabilistic driving style determination by means of a situation based analysis of the vehicle data. *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, 1698–1703. <https://doi.org/10.1109/ITSC.2011.6082924>

- Barlow, T. J., Latham, S., McCrae, I. S., & Boulter, P. G. (2009). A reference book of driving cycles for use in the measurement of road vehicle emissions.
- Basso, R., Kulcsár, B., & Sanchez-Diaz, I. (2021). Electric vehicle routing problem with machine learning for energy prediction. *Transportation Research Part B: Methodological*, 145, 24–55. <https://doi.org/10.1016/j.trb.2020.12.007>
- Bel, G., Bolancé, C., Guillén, M., & Rosell, J. (2015). The environmental effects of changing speed limits: A quantile regression approach. *Transportation Research Part D: Transport and Environment*, 36, 76–85. <https://doi.org/10.1016/j.trd.2015.02.003>
- Ben-Chaim, M., Shmerling, E., & Kuperman, A. (2013). Analytic modeling of vehicle fuel consumption. *Energies*, 6(1), 117–127. <https://doi.org/10.3390/en6010117>
- Berjoza, Dainis, Vilnis Pirs, and Inara Jurgena. "Research into the regenerative braking of an electric car in urban driving." *World Electric Vehicle Journal* 13.11 (2022): 202.
- Berzi, L., Delogu, M., & Pierini, M. (2016). Development of driving cycles for electric vehicles in the context of the city of Florence. *Transportation Research Part D: Transport and Environment*, 47, 299–322. <https://doi.org/10.1016/j.trd.2016.05.010>
- BimmerLink. (2025). *BIMMERLINK*. <https://www.bimmerlink.app/>
- Blades, L. A. W., Matthews, T., McGrath, T. E., Early, J., Cunningham, G., & Harris, A. (2024). Predicting energy consumption of zero emission buses using route feature selection methods. *Transportation Research Part D: Transport and Environment*, 130(September 2023), 104158. <https://doi.org/10.1016/j.trd.2024.104158>
- Bishop, J. D. K., Axon, C. J., & McCulloch, M. D. (2012). A robust, data-driven methodology for real-world driving cycle development. *Transportation Research Part D: Transport and Environment*, 17(5), 389–397. <https://doi.org/10.1016/j.trd.2012.03.003>
- Boroogeni, K. G., Amini, M. H., Bahrami, S., Iyengar, S. S., Sarwat, A. I., & Karabasoglu, O. (2017). A novel multi-time-scale modeling for electric power demand forecasting: From short-term to medium-term horizon. *Electric Power Systems Research*, 142, 58–73. <https://doi.org/10.1016/j.epsr.2016.08.031>
- Botshekan, M., Tootkaboni, M. P., & Louhghalam, A. (2019). Global Sensitivity of Roughness-Induced Fuel Consumption to Road Surface Parameters and Car Dynamic Characteristics. *Transportation Research Record*, 2673(2), 183–193. <https://doi.org/10.1177/0361198118821318>
- Boulter, P. G., & McCrae, I. S. (2007). ARTEMIS: Assessment and reliability of transport emission models and inventory systems-Final Report. TRL Published Project Report.
- Boyles, S. D., Lownes, N. E., & Unnikrishnan, A. (2020). Transportation network analysis. *Volume I: Static and Dynamic Traffic Assignment*.

- Bozorgi, A. M., Farasat, M., & Mahmoud, A. (2017). A Time and Energy Efficient Routing Algorithm for Electric Vehicles Based on Historical Driving Data. *IEEE Transactions on Intelligent Vehicles*, 2(4), 308–320. <https://doi.org/10.1109/TIV.2017.2771233>
- Braband, M., Scherer, M., & Voos, H. (2022). Global Sensitivity Analysis of Economic Model Predictive Longitudinal Motion Control of a Battery Electric Vehicle. *Electronics (Switzerland)*, 11(10), 1–33. <https://doi.org/10.3390/electronics11101574>
- Brady, J., & Mahony, M. O. (2016). Development of a driving cycle to evaluate the energy economy of electric vehicles in urban areas. *Applied Energy*, 177, 165–178. <https://doi.org/10.1016/j.apenergy.2016.05.094>
- Bühler, F., Cocron, P., Neumann, I., Franke, T., & Krems, J. F. (2014). Is EV experience related to EV acceptance? Results from a German field study. *Transportation Research Part F: traffic psychology and behaviour*, 25, 34–49.
- Burkardt, J. (2014). The truncated normal distribution. Department of Scientific Computing Website, Florida State University, 1(35), 58.
- Car Scanner. (2025). *Car Scanner ELM OBD2*. <https://www.carscanner.info/>
- Carmona, J., García, F., Martín, D., de la Escalera, A., & Armingol, J. M. (2015). Data fusion for driver behaviour analysis. *Sensors (Switzerland)*, 15(10), 25968–25991. <https://doi.org/10.3390/s151025968>
- Caulfield, B., Brazil, W., Ni Fitzgerald, K., & Morton, C. (2014). Measuring the success of reducing emissions using an on-board eco-driving feedback tool. *Transportation Research Part D: Transport and Environment*, 32, 253–262. <https://doi.org/10.1016/j.trd.2014.08.011>
- Central Statistics Office. (2021). Ireland's UN SDG's - Goal 9 Industry Innovation and Infrastructure 2021. <https://www.cso.ie/en/releasesandpublications/ep/p-sdg9/irelandsunsdgs-goal9industryinnovationandinfrastructure2021/infrastructure/>
- Central Statistics Office. (2024). Road Traffic Volumes. <https://data.cso.ie/>
- Central Statistics Office. (2025). *Vehicles Licensed for the first time July 2025*. <https://www.cso.ie/en/releasesandpublications/ep/p-vlftm/vehicleslicensedforthefirsttimejuly2025/>
- Cerema. (2020). *Lowering of the maximum authorised speed to 80 km/h - Final assessment report - July 2020*. <https://www.cerema.fr/en/actualites/lowering-maximum-authorised-speed-80-kmh-final-assessment>
- Cetin, V. R., Yilmaz, H. H., & Erkan, V. (2018). The impact of increasing speed limit in Turkey: The case of Ankara-Sivrihisar road section. *Case Studies on Transport Policy*, 6(1), 72–80. <https://doi.org/10.1016/j.cstp.2017.11.004>

- Chen, R., Zhang, T., & Levin, M. W. (2020A). Effects of Variable Speed Limit on Energy Consumption with Autonomous Vehicles on Urban Roads Using Modified Cell-Transmission Model. *Journal of Transportation Engineering, Part A: Systems*, 146(7), 1–12. <https://doi.org/10.1061/jtepbs.0000379>
- Chen, S., Sun, W., Li, Y., & Shi, L. (2019). On the relationship between energy consumption and driving behavior of electric vehicles based on statistical features. *Chinese Control Conference, CCC, 2019-July*, 3782–3787. <https://doi.org/10.23919/ChiCC.2019.8866414>
- Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 13-17-August-2016*, 785–794. <https://doi.org/10.1145/2939672.2939785>
- Chen, X., & Yang, J. (2022). Potential benefit of electric vehicles in counteracting future urban warming: A case study of Hong Kong. *Sustainable Cities and Society*, 87, 104200.
- Chen, Y., Zhu, L., Gonder, J., Young, S., & Walkowicz, K. (2017). Data-driven fuel consumption estimation: A multivariate adaptive regression spline approach. *Transportation Research Part C: Emerging Technologies*, 83, 134–145. <https://doi.org/10.1016/j.trc.2017.08.003>
- Chen, Z., Yang, C. and Fang, S. (2020B) ‘A Convolutional Neural Network-Based Driving Cycle Prediction Method for Plug-in Hybrid Electric Vehicles with Bus Route’, *IEEE Access*, 8, pp. 3255–3264. Available at: <https://doi.org/10.1109/ACCESS.2019.2960771>.
- Chen, Z., Zhang, Q., Lu, J., & Bi, J. (2019). Optimization-based method to develop practical driving cycle for application in electric vehicle power management: A case study in Shenyang, China. *Energy*, 186, 115766. <https://doi.org/10.1016/j.energy.2019.07.096>
- Cheng, Z. J., Jeng, L. W., & Li, K. (2018). Behavioral classification of drivers for driving efficiency related ADAS using artificial neural network. *Proceedings of the 2018 IEEE International Conference on Advanced Manufacturing, ICAM 2018*, 173–176. <https://doi.org/10.1109/AMCON.2018.8614836>
- Chollet, F. Keras. (2015). <https://keras.io>
- Chu, H., Zhuang, H., Wang, W., Na, X., Guo, L., Zhang, J., Gao, B., & Chen, H. (2023). A Review of Driving Style Recognition Methods From Short-Term and Long-Term Perspectives. *IEEE Transactions on Intelligent Vehicles*, 8(11), 4599–4612. <https://doi.org/10.1109/TIV.2023.3279425>
- Citizens Information. (2024). Speed limits in Ireland. <https://www.citizensinformation.ie/en/travel-and-recreation/roads-and-safety/road-traffic-speed-limits-in-ireland/>
- Citizens Information. (2025). *Electric Vehicles*. <https://www.citizensinformation.ie/en/travel-and-recreation/motoring/buying-or-selling-a-vehicle/electric-vehicles/>

- Citizens Information. (2025). Speed Limits in Ireland.
<https://www.citizensinformation.ie/en/travel-and-recreation/roads-and-safety/road-traffic-speed-limits-in-ireland/>
- Cleland, C. L., Baker, G., Turner, K., Jepson, R., Kee, F., Milton, K., Kelly, M. P., Nightingale, G., & Hunter, R. F. (2021). A qualitative exploration of the mechanisms, pathways and public health outcomes of a city centre 20mph speed limit intervention: The case of Belfast, United Kingdom. *Health and Place*, 70(February), 102627. <https://doi.org/10.1016/j.healthplace.2021.102627>
- Cleland, C. L., McComb, K., Kee, F., Jepson, R., Kelly, M. P., Milton, K., Nightingale, G., Kelly, P., Baker, G., Craig, N., Williams, A. J., & Hunter, R. F. (2020). Effects of 20 mph interventions on a range of public health outcomes: A meta-narrative evidence synthesis. *Journal of Transport and Health*, 17(September 2019).
<https://doi.org/10.1016/j.jth.2019.100633>
- Corti, A., Manzoni, V., & Savaresi, S. M. (2012). Vehicle's energy estimation using low frequency speed signal. *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, 626–631. <https://doi.org/10.1109/ITSC.2012.6338758>
- Coutts, S. R., & Yokomizo, H. (2014). Meta-models as a straightforward approach to the sensitivity analysis of complex models. *Population Ecology*, 56(1), 7-19.
- Cukier, R. I., Levine, H. B., & Shuler, K. E. (1978). Nonlinear sensitivity analysis of multiparameter model systems. *Journal of computational physics*, 26(1), 1-42.
- CSS Electronics. (2025). *CANedge1: 2x CAN Bus Data Logger (SD + RTC)*.
<https://www.csselectronics.com/products/can-logger-sd-canedge1>
- Cura, A., Kucuk, H., Ergen, E., & Oksuzoglu, I. B. (2021). Driver Profiling Using Long Short Term Memory (LSTM) and Convolutional Neural Network (CNN) Methods. *IEEE Transactions on Intelligent Transportation Systems*, 22(10), 6572–6582.
<https://doi.org/10.1109/TITS.2020.2995722>
- Czégé, L., Vámosi, A., & Kocsis, I. (2020). Review on Construction Procedures of Driving Cycles. *International Journal of Engineering and Management Sciences*, 5(2), 266–285. <https://doi.org/10.21791/ijems.2020.2.31>.
- Das, K., & Kumar, R. (2023). A Comprehensive Review of Categorization and Perspectives on State-of-Charge Estimation Using Deep Learning Methods for Electric Transportation. *Wireless Personal Communications*, 133(3), 1599–1618.
<https://doi.org/10.1007/s11277-023-10830-5>
- de Andrade, G. M. S., De Araújo, F. W. C., De Novaes Santos, M. P. M., & Magnani, F. S. (2020). Standardized comparison of 40 local driving cycles: Energy and kinematics. *Energies*, 13(20). <https://doi.org/10.3390/en13205434>
- de Andrade, G. M. S., Santos, M. P. M. de N., Araújo, F. W. C. de, & Magnani, F. S. (2023). Evaluation of the Minimum Number of Local Driving Cycles Required to Represent the Traffic of Distinct Cities: A Case Study of Two Brazilian Metropolises. *Transportation Research Record*.
<https://doi.org/10.1177/03611981231186977>

- De Cauwer, C., Verbeke, W., Coosemans, T., Faid, S., & Van Mierlo, J. (2017). A data-driven method for energy consumption prediction and energy-efficient routing of electric vehicles in real-world conditions. *Energies*, 10(5). <https://doi.org/10.3390/en10050608>
- De Cauwer, C., Verbeke, W., Van Mierlo, J., & Coosemans, T. (2020). A Model for Range Estimation and Energy-Efficient Routing of Electric Vehicles in Real-World Conditions. *IEEE Transactions on Intelligent Transportation Systems*, 21(7), 2787–2800. <https://doi.org/10.1109/TITS.2019.2918019>
- Department of Transport. (2023A). Speed Limit Review. <https://assets.gov.ie/static/documents/speed-limit-review-main-report.pdf>
- Department of Transport. (2023B). Modelling Assessment Report. <https://assets.gov.ie/static/documents/speed-limit-review-modelling-assessment-report.pdf>
- Desreux, A., Bouscayrol, A., Castex, E., Trigui, R., Hittinger, E., & Sirbu, G. M. (2020). Annual variation in energy consumption of an electric vehicle used for commuting. *Energies*, 13(18). <https://doi.org/10.3390/en13184639>
- Diaz Alvarez, A., Serradilla Garcia, F., Naranjo, J. E., Anaya, J. J., & Jimenez, F. (2014). Modeling the driving behavior of electric vehicles using smartphones and neural networks. *IEEE Intelligent Transportation Systems Magazine*, 6(3), 44–53.
- Dibbelt, J., Strasser, B., & Wagner, D. (2016). Customizable contraction hierarchies. *Journal of Experimental Algorithmics (JEA)*, 21, 1–49.
- Donkers, A., Yang, D., & Viktorović, M. (2020). Influence of driving style, infrastructure, weather and traffic on electric vehicle performance. *Transportation Research Part D: Transport and Environment*, 88(October). <https://doi.org/10.1016/j.trd.2020.102569>
- Duran, A., & Earleywine, M. (2012). GPS data filtration method for drive cycle analysis applications. *SAE Technical Papers*. <https://doi.org/10.4271/2012-01-0743>
- Edwards, A. J. (2022). Aggressive-dynamics metrics for drive-cycle characterization. *Transportation research interdisciplinary perspectives*, 14, 100592.
- Ekici, Y. E., Akdag, O., Aydin, A. A., & Karadag, T. (2023). A Novel Energy Consumption Prediction Model of Electric Buses Using Real-Time Big Data From Route, Environment, and Vehicle Parameters. *IEEE Access*, 11(August), 104305–104322. <https://doi.org/10.1109/ACCESS.2023.3316362>
- Electric Ireland. (2025). *Find Your Plan*. <https://www.electricireland.ie/switch/new-customer/price-plans?priceType=E>
- Ellison, A. B., Greaves, S. P., & Bliemer, M. C. J. (2015). Driver behaviour profiles for road safety analysis. *Accident Analysis and Prevention*, 76, 118–132. <https://doi.org/10.1016/j.aap.2015.01.009>
- ESB. (2025). *Price Plans*. <https://esb.ie/what-we-do/ecars/price-plans>

- EU Urban Mobility Observatory. (2020). *New Spanish law to require 30 km/h speed limit in urban areas*. https://urban-mobility-observatory.transport.ec.europa.eu/news-events/news/new-spanish-law-require-30-kmh-speed-limit-urban-areas-2020-11-23_en?prefLang=et
- EV Watchdog. (2025). *EV Watchdog*. <https://evwatchdog.net/>
- EVNotify. (2025). *Remote monitoring and notifications for your electric vehicle. For free*. <https://evnotify.de/>
- Fafoutellis, P., Mantouka, E. G., & Vlahogianni, E. I. (2021). Eco-driving and its impacts on fuel efficiency: An overview of technologies and data-driven methods. *Sustainability (Switzerland)*, 13(1), 1–17. <https://doi.org/10.3390/su13010226>
- Faria, M. V., Baptista, P. C., & Farias, T. L. (2017). Identifying driving behavior patterns and their impacts on fuel use. *Transportation Research Procedia*, 27, 953–960. <https://doi.org/10.1016/j.trpro.2017.12.038>
- Faria, M. V., Baptista, P. C., Farias, T. L., & Pereira, J. M. S. (2020). Assessing the impacts of driving environment on driving behavior patterns. *Transportation*, 47(3), 1311–1337. <https://doi.org/10.1007/s11116-018-9965-5>
- Filipi, N., Karlínová, B., & Krčál, O. (2022). The disutility of driving below the speed limit on highways. *Review of Economic Perspectives*, 22(4), 267–277. <https://doi.org/10.2478/revecp-2022-0012>
- Fiori, C., Arcidiacono, V., Fontaras, G., Makridis, M., Mattas, K., Marzano, V., Thiel, C., & Ciuffo, B. (2019). The effect of electrified mobility on the relationship between traffic conditions and energy consumption. *Transportation Research Part D: Transport and Environment*, 67(December 2018), 275–290. <https://doi.org/10.1016/j.trd.2018.11.018>
- Flötteröd, G., & Rohde, J. (2011). Operational macroscopic modeling of complex urban road intersections. *Transportation Research Part B: Methodological*, 45(6), 903–922.
- Felipe, J., Amarillo, J. C., Naranjo, J. E., Serradilla, F., & Diaz, A. (2015). Energy Consumption Estimation in Electric Vehicles Considering Driving Style. *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC, 2015-October*, 101–106. <https://doi.org/10.1109/ITSC.2015.25>
- Foiadelli, F., Longo, M., & Miraftabzadeh, S. (2018). Energy Consumption Prediction of Electric Vehicles Based on Big Data Approach. *Proceedings - 2018 IEEE International Conference on Environment and Electrical Engineering and 2018 IEEE Industrial and Commercial Power Systems Europe, IEEEIC/I and CPS Europe 2018*, 1–6. <https://doi.org/10.1109/EEEIC.2018.8494573>
- Folgerø, I. K., Harding, T., & Westby, B. S. (2020). Going fast or going green? Evidence from environmental speed limits in Norway. *Transportation Research Part D: Transport and Environment*, 82(March), 102261. <https://doi.org/10.1016/j.trd.2020.102261>

- Fondzenyuy, S. K., Turner, B. M., Burlacu, A. F., Jurewicz, C., Usami, D. S., Feudjio, S. L. T., & Persia, L. (2024). The Impact of Speed Limit Change on Emissions: A Systematic Review of Literature. *Sustainability (Switzerland)*, 16(17), 1–17. <https://doi.org/10.3390/su16177712>
- Forbes, C., Evans, M., Hastings, N., & Peacock, B. (2011). *Statistical distributions*. John Wiley & Sons.
- Franke, T., Cocron, P., Bühler, F., Neumann, I., & Krems, J. F. (2012). Adapting to the range of an electric vehicle—the relation of experience to subjectively available mobility resources. In *Proceedings of the European conference on human centred design for intelligent transport systems, Valencia, Spain* (pp. 95-103).
- Friedman, J., Tibshirani, R., & Hastie, T. (2000). Additive logistic regression: a statistical view of boosting (With discussion and a rejoinder by the authors). *The Annals of Statistics*, 28(2), 337–407. <https://doi.org/10.1214/aos/1016120463>
- Friedman, L. S., Hedeker, D., & Richter, E. D. (2009). Long-term effects of repealing the national maximum speed limit in the United States. *American journal of public health*, 99(9), 1626-1631.
- Fukushima, A., Yano, T., Imahara, S., Aisu, H., Shimokawa, Y., & Shibata, Y. (2018). Prediction of energy consumption for new electric vehicle models by machine learning. *IET Intelligent Transport Systems*, 12(9), 1174–1180. <https://doi.org/10.1049/iet-its.2018.5169>
- Galgamuwa, U., Perera, L., & Bandara, S. (2015). Developing a General Methodology for Driving Cycle Construction: Comparison of Various Established Driving Cycles in the World to Propose a General Approach. *Journal of Transportation Technologies*, 05(04), 191–203. <https://doi.org/10.4236/jtts.2015.54018>
- Galvin, R. (2017). Energy consumption effects of speed and acceleration in electric vehicles: Laboratory case studies and implications for drivers and policymakers. *Transportation Research Part D: Transport and Environment*, 53, 234–248. <https://doi.org/10.1016/j.trd.2017.04.020>
- Gawlikowski, J., Tassi, C. R. N., Ali, M., Lee, J., Humt, M., Feng, J., Kruspe, A., Triebel, R., Jung, P., Roscher, R., Shahzad, M., Yang, W., Bamler, R., & Zhu, X. X. (2023). A survey of uncertainty in deep neural networks. In *Artificial Intelligence Review* (Vol. 56, Issue s1). Springer Netherlands. <https://doi.org/10.1007/s10462-023-10562-9>
- Gebisa, A., Gebresenbet, G., Gopal, R., & Nallamotheu, R. B. (2021). Driving Cycles for Estimating Vehicle Emission Levels and Energy Consumption. *Future Transportation*, 1(3), 615–638. <https://doi.org/10.3390/futuretransp1030033>
- Genikomsakis, K. N., & Mitrentsis, G. (2017). A computationally efficient simulation model for estimating energy consumption of electric vehicles in the context of route planning applications. *Transportation Research Part D: Transport and Environment*, 50, 98–118. <https://doi.org/10.1016/j.trd.2016.10.014>

- Gong, H., Zou, Y., Yang, Q., Fan, J., & Sun, F. (2018). Generation of a driving cycle for battery electric vehicles : A case study of Beijing. *Energy*, 150, 901–912. <https://doi.org/10.1016/j.energy.2018.02.092>
- Google Play. (2025). *Leaf Spy Pro*. https://play.google.com/store/apps/details?id=com.Turbo3.Leaf_Spy_Pro&hl
- Government of Ireland. (2019). Design Manual for Urban Roads and Streets. <https://www.dmurs.ie/copy-of-what-is-dmurs>
- Government of Ireland. (2025). Climate Action Plan 2025. <https://www.gov.ie/en/department-of-climate-energy-and-the-environment/publications/climate-action-plan-2025/>
- Guagalango-Gomez, P. J., Balbuena-Campuzano, C. A., Chunab-Rodriguez, M. A., Rodriguez-Arce, J., Vinolas, J., & Santana-Diaz, A. (2023). Raspberry Pi-Based Low-Cost System for Electrical Parameters Gathering and Analysis for Performance Evaluation of Electric Vehicles. *IEEE Transactions on Instrumentation and Measurement*, 72. <https://doi.org/10.1109/TIM.2023.3244815>
- Guimarães, P. P., Arce, J. E., Lopes, E. da S., Fiedler, N. C., Robert, R. C. G., & Seixas, F. (2019). Analysis of fuel consumption sensitivity in forestry road transport. *Floresta*, 49(2), 155–162. <https://doi.org/10.5380/rf.v49i2.49137>
- Günther, R., Wenzel, T., Wegner, M., & Rettig, R. (2017). Big data driven dynamic driving cycle development for busses in urban public transportation. *Transportation Research Part D: Transport and Environment*, 51, 276–289. <https://doi.org/10.1016/j.trd.2017.01.009>
- Guo, F., & Zhang, F. (2016). A study of driving cycle for electric cars on Beijing urban and suburban roads. 2016 IEEE International Conference on Power and Renewable Energy, ICPRE 2016, 319–322. <https://doi.org/10.1109/ICPRE.2016.7871224>
- Guo, J., Jiang, Y., Yu, Y., & Liu, W. (2020). A novel energy consumption prediction model with combination of road information and driving style of BEVs. *Sustainable Energy Technologies and Assessments*, 42(May), 100826. <https://doi.org/10.1016/j.seta.2020.100826>
- Gurusamy, A., Ashok, B., & Mason, B. (2023). Prediction of Electric Vehicle Driving Range and Performance Characteristics: A Review on Analytical Modeling Strategies With Its Influential Factors and Improvisation Techniques. *IEEE Access*, 11(November), 131521–131548. <https://doi.org/10.1109/ACCESS.2023.3334620>
- Han, D. S., Choi, N. W., Cho, S. L., Yang, J. S., Kim, K. S., Yoo, W. S., & Jeon, C. H. (2012). Characterization of driving patterns and development of a driving cycle in a military area. *Transportation Research Part D: Transport and Environment*, 17(7), 519–524. <https://doi.org/10.1016/j.trd.2012.06.004>
- Hao, J. Ruan, S. Wang, W. (2023). Model Predictive Control Based Energy Management Strategy of Series Hybrid Electric Vehicles Considering Driving Pattern Recognition. *Electronics*, 32(1418).

- Hariharan, C., Gunadevan, D., Arun Prakash, S., Latha, K., Antony Aroul Raj, V., & Velraj, R. (2022). Simulation of battery energy consumption in an electric car with traction and HVAC model for a given source and destination for reducing the range anxiety of the driver. *Energy*, 249, 123657. <https://doi.org/10.1016/j.energy.2022.123657>
- Herman, J. and Usher, W. (2017) SALib: An open-source Python library for sensitivity analysis. *Journal of Open Source Software*, 2(9). doi:10.21105/joss.00097
- Herman, J., Usher, W. et al. (2025). *Advanced Examples*. SALib. https://salib.readthedocs.io/en/latest/user_guide/advanced.html
- Ho, S., Wong, Y., & Chang, V. W. (2014). Developing Singapore Driving Cycle for passenger cars to estimate fuel consumption and vehicular emissions. *Atmospheric Environment*, 97, 353–362. <https://doi.org/10.1016/j.atmosenv.2014.08.042>
- Hong, S., Hwang, H., Kim, D., Cui, S., & Joe, I. (2021). Real driving cycle-based state of charge prediction for ev batteries using deep learning methods. *Applied Sciences (Switzerland)*, 11(23). <https://doi.org/10.3390/app112311285>
- Hosseini, M. H., Kheyrabadi, S. A., & Zolfaghari, A. (2015). Determining optimal speed limits in traffic networks. *IATSS Research*, 39(1), 36–41. <https://doi.org/10.1016/j.iatssr.2014.08.003>
- Hu, L., Tian, Q., Zou, C., Huang, J., Ye, Y., & Wu, X. (2022). A study on energy distribution strategy of electric vehicle hybrid energy storage system considering driving style based on real urban driving data. *Renewable and Sustainable Energy Reviews*, 162(January), 112416. <https://doi.org/10.1016/j.rser.2022.112416>
- Hu, L., Zhong, Y., Hao, W., Moghimi, B., Huang, J., Zhang, X., & Du, R. (2018). Optimal route algorithm considering traffic light and energy consumption. *Ieee Access*, 6, 59695-59704.
- Huang, H., Huang, X., Ding, W., Yang, M., Fan, D., & Pang, J. (2022). Uncertainty optimization of pure electric vehicle interior tire/road noise comfort based on data-driven. *Mechanical Systems and Signal Processing*, 165(August 2021), 108300. <https://doi.org/10.1016/j.ymssp.2021.108300>
- Huang, R., Ni, J., Cheng, Z., Wang, Q., Shi, X., & Yao, X. (2021). Assessing the effects of ethanol additive and driving behaviors on fuel economy, particle number, and gaseous emissions of a GDI vehicle under real driving conditions. *Fuel*, 306, 121642.
- Huang, Y., Zhu, L., Sun, R., Yi, J., Liu, L., & Luan, T. H. (2020). Save or Waste: Real Data Based Energy-Efficient Driving. *IEEE Access*, 8, 133936–133950. <https://doi.org/10.1109/ACCESS.2020.3007508>
- Huertas, J. I., Díaz, J., Cordero, D., & Cedillo, K. (2018). A new methodology to determine typical driving cycles for the design of vehicles power trains. *International Journal on Interactive Design and Manufacturing*, 12(1), 319–326. <https://doi.org/10.1007/s12008-017-0379-y>
- Huertas, I., Quirama, L. F., & Giraldo, M. (2019). Comparison of Three Methods for Constructing Real Driving Cycles. <https://doi.org/10.3390/en12040665>

- Hyundai. (2024). *Ioniq 6*. Hyundai. www.hyundai.ie/Ioniq6/
- Iclodean, C., Jurchis, B., Macavei, C., & Volosciuc, E. (2025). Optimizing the Energy Efficiency of Electric Vehicles in Urban and Metropolitan Environments According to Various Driving Cycles and Behavioral Conditions. *Electronics*, 14(2224).
- IEA. (2025). Global EV Outlook 2025. <https://iea.blob.core.windows.net/assets/7ea38b60-3033-42a6-9589-71134f4229f4/GlobalEVOutlook2025.pdf>
- Iora, P., & Tribioli, L. (2019). Effect of ambient temperature on electric vehicles' energy consumption and range: Model definition and sensitivity analysis based on Nissan Leaf data. *World Electric Vehicle Journal*, 10(1), 1–15.
<https://doi.org/10.3390/wevj10010002>
- Iwanaga, T., Usher, W., & Herman, J. (2022). Toward SALib 2.0: Advancing the accessibility and interpretability of global sensitivity analyses. *Socio-Environmental Systems Modelling*, 4, 18155. doi:10.18174/sesmo.18155
- Jacko, T., Bartsch, J., von Diecken, C., & Ueberschär, O. (2024). Validity of Current Smartwatches for Triathlon Training: How Accurate Are Heart Rate, Distance, and Swimming Readings?. *Sensors*, 24(14), 4675.
- Jadali, A. M., Ioannou, A., Salonitis, K., & Kolios, A. (2021). Decommissioning vs. repowering of offshore wind farms—a techno-economic assessment. *International Journal of Advanced Manufacturing Technology*, 112(9–10), 2519–2532.
<https://doi.org/10.1007/s00170-020-06349-9>
- Janpoom, K., Suttakul, P., Achariyaviriya, W., Fongsamootr, T., Katongtung, T., & Tippayawong, N. (2023). Investigating the influential factors in real-world energy consumption of battery electric vehicles. *Energy Reports*, 9(S12), 316–320.
<https://doi.org/10.1016/j.egyr.2023.10.012>
- Jaworski, A., Kuszewski, H., Lew, K., Wojewoda, P., Balawender, K., Woś, P., Longwic, R., & Boichenko, S. (2023). Assessment of the Effect of Road Load on Energy Consumption and Exhaust Emissions of a Hybrid Vehicle in an Urban Road Driving Cycle—Comparison of Road and Chassis Dynamometer Tests. *Energies*, 16(15). <https://doi.org/10.3390/en16155723>
- Jensen, A. F., Cherchi, E., & Mabit, S. L. (2013). On the stability of preferences and attitudes before and after experiencing an electric vehicle. *Transportation Research Part D: Transport and Environment*, 25, 24–32.
- Jeon, H., Kuo, Y. H., Aupetit, M., Ma, K. L., & Seo, J. (2023). Classes are not clusters: Improving label-based evaluation of dimensionality reduction. *IEEE Transactions on Visualization and Computer Graphics*, 30(1), 781–791.
- Ji, J., Bie, Y., Zeng, Z., & Wang, L. (2022). Trip energy consumption estimation for electric buses. *Communications in Transportation Research*, 2(April), 100069.
<https://doi.org/10.1016/j.commtr.2022.100069>

- Jiang, J., Yu, Y., Min, H., Cao, Q., Sun, W., Zhang, Z., & Luo, C. (2023). Trip-level energy consumption prediction model for electric bus combining Markov-based speed profile generation and Gaussian processing regression. *Energy*, 263(PD), 125866. <https://doi.org/10.1016/j.energy.2022.125866>
- Jiménez, D., Hernández, S., Fraile-Ardanuy, J., Serrano, J., Fernández, R., & Alvarez, F. (2018). Modelling the effect of driving events on electrical vehicle energy consumption using inertial sensors in smartphones. *Energies*, 11(2). <https://doi.org/10.3390/en11020412>
- Jing, Z., Wang, G., Zhang, S., & Qiu, C. (2017). Building Tianjin driving cycle based on linear discriminant analysis. *Transportation Research Part D*, 53, 78–87. <https://doi.org/10.1016/j.trd.2017.04.005>
- Jing, Z., Wang, T., Zhang, S., & Wang, G. (2022). Development Method for the Driving Cycle of Electric Vehicles. *Energies*, 15(22). <https://doi.org/10.3390/en15228715>
- Jondrow, J., Bowes, M., & Levy, R. (1983). The optimal speed limit. *Economic Inquiry*, 21(3), 325.
- Joumard, R., André, M., Vidon, R., Tassel, P., & Pruvost, C. (2000). Influence of driving cycles on unit emissions from passenger cars. *Atmospheric Environment*, 34(27), 4621–4628. [https://doi.org/10.1016/S1352-2310\(00\)00118-7](https://doi.org/10.1016/S1352-2310(00)00118-7)
- Jung, M. F., Sirkin, D., Gür, T. M., & Steinert, M. (2015). Displayed uncertainty improves driving experience and behavior: The case of range anxiety in an electric car. *Conference on Human Factors in Computing Systems - Proceedings, 2015-April*(April), 2201–2210. <https://doi.org/10.1145/2702123.2702479>
- Kamble, S. H., Mathew, T. V., & Sharma, G. K. (2009). Development of real-world driving cycle : Case study of Pune , India. *Transportation Research Part D*, 14(2), 132–140. <https://doi.org/10.1016/j.trd.2008.11.008>
- Kambly, K. R., & Bradley, T. H. (2014). Estimating the HVAC energy consumption of plug-in electric vehicles. *Journal of Power Sources*, 259, 117–124. <https://doi.org/10.1016/j.jpowsour.2014.02.033>
- Karamangil, M.; Tekin, M. (2022). Comparison of fuel consumption and recoverable energy according to NEDC and WLTP cycles of a vehicle. *CT&F - Ciencia, Tecnología y Futuro*, 12(December), 31–38.
- Karbowski, D., Pagerit, S., & Calkins, A. (2012). Energy consumption prediction of a vehicle along a user-specified real-world trip. *World Electric Vehicle Journal*, 5(4), 1109–1120. <https://doi.org/10.3390/wevj5041109>
- Kawgan-Kagan, I., & Daubitz, S. (2014). BEVs beyond Range Anxiety - Cognitive Interviews to Elicit the Perception of Mode Choice. *Transportation Research Procedia*, 4, 387–390. <https://doi.org/10.1016/j.trpro.2014.11.029>
- Khanra, M., Chakraborty, D., & Nandi, A. K. (2018). Improvement of regenerative braking energy of fully battery electric vehicle through optimal driving. *Journal of Asian Electric Vehicles*, 16(1), 1789-1798.

- Kia. (2024). *The EV6 Get Inspired*. Kia. www.kia.com/ie/new-cars/ev6/discover/
- Kim, S., Lee, J., & Lee, C. (2017). Does driving range of electric vehicles influence electric vehicle adoption? *Sustainability (Switzerland)*, 9(10).
<https://doi.org/10.3390/su9101783>
- Kim, T. K. (2017). Understanding one-way ANOVA using conceptual figures. *Korean journal of anesthesiology*, 70(1), 22.
- Kim, T. Y., & Cho, S. B. (2019). Predicting residential energy consumption using CNN-LSTM neural networks. *Energy*, 182, 72–81.
<https://doi.org/10.1016/j.energy.2019.05.230>
- Kingma, D. P., & Ba, J. L. (2015). Adam: A method for stochastic optimization. *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, 1–15.
- Kioutsioukis, I., Tarantola, S., Saltelli, A., & Gatelli, D. (2004). Uncertainty and global sensitivity analysis of road transport emission estimates. *Atmospheric Environment*, 38(38), 6609–6620. <https://doi.org/10.1016/j.atmosenv.2004.08.006>
- Kockelman, K., Bottom, J., Kweon, Y. J., Ma, J., & Wang, X. (2006). *Safety impacts and other implications of raised speed limits on high-speed roads* (Vol. 90). Washington, DC, USA: Transportation Research Board.
- Kodinariya, Trupti M., and Prashant R. Makwana. (2013). Review on determining number of Cluster in K-Means Clustering. *International Journal* 1.6: 90-95.
- Komorska, I., Puchalski, A., Niewczas, A., Ślęzak, M., & Szczepański, T. (2021). Adaptive driving cycles of evs for reducing energy consumption. *Energies*, 14(9).
<https://doi.org/10.3390/en14092592>
- Kolarova, V., Kuhnimhof, T., & Trommer, S. (2017). Assessment of real-world vehicle data from electric vehicles—potentials and challenges. *11th International Conference on Transport Survey Methods ISCTSC, September 24- 29, 2017, Québec, Canada*
- Komorska, I., Puchalski, A., Niewczas, A., Ślęzak, M., & Szczepański, T. (2021). Adaptive driving cycles of evs for reducing energy consumption. *Energies*, 14(9).
<https://doi.org/10.3390/en14092592>
- Krasopoulos, C. T., Beniakar, M. E., & Kladas, A. G. (2017). Speed and Torque Limit Profile Optimization of Electric Vehicle Including Limited Overload. *IEEE Transactions on Industry Applications*, 53(4), 3907–3916.
<https://doi.org/10.1109/TIA.2017.2680405>
- Krivda, V., Petru, J., Macha, D., & Novak, J. (2021). Use of microsimulation traffic models as means for ensuring public transport sustainability and accessibility. *Sustainability*, 13(5), 2709.

- Kumar, D., Singh, A., Kumar, P., Jha, R. K., Sahoo, S. K., & Jha, V. (2020). Sobol sensitivity analysis for risk assessment of uranium in groundwater. *Environmental Geochemistry and health*, 42(6), 1789-1801.
- Lai, S., Zhang, W., & Liu, H. (2023). A Preference-aware Meta-optimization Framework for Personalized Vehicle Energy Consumption Estimation. In *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (Vol. 1, Issue 1). Association for Computing Machinery.
<https://doi.org/10.1145/3580305.3599767>
- Lai, X., Meng, Z., Wang, S., Han, X., Zhou, L., Sun, T., Li, X., Wang, X., Ma, Y., & Zheng, Y. (2021). Global parametric sensitivity analysis of equivalent circuit model based on Sobol' method for lithium-ion batteries in electric vehicles. *Journal of Cleaner Production*, 294, 126246. <https://doi.org/10.1016/j.jclepro.2021.126246>
- Lee, C. H., & Wu, C. H. (2015). A Novel Big Data Modeling Method for Improving Driving Range Estimation of EVs. *IEEE Access*, 3, 1980–1993.
<https://doi.org/10.1109/ACCESS.2015.2492923>
- Lee, G., Song, J., Han, J., Lim, Y., & Park, S. (2023). Study on energy consumption characteristics of passenger electric vehicle according to the regenerative braking stages during real-world driving conditions. *Energy*, 283(August), 128745.
<https://doi.org/10.1016/j.energy.2023.128745>
- Lee, G., Song, J., Lim, Y., & Park, S. (2024). Energy consumption evaluation of passenger electric vehicle based on ambient temperature under Real-World driving conditions. *Energy Conversion and Management*, 306(December 2023), 118289.
<https://doi.org/10.1016/j.enconman.2024.118289>
- Lee, T. K., Adornato, B., & Filipi, Z. S. (2011). Synthesis of real-world driving cycles and their use for estimating PHEV energy consumption and charging opportunities: Case study for midwest/U.S. *IEEE Transactions on Vehicular Technology*, 60(9), 4153–4163. <https://doi.org/10.1109/TVT.2011.2168251>
- Lewis, J. R. (1993). Multipoint scales: Mean and median differences and observed significance levels. *International Journal of Human-Computer Interaction*, 5(4), 383-392.
- Li, H., & Wen, Y. (2020). Research on Highway Tunnel Safe Speed Limit Setting Method Considering Energy Conservation and Emission Reduction. In *IOP Conference Series: Materials Science and Engineering* (Vol. 782, No. 5, p. 052013). IOP Publishing.
- Li, P., Zhang, Y., Zhang, Y., & Zhang, K. (2021). Prediction of electric bus energy consumption with stochastic speed profile generation modelling and data driven method based on real-world big data. *Applied Energy*, 298(June), 117204.
<https://doi.org/10.1016/j.apenergy.2021.117204>
- Li, X., Wang, Z., Zhang, L., Sun, F., Cui, D., Hecht, C., Figgner, J., & Sauer, D. U. (2023). Electric vehicle behavior modeling and applications in vehicle-grid integration: An overview. *Energy*, 268(5), 126647.
<https://doi.org/10.1016/j.energy.2023.126647>

- Lin, J. (2002). A Markov process approach to driving cycle development. University of California, Davis.
- Lin, J., & Niemeier, D. A. (2002). An exploratory analysis comparing a stochastic driving cycle to California's regulatory cycle. 36, 5759–5770.
- Liu, J., Shi, P., & Fan, J. (2022). Electric Vehicle Energy Consumption Model Based on Working Condition Recognition and K-means Cluster Analysis. *Proceedings - 2022 International Conference on Manufacturing, Industrial Automation and Electronics, ICMIAE 2022*, 130–134. <https://doi.org/10.1109/ICMIAE57032.2022.00032>
- Liu, K., Wang, J., Yamamoto, T., & Morikawa, T. (2016). Modelling the multilevel structure and mixed effects of the factors influencing the energy consumption of electric vehicles. *Applied Energy*, 183, 1351–1360. <https://doi.org/10.1016/j.apenergy.2016.09.082>
- Liu, K., Wang, J., Yamamoto, T., & Morikawa, T. (2018). Exploring the interactive effects of ambient temperature and vehicle auxiliary loads on electric vehicle energy consumption. *Applied Energy*, 227(August 2017), 324–331. <https://doi.org/10.1016/j.apenergy.2017.08.074>
- Liu, K., Yamamoto, T., & Morikawa, T. (2017). Impact of road gradient on energy consumption of electric vehicles. *Transportation Research Part D: Transport and Environment*, 54, 74–81. <https://doi.org/10.1016/j.trd.2017.05.005>
- Liu, Q., Gao, C., Peng, Z., Zhang, R., & Shang, R. (2021). Smartphone positioning and accuracy analysis based on real-time regional ionospheric correction model. *Sensors*, 21(11). <https://doi.org/10.3390/s21113879>
- Liu, X., Ma, J., Zhao, X., Du, J., & Xiong, Y. (2020). Study on Driving Cycle Synthesis Method for City Buses considering Random Passenger Load. *Journal of Advanced Transportation*, 2020. <https://doi.org/10.1155/2020/3871703>
- Loganathan, M. K., Tan, C. M., Sultana, S., Hsieh, I. Y. L., Kumaraswamidhas, L. A., & Rai, R. N. (2021). Parametric performance analysis of battery operated electric vehicle. *2021 International Conference on Sustainable Energy and Future Electric Transportation, SeFet 2021*, 3–8. <https://doi.org/10.1109/SeFet48154.2021.9375788>
- Lu, C., Dong, J., & Hu, L. (2019). Energy-Efficient Adaptive Cruise Control for Electric Connected and Autonomous Vehicles. *IEEE Intelligent Transportation Systems Magazine*, 11(3), 42–55. <https://doi.org/10.1109/MITS.2019.2919556>
- Lu, Z., Zhang, Q., Yuan, Y., & Tong, W. (2020). Optimal Driving Range for Battery Electric Vehicles Based on Modeling Users' Driving and Charging Behavior. *Journal of Advanced Transportation*, 2020(1). <https://doi.org/10.1155/2020/8813137>
- Ma, Z., Jørgensen, B. N., & Ma, Z. (2024). A Scoping Review of Energy-Efficient Driving Behaviors and Applied State-of-the-Art AI Methods. *Energies*, 17(2). <https://doi.org/10.3390/en17020500>

- Mac Uidhir, T. & Rogan, F. (2024). Will 41,000 electric cars be sold in Ireland in 2024? *RTE*. <https://www.rte.ie/brainstorm/2024/0325/1439779-electric-cars-sales-ireland-2030-analysis/>
- Machado, M. R., Karray, S., & De Sousa, I. T. (2019). LightGBM: An effective decision tree gradient boosting method to predict customer loyalty in the finance industry. *14th International Conference on Computer Science and Education, ICCSE 2019, Nips*, 1111–1116. <https://doi.org/10.1109/ICCSE.2019.8845529>
- Madhusudhanan, A. K., & Na, X. (2020). Effect of a traffic speed based cruise control on an electric vehicles performance and an energy consumption model of an electric vehicle. *IEEE/CAA Journal of Automatica Sinica*, 7(2), 386–394. <https://doi.org/10.1109/JAS.2020.1003030>
- Mądział, M. (2024). Energy Modeling for Electric Vehicles Based on Real Driving Cycles: An Artificial Intelligence Approach for Microscale Analyses. *Energies*, 17(5). <https://doi.org/10.3390/en17051148>
- Mądział, M., & Campisi, T. (2023). Energy Consumption of Electric Vehicles: Analysis of Selected Parameters Based on Created Database. *Energies*, 16(3). <https://doi.org/10.3390/en16031437>
- Maity, A., & Sarkar, S. (2024). Data-driven probabilistic energy consumption estimation for battery electric vehicles with model uncertainty. *International Journal of Green Energy*, 21(9), 1986–2003. <https://doi.org/10.1080/15435075.2023.2276174>
- Mamala, J., Graba, M., Mitrovic, J., Prażnowski, K., & Stasiak, P. (2023). Analysis of speed limit and energy consumption in electric vehicles. *Combustion Engines*, 195(4), 83–89. <https://doi.org/10.19206/CE-169370>
- Manoharan, A., Begam, K. M., Aparow, V. R., & Sooriamoorthy, D. (2022). Artificial Neural Networks, Gradient Boosting and Support Vector Machines for electric vehicle battery state estimation: A review. *Journal of Energy Storage*, 55(PA), 105384. <https://doi.org/10.1016/j.est.2022.105384>
- Mansour, S., & Raeesi, M. (2023). Performance assessment of fuel cell and electric vehicles taking into account the fuel cell degradation, battery lifetime, and heating, ventilation, and air conditioning system. *International Journal of Hydrogen Energy* <https://doi.org/10.1016/j.ijhydene.2023.05.315>
- Martins, L. D. C., Tordecilla, R. D., Castaneda, J., Juan, A. A., & Faulin, J. (2021). Electric vehicle routing, arc routing, and team orienteering problems in sustainable transportation. *Energies*, 14(16), 1–30. <https://doi.org/10.3390/en14165131>
- Marzbnai, F., Osman, A. H., & Hassan, M. S. (2023). Electric Vehicle Energy Demand Prediction Techniques: An In-Depth and Critical Systematic Review. *IEEE Access*, 11(August), 96242–96255. <https://doi.org/10.1109/ACCESS.2023.3308928>

- Maurício Pereira Magalhães de Novaes Santos, Fernando Wesley Cavalcanti de Araújo, Guilherme Medeiros Soares de Andrade, F. S. M. (2020). CONSTRUCTION OF DRIVING CYCLES: CASE STUDY FOR MICROTRIP AND MARKOV CHAIN METHODS' USING REAL DATA. 34o Congresso de Pesquisa e Ensino Em Transporte Da ANPET, Ldc, 12.
- Mavlonov, J., Ruzimov, S., Tonoli, A., Amati, N., & Mukhitdinov, A. (2023). Sensitivity Analysis of Electric Energy Consumption in Battery Electric Vehicles with Different Electric Motors. *World Electric Vehicle Journal*, 14(2), 1–16. <https://doi.org/10.3390/wevj14020036>
- Mayakuntla, S. K., & Verma, A. (2018). A novel methodology for construction of driving cycles for Indian cities. *Transportation Research Part D: Transport and Environment*, 65(October), 725–735. <https://doi.org/10.1016/j.trd.2018.10.013>
- McNally, M. G. (2007). The four-step model. In *Handbook of transport modelling* (Vol. 1, pp. 35-53). Emerald Group Publishing Limited.
- Mediouni, H., Ezzouhri, A., Charouh, Z., El Harouri, K., El Hani, S., & Ghogho, M. (2022). Energy Consumption Prediction and Analysis for Electric Vehicles: A Hybrid Approach. *Energies*, 15(17), 1–17. <https://doi.org/10.3390/en15176490>
- Mei, P., Karimi, H. R., Huang, C., Chen, F., & Yang, S. (2023). Remaining driving range prediction for electric vehicles: Key challenges and outlook. *IET Control Theory and Applications*, 17(14), 1875–1893. <https://doi.org/10.1049/cth2.12486>
- Mekky, M. F., & Collins, A. R. (2024). The Impact of state policies on electric vehicle adoption -A panel data analysis. *Renewable and Sustainable Energy Reviews*, 191(July 2023), 114014. <https://doi.org/10.1016/j.rser.2023.114014>
- Meseguer, J. E., Toh, C. K., Calafate, C. T., Cano, J. C., & Manzoni, P. (2017). Drivingstyles: A mobile platform for driving styles and fuel consumption characterization. *Journal of Communications and Networks*, 19(2), 162–168. <https://doi.org/10.1109/JCN.2017.000025>
- Mieth, R., Gauterin, F., Pauli, F., & Kraus, H. (2022). Transfer of Statistical Customer Data into Relevant Parameters for the Design of Vehicle Drive Systems. *Vehicles*, 4(1), 137–144. <https://doi.org/10.3390/vehicles4010009>
- Miri, I., Fotouhi, A., & Ewin, N. (2021). Electric vehicle energy consumption modelling and estimation—A case study. *International Journal of Energy Research*, 45(1), 501–520. <https://doi.org/10.1002/er.5700>
- Modi, S., & Bhattacharya, J. (2022). A system for electric vehicle's energy-aware routing in a transportation network through real-time prediction of energy consumption. *Complex and Intelligent Systems*, 8(6), 4727–4751. <https://doi.org/10.1007/s40747-022-00727-4>
- Modi, S., Bhattacharya, J., & Basak, P. (2020). Estimation of energy consumption of electric vehicles using Deep Convolutional Neural Network to reduce driver's range anxiety. *ISA Transactions*, 98, 454–470. <https://doi.org/10.1016/j.isatra.2019.08.055>

- Morey, J. E., Limanond, T., & Niemeier, D. A. (2009). Validity of chase car data used in developing emissions cycles. *Journal of Transportation and Statistics*, 3(2), 15–28.
- Morlock, F., Rolle, B., Bauer, M., & Sawodny, O. (2020). Forecasts of Electric Vehicle Energy Consumption Based on Characteristic Speed Profiles and Real-Time Traffic Data. *IEEE Transactions on Vehicular Technology*, 69(2), 1404–1418. <https://doi.org/10.1109/TVT.2019.2957536>
- Mouzouras, M. (2022). *New methodologies for assessing the behavior of vehicles under real driving emissions testing*. [Doctoral Dissertation, University of Bath]. https://purehost.bath.ac.uk/ws/portalfiles/portal/236904364/Marios_Mouzouras_PhD_thesis_corrected.pdf
- Nan, S., Tu, R., Li, T., Sun, J., & Chen, H. (2022). From driving behavior to energy consumption: A novel method to predict the energy consumption of electric bus. *Energy*, 261(PA), 125188. <https://doi.org/10.1016/j.energy.2022.125188>
- Neubauer, J., & Wood, E. (2014). Thru-life impacts of driver aggression, climate, cabin thermal management, and battery thermal management on battery electric vehicle utility. *Journal of Power Sources*, 259, 262–275. <https://doi.org/10.1016/j.jpowsour.2014.02.083>
- Noel, L., Zarazua de Rubens, G., Sovacool, B. K., & Kester, J. (2019). Fear and loathing of electric vehicles: The reactionary rhetoric of range anxiety. *Energy Research and Social Science*, 48(April 2018), 96–107. <https://doi.org/10.1016/j.erss.2018.10.001>
- Nossent, J., Elsen, P., & Bauwens, W. (2011). Sobol'sensitivity analysis of a complex environmental model. *Environmental modelling & software*, 26(12), 1515-1525.
- Nouri, P., & Morency, C. (2017). Evaluating microtrip definitions for developing driving cycles. *Transportation Research Record*, 2627, 86–92. <https://doi.org/10.3141/2627-10>
- Nyberg, P., Frisk, E., & Nielsen, L. (2014). Generation of Equivalent Driving Cycles Using Markov Chains and Mean Tractive Force Components. In IFAC Proceedings Volumes (Vol. 47, Issue 3). IFAC. <https://doi.org/10.3182/20140824-6-ZA-1003.02239>
- OBD Solutions. (2025). *OBDLINK MX+*. www.obdlink.com/products/obdlink-mxp/
- OpenStreetMap. (2024A). Ireland and Northern Ireland. <https://www.openstreetmap.org/#map>
- OpenStreetMap. (2024B). Tag:highway=primary. <https://wiki.openstreetmap.org/wiki/Tag:highway%3Dprimary>
- Oubelaid, A., Taib, N., Nikolovski, S., Alharbi, T. E. A., Rekioua, T., Flah, A., & Ghoneim, S. S. M. (2022). Intelligent Speed Control and Performance Investigation of a Vector Controlled Electric Vehicle Considering Driving Cycles. *Electronics* (Switzerland), 11(13). <https://doi.org/10.3390/electronics11131925>

- Owen, A. B. (2014). Sobol' indices and Shapley value. *SIAM/ASA Journal on Uncertainty Quantification*, 2(1), 245-251.
- Pacheco, F., Cerrada, M., & Huertas, J. I. (2024). Threshold-guided multi-objective Generative Adversarial Networks for constructing artificial yet representative driving cycles. *Engineering Applications of Artificial Intelligence*, 129(December 2023), 107665. <https://doi.org/10.1016/j.engappai.2023.107665>
- Pal, P., Probst, D., Pei, Y., Zhang, Y., Traver, M., Cleary, D., & Som, S. (2017). Numerical Investigation of a Gasoline-Like Fuel in a Heavy-Duty Compression Ignition Engine Using Global Sensitivity Analysis. *SAE International Journal of Fuels and Lubricants*, 10(1), 56–68. <https://doi.org/10.4271/2017-01-0578>
- Palit, T., Bari, A. B. M. M., & Karmaker, C. L. (2022). An integrated Principal Component Analysis and Interpretive Structural Modeling approach for electric vehicle adoption decisions in sustainable transportation systems. *Decision Analytics Journal*, 4(June), 100119. <https://doi.org/10.1016/j.dajour.2022.100119>
- Pamidimukkala, A., Kermanshachi, S., Rosenberger, J. M., & Hladik, G. (2024). Barriers and motivators to the adoption of electric vehicles: A global review. *Green Energy and Intelligent Transportation*, 3(2), 100153. <https://doi.org/10.1016/j.geits.2024.100153>
- Pan, C., Dai, W., Chen, L., Chen, L., & Wang, L. (2017). Driving range estimation for electric vehicles based on driving condition identification and forecast. *AIP Advances*, 105206(September 2017). <https://doi.org/10.1063/1.4993945>
- Pan, D., Sun, F., He, H., & Peng, J. (2015). *Driving Cycle Construction of City Road for Hybrid Bus Based on Markov Process*. *Iiicec*, 1159–1164.
- Pan, Y., Fang, W., & Zhang, W. (2023). Development of an energy consumption prediction model for battery electric vehicles in real-world driving: A combined approach of short-trip segment division and deep learning. *Journal of Cleaner Production*, 400(December 2022), 136742. <https://doi.org/10.1016/j.jclepro.2023.136742>
- Patel, P., & Aggarwal, S. G. (2022). On the techniques and standards of particulate matter sampling. *Journal of the Air & Waste Management Association*, 72(8), 791-814.
- Pavlovic, J., Fontaras, G., Broekaert, S., Ciuffo, B., Ktistakis, M. A., & Grigoratos, T. (2021). How accurately can we measure vehicle fuel consumption in real world operation? *Transportation Research Part D: Transport and Environment*, 90(December 2020), 102666. <https://doi.org/10.1016/j.trd.2020.102666>
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. *the Journal of machine Learning research*, 12, 2825-2830.
- Petersen, P., & Sax, E. (2022). A Fully Automated Methodology for the Selection and Extraction of Energy-Relevant Features for the Energy Consumption of Battery Electric Vehicles. *SN Computer Science*, 3(5), 1–13. <https://doi.org/10.1007/s42979-022-01233-7>

- Pettersson, P., Berglund, S., Jacobson, B., Fast, L., Johannesson, P., & Santandrea, F. (2018). A proposal for an operating cycle description format for road transport missions. *European Transport Research Review*, 10(2). <https://doi.org/10.1186/s12544-018-0298-4>
- Pevec, D., Babic, J., Carvalho, A., Ghiassi-Farrokhfal, Y., Ketter, W., & Podobnik, V. (2019). Electric vehicle range anxiety: An obstacle for the personal transportation (r)evolution? *2019 4th International Conference on Smart and Sustainable Technologies, SpliTech 2019*, i. <https://doi.org/10.23919/SpliTech.2019.8783178>
- Pianosi, F., & Wagener, T. (2015). A simple and efficient method for global sensitivity analysis based on cumulative distribution functions. *Environmental Modelling & Software*, 67, 1-11.
- Ping, P., Qin, W., Xu, Y., Miyajima, C., & Takeda, K. (2019). Impact of driver behavior on fuel consumption: Classification, evaluation and prediction using machine learning. *IEEE Access*, 7, 78515–78532. <https://doi.org/10.1109/ACCESS.2019.2920489>
- Prehofer, C., & Mehmood, S. (2020). Big Data Architectures for Vehicle Data Analysis. *Proceedings - 2020 IEEE International Conference on Big Data, Big Data 2020*, 3404–3412. <https://doi.org/10.1109/BigData50022.2020.9378397>
- PTV Group. (2025). PTV Visum – The world’s leading transport planning software. <https://www.ptvgroup.com/en/products/ptv-visum>
- POWERDRIVE. (2025). *POWERDRIVE - Power electronics optimisation for next generation electric vehicle components*. <https://www.powerdriveproject.eu/>
- QGIS.org. (2024). QGIS Geographic Information System. QGIS Association. <http://www.qgis.org>
- Qi, X., Wu, G., Boriboonsomsin, K., & Barth, M. J. (2018). Data-driven decomposition analysis and estimation of link-level electric vehicle energy consumption under real-world traffic conditions. *Transportation Research Part D: Transport and Environment*, 64(August 2017), 36–52. <https://doi.org/10.1016/j.trd.2017.08.008>
- Raghu, M., Blumer, K., Sayres, R., Obermeyer, Z., Kleinberg, B., Mullainathan, S., & Kleinberg, J. (2019, May). Direct uncertainty prediction for medical second opinions. In *International Conference on Machine Learning* (pp. 5281-5290). PMLR.
- Ramachandaramurthy, V. K., Ajmal, A. M., Kasinathan, P., Tan, K. M., Yong, J. Y., & Vinoth, R. (2023). Social acceptance and preference of EV users—a review. *IEEE Access*, 11, 11956-11972.
- Ramai, C., Ramnarine, V., Ramharack, S., Bahadoorsingh, S., & Sharma, C. (2022). Framework for Building Low-Cost OBD-II Data-Logging Systems for Battery Electric Vehicles. *Vehicles*, 4(4), 1209–1222. <https://doi.org/10.3390/vehicles4040064>

- Rangaraju, S., De Vroey, L., Messagie, M., Mertens, J., & Van Mierlo, J. (2015). Impacts of electricity mix, charging profile, and driving behavior on the emissions performance of battery electric vehicles: A Belgian case study. *Applied Energy*, 148(X), 496–505. <https://doi.org/10.1016/j.apenergy.2015.01.121>
- Rauh, N., Franke, T., & Krems, J. F. (2015). Understanding the impact of electric vehicle driving experience on range anxiety. *Human Factors*, 57(1), 177–187. <https://doi.org/10.1177/0018720814546372>
- Razavi, Saman; Gupta, H. V. (2016). A new framework for comprehensive, robust, and efficient global sensitivity analysis: 2. Application. *Water Resources Research*, 52, 440–455. <https://doi.org/10.1111/j.1752-1688.1969.tb04897.x>
- Razavi, S., Jakeman, A., Saltelli, A., Prieur, C., Iooss, B., Borgonovo, E., Plischke, E., Lo Piano, S., Iwanaga, T., Becker, W., Tarantola, S., Guillaume, J. H. A., Jakeman, J., Gupta, H., Melillo, N., Rabitti, G., Chabridon, V., Duan, Q., Sun, X., ... Maier, H. R. (2021). The Future of Sensitivity Analysis: An essential discipline for systems modeling and policy support. *Environmental Modelling and Software*, 137(December 2020). <https://doi.org/10.1016/j.envsoft.2020.104954>
- Regulation 2016/646. COMMISSION REGULATION (EU) 2016/646 of 20 April 2016 amending Regulation (EC) No 692/2008 as regards emissions from light passenger and commercial vehicles (Euro 6). <https://eur-lex.europa.eu/eli/reg/2016/646/oj/eng>
- Renardy, M., Joslyn, L. R., Millar, J. A., & Kirschner, D. E. (2021). To Sobol or not to Sobol? The effects of sampling schemes in systems biology applications. *Mathematical biosciences*, 337, 108593.
- Rhode, S., Van Vaerenbergh, S., & Pfriem, M. (2020). Power prediction for electric vehicles using online machine learning. *Engineering Applications of Artificial Intelligence*, 87(October 2019), 103278. <https://doi.org/10.1016/j.engappai.2019.103278>
- Ricotti, M. E., & Zio, E. (1999). Neural network approach to sensitivity and uncertainty analysis. *Reliability Engineering and System Safety*, 64(1), 59–71. [https://doi.org/10.1016/S0951-8320\(98\)00057-X](https://doi.org/10.1016/S0951-8320(98)00057-X)
- Rolim, C. C., Gonçalves, G. N., Farias, T. L., & Rodrigues, Ó. (2012). Impacts of Electric Vehicle Adoption on Driver Behavior and Environmental Performance. *Procedia - Social and Behavioral Sciences*, 54, 706–715. <https://doi.org/10.1016/j.sbspro.2012.09.788>
- Romano, L., Godio, M., Johannesson, P., Bruzelius, F., Ghandriz, T., & Jacobson, B. (2023). Development of the Västra Götaland Operating Cycle for Long-Haul Heavy-Duty Vehicles. *IEEE Access*, 11(June), 73268–73302. <https://doi.org/10.1109/ACCESS.2023.3295989>
- Romero, C. A., Mejía, L. A., & Acosta, R. (2017). Engine data collection and development of a pilot driving cycle for Pereira city by using low cost diagnostic tools. *Ingeniería Y Competitividad*, 19(2), 11–24. <https://doi.org/10.25100/iyc.v19i2.5287>

- Ruan, J., Walker, P. D., Watterson, P. A., & Zhang, N. (2016). The dynamic performance and economic benefit of a blended braking system in a multi-speed battery electric vehicle. *Applied Energy*, 183, 1240–1258.
<https://doi.org/10.1016/j.apenergy.2016.09.057>
- Salihu, F., & Demir, Y. K. (2023). Driving Cycle for Passenger Cars on Urban Roads in Pristina, Kosovo. *Baltic Journal of Road and Bridge Engineering*, 18(1), 69–93.
<https://doi.org/10.7250/bjrbe.2023-18.589>
- Salihu, F., Demir, Y. K., & Demir, H. G. (2023). Effect of road slope on driving cycle parameters of urban roads. *Transportation Research Part D: Transport and Environment*, 118(January), 103676. <https://doi.org/10.1016/j.trd.2023.103676>
- Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M., & Tarantola, S. (2010). Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. *Computer physics communications*, 181(2), 259–270.
- Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., ... & Tarantola, S. (2008). *Global sensitivity analysis: the primer*. John Wiley & Sons.
- Sanguinetti, A., Queen, E., Yee, C., & Akanesuvan, K. (2020). Average impact and important features of onboard eco-driving feedback: A meta-analysis. *Transportation Research Part F: Traffic Psychology and Behaviour*, 70, 1–14.
- Santhosh, B. J., Benvin, S., Arvind, R., Aswini, V., Maranan, R., & Thrinath, B. V. S. (2023). AI Based Electric Automobile Battery Drain Forecasting System. *Proceedings of the 4th International Conference on Smart Electronics and Communication, ICOSEC 2023*, 756–761.
<https://doi.org/10.1109/ICOSEC58147.2023.10276243>
- Saputra, Y. M., Hoang, D. T., Nguyen, D. N., Dutkiewicz, E., Mueck, M. D., & Srikanthswara, S. (2019, December). Energy demand prediction with federated learning for electric vehicle networks. In *2019 IEEE global communications conference (GLOBECOM)* (pp. 1–6). IEEE.
- Sasank, T. V., Akash, K. B., Siddharth, I. S., & Susila, M. (2023). State of Charge Estimation and Analysis For an Electric Vehicle. *2023 International Conference on Recent Advances in Electrical, Electronics, Ubiquitous Communication, and Computational Intelligence, RAEEUCCI 2023*, 1–6.
<https://doi.org/10.1109/RAEEUCCI57140.2023.10134396>
- Schäfers, L., Iravanimanesh, S., Franke, K., Savelsberg, R., & Pischinger, S. (2024). Holistic Sensitivity Analysis for Long-Term Energy Demand Prediction of Battery Electric Vehicles. *Automotive Innovation*, 7(4), 669–681.
<https://doi.org/10.1007/s42154-024-00292-1>
- Schäfers, L., Silber, D., Savelsberg, R., & Pischinger, S. (2023). Efficient determination of driving resistance through system identification based on driving route information and weather data. *Expert Systems with Applications*, 232(June), 0–2.
<https://doi.org/10.1016/j.eswa.2023.120755>

- Schnabel, W., & Lohse, D. (1997). GRUNDLAGEN DER STRASSENVERKEHRSTECHNIK UND DER VERKEHRSPPLANUNG. 2 BAENDE.
- Schneck, A., & Nökel, K. (2020). Accelerating traffic assignment with customizable contraction hierarchies. *Transportation research record*, 2674(1), 188-196.
- SenSight Technologies. (2025). *AutoPulse: Affordable OBD Telematics*. <https://www.autowiz.in/autopulse.html>
- Settey, T., Gnap, J., Synák, F., Skrucány, T., & Dočkalík, M. (2021). Research into the impacts of driving cycles and load weight on the operation of a light commercial electric vehicle. *Sustainability (Switzerland)*, 13(24). <https://doi.org/10.3390/su132413872>
- Shankar, R., Marco, J., & Assadian, F. (2012). A methodology to determine drivetrain efficiency based on external environment. 2012 IEEE International Electric Vehicle Conference, IEVC 2012. <https://doi.org/10.1109/IEVC.2012.6183192>
- Shapley, L. S. (1953). A value for n-person games.
- Shehab El Din, M., Hussein, A. A., & Abdel-Hafez, M. F. (2018). Improved battery SOC estimation accuracy using a modified UKF with an adaptive cell model under real EV operating conditions. *IEEE Transactions on Transportation Electrification*, 4(2), 408–417. <https://doi.org/10.1109/TTE.2018.2802043>
- Shibata, S. & Nakagawa, T. (2015). Mathematical Model of Electric Vehicle Power Consumption for Traveling and Air-Conditioning. *Journal of Energy and Power Engineering*, 9(3), 269–275. <https://doi.org/10.17265/1934-8975/2015.03.006>
- Shiels, Philip (2013). National Transport Model – Overview and Applications. https://www.tii.ie/media/suuhlej3/26_philip-shiels_nra-national-transport-model.pdf
- Sim, A. X. A., & Sitohang, B. (2014). OBD-II standard car engine diagnostic software development. *Proceedings of 2014 International Conference on Data and Software Engineering, ICODSE 2014*, 1–5. <https://doi.org/10.1109/ICODSE.2014.7062704>
- Singirikonda, S., & Obuselu, YP. (2023). Investigation on performance evaluation of electric vehicle batteries under different drive cycles. *Journal of Energy Storage*, 63(February), 106966. <https://doi.org/10.1016/j.est.2023.106966>
- Sivak, M., & Schoettle, B. (2012). Eco-driving: Strategic, tactical, and operational decisions of the driver that influence vehicle fuel economy. *Transport Policy*, 22, 96–99. <https://doi.org/10.1016/j.tranpol.2012.05.010>
- Skuzza, A., & Jurecki, R. S. (2022). Analysis of factors affecting the energy consumption of an EV vehicle - a literature study. *IOP Conference Series: Materials Science and Engineering*, 1247(1), 012001. <https://doi.org/10.1088/1757-899x/1247/1/012001>

- Smith, L. N. (2017). Cyclical learning rates for training neural networks. *Proceedings - 2017 IEEE Winter Conference on Applications of Computer Vision, WACV 2017, April 2015*, 464–472. <https://doi.org/10.1109/WACV.2017.58>
- Smith, II, H., Akhtar, S., Caulfield, B., & O'Mahony, M. (2024). Validity of GPS data in driving cycles. *IET Intelligent Transport Systems*, 18, 3034-3040.
- Sobol', I. M. (1967). The distribution of points in a cube and the approximate evaluation of integrals. *USSR Computational mathematics and mathematical physics*, 7, 86-112.
- Sobol', I. Y. M. (1990). On sensitivity estimation for nonlinear mathematical models. *Matematicheskoe modelirovanie*, 2(1), 112-118.
- Song, E., Nelson, B. L., & Staum, J. (2016). Shapley effects for global sensitivity analysis: Theory and computation. *SIAM/ASA Journal on Uncertainty Quantification*, 4(1), 1060-1083.
- Sovacool, B. K., Kester, J., Noel, L., & de Rubens, G. Z. (2018). The demographics of decarbonizing transport: The influence of gender, education, occupation, age, and household size on electric mobility preferences in the Nordic region. *Global Environmental Change*, 52(June), 86–100. <https://doi.org/10.1016/j.gloenvcha.2018.06.008>
- Stiniez. (2025). *Piston - OBD2 diagnostic app for Android*. <https://piston.app/>
- Straub, T., Frey, M., & Gauterin, F. (2022). Learning from the Fleet: Map Attributes for Energetic Representation of Driving Profiles. *IEEE Transactions on Intelligent Transportation Systems*, 23(1), 471–482. <https://doi.org/10.1109/TITS.2020.3011976>
- Sugiyanto, G., Jajang, & Santi, M. Y. (2019). The impact of lowering speed limit on mobility and the environment. In *AIP Conference Proceedings* (Vol. 2094, No. 1, p. 020019). AIP Publishing LLC.
- Sun, S., Zhang, J., Bi, J., Wang, Y., & Moghaddam, M. H. Y. (2019). A Machine Learning Method for Predicting Driving Range of Battery Electric Vehicles. *Journal of Advanced Transportation*, 2019. <https://doi.org/10.1155/2019/4109148>
- Sun, X., & Fu, J. (2024). Many-objective optimization of BEV design parameters based on gradient boosting decision tree models and the NSGA-III algorithm considering the ambient temperature. *Energy*, 288(December 2023), 129840. <https://doi.org/10.1016/j.energy.2023.129840>
- Suryavanshi, S. S., & Ghanegaonkar, P. M. (2023). Repercussion of effect of different drive cycles on evaluation of electrical consumption for electric four-wheeler to achieve optimal performance of electric vehicle. *Energy Storage*, 5(7), 1–14. <https://doi.org/10.1002/est2.468>
- Sweeting, W. J., Hutchinson, A. R., & Savage, S. D. (2011). Factors affecting electric vehicle energy consumption. *International Journal of Sustainable Engineering*, 4(3), 192–201. <https://doi.org/10.1080/19397038.2011.592956>

- Syakur, M. A., Khotimah, B. K., Rochman, E. M., & Satoto, B. D. (2018, April). Integration k-means clustering method and elbow method for identification of the best customer profile cluster. In *IOP conference series: materials science and engineering* (Vol. 336, p. 012017). IOP Publishing.
- Szumaska, E. M. (2025). Regenerative Braking Systems in Electric Vehicles: A Comprehensive Review of Design, Control Strategies, and Efficiency Challenges. *Energies*, 18(10), 2422.
- Szumaska, E. M., & Jurecki, R. S. (2021). Parameters influencing on electric vehicle range. *Energies*, 14(16). <https://doi.org/10.3390/en14164821>
- Tang, J., McNabola, A., & Misstear, B. (2020). The potential impacts of different traffic management strategies on air pollution and public health for a more sustainable city: A modelling case study from Dublin, Ireland. *Sustainable Cities and Society*, 60(May 2019), 102229. <https://doi.org/10.1016/j.scs.2020.102229>
- Teymouri, P., Sarbakhsh, P., Taghipour, H., Roshani, M., & Dehghanzadeh, R. (2023). Modeling fuel consumption and emissions by taxis in Tabriz, Iran: uncertainty and sensitivity analyses. *Environmental Science and Pollution Research*, 30(33), 80655–80675. <https://doi.org/10.1007/s11356-023-27883-5>
- Thorgeirsson, A. T. (2024). *Probabilistic prediction of energy demand and driving range for electric vehicles with federated learning* (p. 72). KIT Scientific Publishing.
- Thorgeirsson, A. T., Scheubner, S., Funfgeld, S., & Gauterin, F. (2021). Probabilistic Prediction of Energy Demand and Driving Range for Electric Vehicles with Federated Learning. *IEEE Open Journal of Vehicular Technology*, 2(April), 151–161. <https://doi.org/10.1109/OJVT.2021.3065529>
- Timmers, V. R., & Achten, P. A. (2016). Non-exhaust PM emissions from electric vehicles. *Atmospheric environment*, 134, 10–17.
- Tong, H. Y., & Hung, W. T. (2010). A framework for developing driving cycles with on-road driving data. *Transport Reviews*, 30(5), 589–615. <https://doi.org/10.1080/01441640903286134>
- Torque. 2025. *Torque*. <https://torque-bhp.com/>
- Tosin, M., Côrtes, A. & Cunha Jr., A. (2020). A Tutorial on Sobol' Global Sensitivity Analysis Applied to Biological Models. da Silva F.A.B., Carels N., Trindade dos Santos M., Lopes F.J.P. Networks in Systems Biology: Applications for Disease Modeling, Springer, Cham, 2020, *Mechanisms and Machine Science*, ff10.1007/978-3-030-51862-2_6ff. fffal-02967410f
- Tran, T. B., Kolmanovsky, I., Biberstein, E., Makke, O., Tharayil, M., & Gusikhin, O. (2024). Effect of Wind on Electric Vehicle Energy Consumption: Sensitivity Analyses and Implications for Range Estimation and Optimal Routing. *ACM Journal on Autonomous Transportation Systems*, 1(2), 1–31. <https://doi.org/10.1145/3633460>
- Transport Infrastructure Ireland. (2024A). Weather Sites. <https://data.tii.ie/Datasets/Its/DatexII/WeatherSites/index.html>

- Transportation Infrastructure Ireland. (2024B). Our National Road Network.
<https://www.tii.ie/en/roads-tolling/operations-and-maintenance/>
- Tsai, J. H., Chiang, H. L., Hsu, Y. C., Peng, B. J., & Hung, R. F. (2005). Development of a local real world driving cycle for motorcycles for emission factor measurements. *Atmospheric Environment*, 39(35), 6631–6641.
<https://doi.org/10.1016/j.atmosenv.2005.07.040>
- Tsang, M., Cheng, D., & Liu, Y. (2018). Detecting Statistical Interactions From Neural Network Weights. *ICLR 2018*, 1–21.
- Tsvetkova, O., & Ouarda, T. B. M. J. (2021). A review of sensitivity analysis practices in wind resource assessment. *Energy Conversion and Management*, 238(April), 114112.
<https://doi.org/10.1016/j.enconman.2021.114112>
- Tuan, T. D., Phuoc, S. H., & Quang, T. N. (2023). An Experimental Research to Evaluate the Effective of Regenerative Braking System Based on Actual Driving Conditions. *Proceedings of 2023 International Conference on System Science and Engineering, ICSSE 2023*, 249–253. <https://doi.org/10.1109/ICSSE58758.2023.10227190>
- Tutuianu, M. *et al.* (2015) ‘Development of the World-wide harmonized Light duty Test Cycle (WLTC) and a possible pathway for its introduction in the European legislation’, *Transportation Research Part D: Transport and Environment*, 40, pp. 61–75. <https://doi.org/10.1016/j.trd.2015.07.011>.
- UCD Library. (2015). GIS at UCD and on the Web: Find Spatial Data & Other Datasets for Ireland: Legal Towns and Cities – 2015.
<https://libguides.ucd.ie/gisguide/findspatialdata>
- Ullah, I., Liu, K., Yamamoto, T., Al Mamlook, R. E., & Jamal, A. (2022). A comparative performance of machine learning algorithm to predict electric vehicles energy consumption: A path towards sustainability. *Energy and Environment*, 33(8), 1583–1612. <https://doi.org/10.1177/0958305X211044998>
- Unni, K., & Thale, S. (2023). Energy Consumption Analysis for the Prediction of Battery Residual Energy in Electric Vehicles. *Engineering, Technology and Applied Science Research*, 13(3), 11011–11019. <https://doi.org/10.48084/etasr.5868>
- van Eekelen, W., den Hertog, D., & van Leeuwen, J. S. (2022). MAD dispersion measure makes extremal queue analysis simple. *INFORMS Journal on Computing*, 34(3), 1681–1692.
- Varga, B. O., Sagoian, A., & Mariasiu, F. (2019). Prediction of electric vehicle range: A comprehensive review of current issues and challenges. *Energies*, 12(5).
<https://doi.org/10.3390/en12050946>
- Vatanparvar, K., Faezi, S., Burago, I., Levorato, M., & Al Faruque, M. A. (2019). Extended Range Electric Vehicle with Driving Behavior Estimation in Energy Management. *IEEE Transactions on Smart Grid*, 10(3), 2959–2968.
<https://doi.org/10.1109/TSG.2018.2815689>

- Veepeak Inc. 2025. *OBDCheck BLE+ Plus*.
<https://veepeak.com/products/obdcheck-ble-plus>
- Vepsäläinen, J., Otto, K., Lajunen, A., & Tammi, K. (2019). Computationally efficient model for energy demand prediction of electric city bus in varying operating conditions. *Energy*, 169, 433–443. <https://doi.org/10.1016/j.energy.2018.12.064>
- Vepsäläinen, J., Ritari, A., Lajunen, A., Kivekäs, K., & Tammi, K. (2018). Energy uncertainty analysis of electric buses. *Energies*, 11(12).
<https://doi.org/10.3390/en11123267>
- Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., ... SciPy 1.0 Contributors. (2020). SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17, 261–272.
<https://doi.org/10.1038/s41592-019-0686-2>
- Vrbanić, F., Miletić, M., Tišljarić, L., & Ivanjko, E. (2022). Influence of Variable Speed Limit Control on Fuel and Electric Energy Consumption, and Exhaust Gas Emissions in Mixed Traffic Flows. *Sustainability (Switzerland)*, 14(2).
<https://doi.org/10.3390/su14020932>
- Wager, G., Whale, J., & Braunl, T. (2016). Driving electric vehicles at highway speeds: The effect of higher driving speeds on energy consumption and driving range for electric vehicles in Australia. *Renewable and sustainable energy reviews*, 63, 158–165.
- Waldowski, P., Marker, S., Schulz, A., & Schindler, V. (2011). BEV , REEV or PHEV ? – Deciding Between Different Alternative Drives Based on Measured Individual Operational Profiles. Time, December
- Wang, G., Makino, K., Harmandayan, A., & Wu, X. (2020). Eco-driving behaviors of electric vehicle users: A survey study. *Transportation Research Part D: Transport and Environment*, 78(December 2019), 102188.
<https://doi.org/10.1016/j.trd.2019.11.017>
- Wang, H., Zhang, X., & Ouyang, M. (2015). Energy consumption of electric vehicles based on real-world driving patterns: A case study of Beijing. *Applied Energy*, 157(2015), 710–719. <https://doi.org/10.1016/j.apenergy.2015.05.057>
- Wang, L., Ma, J., Zhao, X., & Li, X. (2021). Development of a typical urban driving cycle for battery electric vehicles based on kernel principal component analysis and random forest. *IEEE Access*, 9, 15053–15065.
<https://doi.org/10.1109/ACCESS.2021.3052820>
- Wang, J., Besselink, I., & Nijmeijer, H. (2018). Battery electric vehicle energy consumption prediction for a trip based on route information. *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, 232(11), 1528–1542. <https://doi.org/10.1177/0954407017729938>
- Wang, X., Ye, P., Deng, Y., Yuan, Y., Zhu, Y., & Ni, H. (2023). Influence of Different Data Interpolation Methods for Sparse Data on the Construction Accuracy of Electric Bus Driving Cycle. *Electronics (Switzerland)*, 12(6).
<https://doi.org/10.3390/electronics12061377>

- Wang, Z., Zhang, J., Liu, P., Qu, C., Li, X., Wang, Z., Pina, A., Ferrão, P., Fournier, J., Lacarrière, B., & Corre, O. Le. (2019). Driving Cycle Construction for Electric Vehicles Based on Markov and Cooling Driving Cycle Construction for Electric Vehicles on Markov Chain and Monte Carlo Method : A Case Study in Beijing Chain and Monte Carlo Method : A Case Study in Beijing Assessing. *Energy Procedia*, 158, 2494–2499. <https://doi.org/10.1016/j.egypro.2019.01.389>
- Wei, C., Wang, X., Chen, Y., Wu, H., & Chen, Y. (2023). Adaptive Fuzzy Power Management Strategy for Extended-Range Electric Logistics Vehicles Based on Driving Pattern Recognition. *Actuators*, 12(11). <https://doi.org/10.3390/act12110410>
- Wei, T. (2013). A review of sensitivity analysis methods in building energy analysis. *Renewable and Sustainable Energy Reviews*, 20, 411–419. <https://doi.org/10.1016/j.rser.2012.12.014>
- Weldon, P., Morrissey, P., Brady, J., & O'Mahony, M. (2016). An investigation into usage patterns of electric vehicles in Ireland. *Transportation Research Part D: Transport and Environment*, 43(2016), 207–225. <https://doi.org/10.1016/j.trd.2015.12.013>
- Welsh Government, (2020). Welsh 20mph Task Force Group – Final Report. <https://www.gov.wales/sites/default/files/publications/2020-07/20mph-task-force-group-report.pdf>
- Wooley, J. E., Zito, R., Dyson, C. B., STAZIC, B., & TAYLOR, M. A. P. (2002). Impacts of Lower Speed Limits in South Australia. *IATSS Research*, 26(2), 6–17. [https://doi.org/10.1016/s0386-1112\(14\)60038-8](https://doi.org/10.1016/s0386-1112(14)60038-8)
- Wu, X., Freese, D., Cabrera, A., & Kitch, W. A. (2015). Electric vehicles' energy consumption measurement and estimation. *Transportation Research Part D: Transport and Environment*, 34, 52–67. <https://doi.org/10.1016/j.trd.2014.10.007>
- Wu, Y., Huang, Z., Zhang, R., Huang, P., Gao, Y., Li, H., Liu, Y., & Peng, J. (2023). Driving style-aware energy management for battery/supercapacitor electric vehicles using deep reinforcement learning. *Journal of Energy Storage*, 73(PC), 109199. <https://doi.org/10.1016/j.est.2023.109199>
- Xiang, Y., Liu, J., Li, R., Li, F., Gu, C., & Tang, S. (2016). Economic planning of electric vehicle charging stations considering traffic constraints and load profile templates. *Applied Energy*, 178, 647–659. <https://doi.org/10.1016/j.apenergy.2016.06.021>
- Xiao, Z., Dui-jia, Z., & Jun-min, S. (2012). Energy Procedia A Synthesis of Methodologies and Practices for Developing Driving Cycles. 16, 1868–1873. <https://doi.org/10.1016/j.egypro.2012.01.286>
- Xie, Y., Li, Y., Zhao, Z., Dong, H., Wang, S., Liu, J., Guan, J., & Duan, X. (2020). Microsimulation of electric vehicle energy consumption and driving range. *Applied Energy*, 267(December 2019), 115081. <https://doi.org/10.1016/j.apenergy.2020.115081>

- Xing, Y., Lv, C., Cao, D., & Lu, C. (2020). Energy oriented driving behavior analysis and personalized prediction of vehicle states with joint time series modeling. *Applied Energy*, 261(December 2019), 114471. <https://doi.org/10.1016/j.apenergy.2019.114471>
- Xu, Y., Zheng, Y., & Yang, Y. (2021). On the movement simulations of electric vehicles: A behavioral model-based approach. *Applied Energy*, 283(October 2020), 116356. <https://doi.org/10.1016/j.apenergy.2020.116356>
- Yang, D., Liu, H., Li, M., & Xu, H. (2023). Data-driven analysis of battery electric vehicle energy consumption under real-world temperature conditions. *Journal of Energy Storage*, 72(PD), 108590. <https://doi.org/10.1016/j.est.2023.108590>
- Yang, D.P. et al. (2023). ‘Construction of high-precision driving cycle based on Metropolis-Hastings sampling and genetic algorithm’, *Transportation Research Part D: Transport and Environment*, 118(March), p. 103715. Available at: <https://doi.org/10.1016/j.trd.2023.103715>.
- Yang, S., Wang, W., Zhang, F., Hu, Y., & Xi, J. (2018). Driving-style-oriented adaptive equivalent consumption minimization strategies for HEVs. *IEEE Transactions on Vehicular Technology*, 67(10), 9249–9261. <https://doi.org/10.1109/TVT.2018.2855146>
- Yang, Y., Zhang, X., Zhang, J., & Shang, H. (2021). Deriving environment-friendly speed limit scheme using an integrated traffic simulation framework. *Journal of the Air and Waste Management Association*, 71(5), 633–649. <https://doi.org/10.1080/10962247.2021.1874567>
- Yannis, G., & Michelaraki, E. (2025). Effectiveness of 30 km/h speed limit – A literature review. *Journal of Safety Research*, 92(April 2024), 490–503. <https://doi.org/10.1016/j.jsr.2024.11.003>
- Yavasoglu, H. A., Tetik, Y. E., & Gokce, K. (2019). Implementation of machine learning based real time range estimation method without destination knowledge for BEVs. *Energy*, 172, 1179–1186. <https://doi.org/10.1016/j.energy.2019.02.032>
- Ye, Y., Zhao, X., & Zhang, J. (2023). Driving cycle electrification and comparison. *Transportation Research Part D: Transport and Environment*, 123(August), 103900. <https://doi.org/10.1016/j.trd.2023.103900>
- Yi, Z., & Bauer, P. H. (2017). Effects of environmental factors on electric vehicle energy consumption: A sensitivity analysis. *IET Electrical Systems in Transportation*, 7(1), 3–13. <https://doi.org/10.1049/iet-est.2016.0011>
- Yi, Z., Smart, J., & Shirk, M. (2018). Energy impact evaluation for eco-routing and charging of autonomous electric vehicle fleet: Ambient temperature consideration. *Transportation Research Part C: Emerging Technologies*, 89(October 2017), 344–363. <https://doi.org/10.1016/j.trc.2018.02.018>
- Yu, H., Tseng, F., & McGee, R. (2012). Driving pattern identification for EV range estimation. *2012 IEEE International Electric Vehicle Conference, IEVC 2012*, 1–7. <https://doi.org/10.1109/IEVC.2012.6183207>

- Yu, L., Zhang, X., Qiao, F., & Qi, Y. (2010). Genetic algorithm-based approach to develop driving schedules to evaluate greenhouse gas emissions from light-duty vehicles. *Transportation Research Record*, 2191, 166–173. <https://doi.org/10.3141/2191-21>
- Yu, Z., Shi, S., Wang, S., Mu, Y., Li, W., Xu, C., & Zhang, M. (2017). Statistical inference-based research on sampling time of vehicle driving cycle experiments. *Transportation Research Part D: Transport and Environment*, 54, 114–141. <https://doi.org/10.1016/j.trd.2017.04.029>
- Yusof, N., Abas, M. A., Rohana, N. A., & Mohd Perang, M. R. (2023). Route selection for road vehicle real-driving emissions test using GIS. *IOP Conference Series: Earth and Environmental Science*, 1143(1). <https://doi.org/10.1088/1755-1315/1143/1/012027>
- ZenSung. (2025). *Mobile based Telematics App*. <https://www.zensung.com/>
- Zhang, J., Wang, Z., Liu, P., Zhang, Z., Li, X., & Qu, C. (2019). Driving cycles construction for electric vehicles considering road environment: A case study in Beijing. *Applied Energy*, 253(5), 113514. <https://doi.org/10.1016/j.apenergy.2019.113514>
- Zhang, J., Wang, Z., Liu, P., & Zhang, Z. (2020A). Energy consumption analysis and prediction of electric vehicles based on real-world driving data. *Applied Energy*, 275(5), 115408. <https://doi.org/10.1016/j.apenergy.2020.115408>
- Zhang, J., Yan, J., Liu, Y., Zhang, H., & Lv, G. (2020B). Daily electric vehicle charging load profiles considering demographics of vehicle users. *Applied Energy*, 274(June), 115063. <https://doi.org/10.1016/j.apenergy.2020.115063>
- Zhang, L., Huang, Z., Yu, F., Liao, S., Luo, H., Zhong, Z., Zhu, M., Li, Z., Cui, X., Yan, M., & Zheng, J. (2021). Road type-based driving cycle development and application to estimate vehicle emissions for passenger cars in Guangzhou. *Atmospheric Pollution Research*, 12(8), 101138. <https://doi.org/10.1016/j.apr.2021.101138>
- Zhang, L., Zhu, Z., Zhang, Z., Song, G., Zhai, Z., & Yu, L. (2022). An improved method for evaluating eco-driving behavior based-on speed-specific vehicle-specific power distributions. *Transportation Research Part D: Transport and Environment*, 113, 103476.
- Zhang, S., Fatih, D., Abdulqadir, F., Schwarz, T., & Ma, X. (2022). Extended vehicle energy dataset (eVED): an enhanced large-scale dataset for deep learning on vehicle trip energy consumption. *ArXiv*, 1–16. <http://arxiv.org/abs/2203.08630>
- Zhang, S., Li, T. and Chen, L. (2023) ‘Fuzzy Logic Control of External Heating System for Electric Vehicle Batteries at Low Temperature’, *World Electric Vehicle Journal*, 14(4), pp. 1–18. Available at: <https://doi.org/10.3390/wevj14040099>.
- Zhang, X., Zhao, D. J., & Shen, J. M. (2011). A synthesis of methodologies and practices for developing driving cycles. *Energy Procedia*, 16(PART C), 1868–1873. <https://doi.org/10.1016/j.egypro.2012.01.286>

- Zheng, F., Li, J., Van Zuylen, H. J., & Lu, C. (2019). Influence of driver characteristics on emissions and fuel consumption. *IET Intelligent Transport Systems*, 13(12), 1770–1779.
- Zhao, L., Yao, W., Wang, Y., & Hu, J. (2020). Machine Learning-Based Method for Remaining Range Prediction of Electric Vehicles. *IEEE Access*, 8, 212423–212441. <https://doi.org/10.1109/ACCESS.2020.3039815>
- Zhao, X., Ma, J., Wang, S., Ye, Y., Wu, Y., & Yu, M. (2019). Developing an electric vehicle urban driving cycle to study differences in energy consumption. *Environmental Science and Pollution Research*, 26(14), 13839–13853. <https://doi.org/10.1007/s11356-018-3541-6>
- Zhao, X., Ye, Y., Ma, J., Shi, P., & Chen, H. (2020A). Construction of electric vehicle driving cycle for studying electric vehicle energy consumption and equivalent emissions. *Environmental Science and Pollution Research*, 27(30), 37395–37409. <https://doi.org/10.1007/s11356-020-09094-4>
- Zhao, X., Zhao, X., Yu, Q., Ye, Y., & Yu, M. (2020; 2020B). Development of a representative urban driving cycle construction methodology for electric vehicles: A case study in Xi'an. *Transportation Research Part D: Transport and Environment*, 81(February), 102279. <https://doi.org/10.1016/j.trd.2020.102279>
- Zhou, L., Lai, X., Li, B., Yao, Y., Yuan, M., Weng, J., & Zheng, Y. (2023). State Estimation Models of Lithium-Ion Batteries for Battery Management System: Status, Challenges, and Future Trends. *Batteries*, 9(2). <https://doi.org/10.3390/batteries9020131>
- Zhou, M., Jin, H., & Wang, W. (2016A). A review of vehicle fuel consumption models to evaluate eco-driving and eco-routing. *Transportation Research Part D: Transport and Environment*, 49(5), 203–218. <https://doi.org/10.1016/j.trd.2016.09.008>
- Zhou, S., Xu, Z., & Liu, F. (2016B). Method for determining the optimal number of clusters based on agglomerative hierarchical clustering. *IEEE transactions on neural networks and learning systems*, 28(12), 3007–3017.
- Ziegmann, J., Denk, F., Vogele, U., & Endisch, C. (2018). Stochastic Driver Speed Prediction with Environmental Features on Naturalistic Driving Data. *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC, 2018-Novem*, 1807–1814. <https://doi.org/10.1109/ITSC.2018.8569767>

Appendix A. Summary of Previous DCs

The abbreviations for the CPs in Table 1 are as follows: Acceleration (A), Deceleration (D), Cruise (C), Idle (I), Mean (M), Maximum (Max), Minimum (Min), Standard Deviation (S), Time (T), Speed (V), Number (No), Root Mean Square (RMS), Positive Kinetic Energy (PKE)

Study and Location	Data Collection Method; Equipment	Vehicle(s); Designated Routes	Database Size	Construction Method	Characteristic Parameters	Validation Threshold %; DC Length
Settey et al. 2021 Žilina	Not Stated	VW Golf 1.9 TDI; 7	Not Stated	Weighted Arithmetic Mean	Total T; Driving T; % T I, C, A, D; intervals of V; No stops	None, 1 DC produced; 1546s
Wang et al. 2019 Beijing	From national database; None	Unspecified EVs; 1	30,000 MTs	Markov Chain and Monte Carlo	% T I, C, A, D; M, Max, Min, and S V and A	None, 1 DC produced; 1200s
Nyberg et al. 2014 Sweden	Instrumented Vehicle; Not Stated	Not Stated; None	466 trips	MC	Mean Tractive Force	Not Stated; Not Stated
Kamble et al. 2009 Pune	Chase car; GPS	Not Stated; 5	46 MTs	MT-parameter based	T; M V	Not Stated; 1533s
Komorska et al. 2021 Warsaw	Instrumented Vehicle; On-board GPS	Tesla Model X; 1	2 trips	MCMC over segments, TPM between distance not time	Total T; M V, V>0, A, and D; S V, A, and D; Max A and D; % T I; V x A in brackets; M energy consumption	None, 1 DC produced; Not Stated
Wang et al. 2021 Xi'an	Combination Chase Car and Instrumented Vehicle; GPS, OBD, and Racelogic VBOX	BYD e6; 1	1414 MTs	K-Means and Random Forest Clustering and Kernel PCA	Total T; T of I, C, A, and D; M, Max, and S V, A, and D; M V>0	10; 1200s
Jing et al. 2022 Zhenjiang	Instrumented Vehicle; CATC terminal recording OBD and GPS at 1hz	Not Stated; None	1162 MTs	Wavelet domain denoising, hierarchical clustering	3 PCA from 15: M and Max V,A, and D; % T I, C, A, and D; T of I, C, A, and D; Total T;	None; 1200s
Lin and Niemeier 2002 Los Angeles	Chase car from previous DC; None	Not Stated; 102 (not designated)	833 MTs	MC of modal events (I, C, A, D)	None	Lowest from candidate cycles; 1381s

Study and Location	Data Collection Method; Equipment	Vehicle(s); Designated Routes	Database Size	Construction Method	Characteristic Parameters	Validation Threshold %; DC Length
Brady and O'Mahony 2016 Dublin	Instrumented Vehicle; GPS and CAN data logger	Mitsubishi iMiEV; None	1485 trips	Neural network and K-means clustering; MC	M and S V; M, Max, and Min A; % T I and C; No stops and stops/km; M specific power; T of C	10; 1650s
Huertas et al. 2019 Mexico City	Instrumented Vehicle; Garmin 16x GPS	Diesel Buses; None	138 trips	3 Different: quasi-random MT; MC; fuel-based	M V, % T I	5; N/A
Amirjamshidi and Roorda 2015 Pune	Microsimulation; None	None; None	Not Stated	MT-parameter based	M V, A, D, and V>0; % T I, C, A, D, and Creep (0<V,4 kmph); M MT duration; M % A-D changes; RMS A; RMS PKE/Weight	15; N/A
Arun et al. 2017 Chennai	Instrumented Vehicle; OBD and Garmin etrex 10 GPS	Tata Indica; 1	756 MTS	MT-parameter based	M V, A, D, and V>0; % T I, C, A, D, and Creep (0<V<5kmph); RMS A; RMS PKE	5; 1448s
Mayakuntla and Verma, 2018 Bangalore	Instrumented Vehicle; VBOX	Not Stated; 2	250km/1 8hrs	Segment based, K-means clustering	M V and MT duration; PKE; RMS A; % T C and Creep	Lowest from candidate cycles; 2088s
Berzi et al. 2016 Florence	Instrumented Vehicle; OBD with GPS for location analysis	Renault Kangoo ZE, Renault Twizy, Peugeot iOn, Citroen C-Zero; None	2500km	Kernal smoothing; K-means into 9 clusters; random MT	Total T; Distance; % T I, C, A, and D; M V, A, D and V>0; RMS V; PKE; Relative A; Stop Rate; SAFD edges;	5 for 11/12 values; N/A

Study and Location	Data Collection Method; Equipment	Vehicle(s); Designated Routes	Database Size	Construction Method	Characteristic Parameters	Validation Threshold %; DC Length
Nouri and Morency 2017	Instrumented Vehicle; OBD and GPS	Toyota Yaris; 1	45 trips/35hrs	Define MTs in 6 different ways; k-means into 7 clusters based on SAFD; MC	Difference in SAFD; M V, A, D, and V>0; Max V; M No of A-D changes; RMS A; road power; % T I, C, A, D, and creep	Lowest from candidate cycles; 2760s
Huertas et al. 2018 Mexico City	Instrumented Vehicle; GPS validated by OBD	Diesel Buses; 1	1100 MTs	new MWD-CP method that takes energy consumption into account but only uses speed data constructing DC	V on hilly terrain, T I; road grade, aerodynamic forces; M and Max A; No stops, and Max altitude	Lowest from candidate cycles; 10000s
Yu et al. 2010 Houston	Instrumented Vehicle; GPS and OEM-2100 OBD	7 gas cars; None	264 MTs/425.31mi/24.33hrs	Genetic algorithm improved MT	% T I, C, A, and D; M V, A, D, and V>0; Max V	Lowest from candidate cycles; 600-1200s
Bishop et al. 2012 Oxford	Instrumented Vehicle; Globalsat DG-100 GPS	10 electric scooters; None	167 Trips/1180km	MC	M V and V>0; No A-D changes/km; M and S specific power; % I, C, A, and D	5, 10 for analysis purposes; 500s
Waldowski et al. 2011 Berlin	Instrumented Vehicle; TU-veLOG GPS	Not Stated; None	Not Stated	MT-parameter based	M V and A; stop rate; stop frequency	Not Stated; N/A
Ashtari et al. 2014 Winnipeg	Instrumented Vehicle; PERSENTECH GPS	76 vehicles; None	44 million data points	K-means into 6 clusters; 2-step MC to account for previous MTs	Diff in SAFD; M V, A, D, V >0, and No A-D changes; Max V; RMS A; road power; % T I, C, A, D, and creep	None, 1 DC produced; 1200s

Study and Location	Data Collection Method; Equipment	Vehicle(s); Designated Routes	Database Size	Construction Method	Characteristic Parameters	Validation Threshold %; DC Length
Pereira et al. 2020 Recife	Instrumented Vehicle; Phone GPS	Hyundai HB20, Volkswagen Golf, Fiat Argo, and motorcycles; 1	405 MTs/240 km	MT and Markov to compare	M V, A, D, and V>0; S V; % T I, C, A, and D	4; 1800s
Tsai et al. 2005 Kaohsiung	Chase Car; Frequency-voltage transducer	Motorcycle; 4	316 MTs	randomly add 10 MTs 40 times	M V, A, D, V>0, No A-D changes, and MT duration; % T I, C, A, and D; RMS A	Lowest from candidate cycles; 111s
Zhang et al. 2019 Beijing	Instrumented Vehicle; National Database	40 electric taxis; None	750,000 km	MT and MC to compare; k-means into 3 and random MT; MC matching road conditions	T V>0; Dist; % T I, C, A, and D; M V, A, D, V>0, pos and neg grade; Max and S V; 95 th % V, A, and D; S grade; % pos, neg, and 0 grade; % regeneration energy	Lowest from candidate cycles; 1500s
Zhao et al. 2020 Xi'an	Combination Chase Car and Instrumented Vehicle; GPS, OBD, and Racelogic VBOX	BYD e6; 1	27,866 kinematic segments	Wavelet denoising; kinematic segments clustered with K-means and SVM hybrid algorithms into 6; MC	4 PCA from M, Max, Min, and S V; M, Max, and S A; Max D; speed difference; running time	Lowest from candidate cycles; 1200s
Chen et al. 2019 Shenyang	Chase Car; P-box GPS	Not Stated; 2	393km	Divide MTs by speed into 3; genetic algorithm	M distance of V, V ³ , A x V, and grade x V; all 4 over A and D periods; RMSE SAFD; AEC	Lowest from candidate cycles; 2984s
Lee et al. 2011 Ann Arbor	Instrumented Vehicle; GPS	Not Stated; None	181 trips	Divide trips by length; synthesise DC using MC	S V and A; M A and V>0; % T I, A, and D; No stops/mi;	None, 1 DC produced; N/A

Study and Location	Data Collection Method; Equipment	Vehicle(s); Designated Routes	Database Size	Construction Method	Characteristic Parameters	Validation Threshold %; DC Length
Günther et al. 2017 Hamburg	Instrumented Vehicle; Original datalogger	12 buses; None	836,649 MTs/315, 583km/2 7,365hra	2 MT methods	M V and V during C; 1st and 2nd A and D rate; peak time; T of C and I	Lowest from candidate cycles; 1200s
Guo and Zhang 2016 Beijing	Instrumented Vehicle; GPS and CAN	E150EV EV200 Midi series and BYD e6; 2	56,801 MTs	Random MT with PCA and clustering	Total T; T of I, C, A, and D; Max V, A, and D; S V and A; M V, A, D, and V>0; % T I, C, A, and D; V distribution brackets	Not Stated; 1150s
Pan et al. 2015 Beijing	Instrumented Vehicle; GPS and INS OXTS Inertial+	Hybrid Bus; Not Stated	Not Stated	MCMC, divide fragments into I, C, a, and D; cluster by speed into 7;	% T I, C, A, and D; M, Max, and S V; Max, Min, and S A; M, Max, and Min grade and road power	Lowest from candidate cycles; 1200s
Ho et al. 2014 Singapore	Combination Chase Car and Instrumented Vehicle; OBD	Not Stated; 12	2237km/ 42.5hrs	Categorise MT using GPS; select best MT	M A, D, and V>0; % T C, A, and D	10; 2344s
Jing et al. 2017 Tianjin	Instrumented Vehicle; CAN	BAIC; Yes	674 MTs	linear discrimination with PCA; quasi-random MT method	4 PCA from M and Max V, A, and D; Total T; T and %T I, C, A, and D	10; 1050s
Gong et al. 2018 Beijing	Instrumented Vehicle; CAN	10 Evs; None	3885 trips/ 45384km	Regression analysis on energy consumption; 2-level MC for state then speed	S V and A; M A and D; % T I, A, and D; No stops/km; Max and Min specific power	10; 2536s
Romero et al. 2017 Pereira	Instrumented Vehicle; ELM327 OBD with Torque app and Garmin GPS for elevation	4 ice cars; 2	Not Stated	MT binned by average speed; MTs with highest correlation to CP of bin added	M V; % T I, C, A, and D	Lowest from candidate cycles; 1100s

Study and Location	Data Collection Method; Equipment	Vehicle(s); Designated Routes	Database Size	Construction Method	Characteristic Parameters	Validation Threshold %; DC Length
Han et al. 2012 Korean DMZ	Combination Chase Car and Instrumented Vehicle; GPS and sensors	Military Vehicle; 16	430 MTs/245 km	Random MT	M and S V and grade	Lowest from candidate cycles; 1436s
Ali and Asfoor 2023 Cairo	Instrumented Vehicle; Not Stated	Not Stated; 8	6000 trips	Multivariate K-means clustering for V, dist, and time of day; DC made by M V every 10m	M and S V; distance	None, 1 DC produced; N/A
Salihu et al. 2023 Pristina	Instrumented Vehicle; Canedge1	Not Stated; 2	525 MTs	K-means clustering and random MT	% T I, C, A, and D; total T; M and S V; grade;	None; N/A
Yang et al. 2023 Qingdao	Instrumented Vehicle; T-Box	Hybrid Trucks; None	100,000k m	ST-DBSCAN and Epanechnikov kernel denoising; random state transfer by Metropolis-Hastings sampling; genetic algorithm to optimise DCs	Linear Regression on fuel consumption to get M V, A, D, and V>0; S V; % T I, D, C, various V brackets; S A; No stops/km; M driving power (KW)	5; 1600s
Wang et al. 2023 Nantong	Instrumented Vehicle; Bus operations platform	Electric Bus; 1	1870 MTs	K-means into 2 clusters; optimal MTs spliced	3 PCA from M, Max, and S V; M A, D, and V>0; Max A and D; % T I, C, A, and D	None, 1 DC produced; 1800s

Appendix B. Empirical Method Applied to Route DCs

Table B.1. Average Statistics by EC Group: Dub-Don Route

Parameter	Energy Consumption		
	Low	Moderate	High
Duration (Sec)	10620	10611	10590
Average Speed (km/h)	78.38	78.46	78.62
Std Dev Speed (km/h)	2.51	2.46	2.50
Average Acceleration (m/s ²)	0.32	0.32	0.33
Maximum Acceleration (m/s ²)	3.37	3.70	3.65
Std Dev Acceleration (m/s ²)	0.13	0.13	0.14
Average Deceleration (m/s ²)	-0.36	-0.36	-0.36
Maximum Deceleration (m/s ²)	-3.99	-3.58	-4.13
Std Dev Deceleration (m/s ²)	0.15	0.15	0.15
Cruise Time (%)	66.80	66.74	66.48
Acceleration Time (%)	17.62	17.60	17.66
Deceleration Time (%)	15.58	15.65	15.86
Energy Consumption (kWh)	33.32	33.58	33.92

Table B.2. Average Statistics by EC Group: Dub-Gal Route

Parameter	Energy Consumption		
	Low	Moderate	High
Duration (Sec)	8011	7993	7981
Average Speed (km/h)	94.45	94.69	94.83
Std Dev Speed (km/h)	1.79	1.75	1.80
Average Acceleration (m/s ²)	0.32	0.32	0.31
Maximum Acceleration (m/s ²)	3.23	3.40	3.40
Std Dev Acceleration (m/s ²)	0.11	0.11	0.11
Average Deceleration (m/s ²)	-0.34	-0.34	-0.34
Maximum Deceleration (m/s ²)	-3.75	-4.01	-3.85
Std Dev Deceleration (m/s ²)	0.12	0.12	0.12
Cruise Time (%)	69.44	70.10	70.20
Acceleration Time (%)	15.79	15.53	15.56
Deceleration Time (%)	14.77	14.37	14.24
Energy Consumption (kWh)	37.15	37.37	37.63

Table B.1. Average Statistics by EC Group: Dub2 Route

Parameter	Energy Consumption		
	Low	Moderate	High
Duration (Sec)	1142	1140	1146
Average Speed (km/h)	23.25	23.29	23.19
Std Dev Speed (km/h)	5.15	4.98	5.11
Average Acceleration (m/s ²)	0.74	0.73	0.74
Maximum Acceleration (m/s ²)	3.13	3.06	3.68
Std Dev Acceleration (m/s ²)	0.49	0.49	0.51
Average Deceleration (m/s ²)	-0.77	-0.77	-0.77
Maximum Deceleration (m/s ²)	-3.47	-3.40	-3.19
Std Dev Deceleration (m/s ²)	0.54	0.52	0.52
Cruise Time (%)	23.11	22.79	22.30
Acceleration Time (%)	39.36	39.49	39.73
Deceleration Time (%)	37.52	37.72	37.97
Energy Consumption (kWh)	0.92	0.97	1.03

Table B.1. Average Statistics by EC Group: Dub3 Route

Parameter	Energy Consumption		
	Low	Moderate	High
Duration (Sec)	1162	1164	1165
Average Speed (km/h)	25.87	25.86	25.82
Std Dev Speed (km/h)	4.96	4.71	4.78
Average Acceleration (m/s ²)	0.72	0.70	0.72
Maximum Acceleration (m/s ²)	3.12	3.24	3.44
Std Dev Acceleration (m/s ²)	0.48	0.47	0.49
Average Deceleration (m/s ²)	-0.72	-0.73	-0.75
Maximum Deceleration (m/s ²)	-3.72	-3.15	-3.61
Std Dev Deceleration (m/s ²)	0.51	0.49	0.51
Cruise Time (%)	23.71	25.01	24.55
Acceleration Time (%)	38.22	38.35	38.77
Deceleration Time (%)	38.07	36.64	36.67
Energy Consumption (kWh)	1.02	1.07	1.11

Appendix C. Summary of EC Models

Appendix C provides additional reference information for studies which build EC models in the literature review of Chapter 3.

Study	Target Variable	Predictor Variables	Methods	Model Scale
Liu et al. (2018)	kWh/km	temp, HVAC, speed, elev	OLS	Macroscopic
Lee et al. (2024)	km/kWh	speed, battery, motor torque, temp	Dynamometer	Microscopic
Yi et al. (2018)	kWh/mi	temp, route choice, average speed	OLS, grid stochastic	Macroscopic
Desrevaux et al. (2020)	kWh/100km	temp, resistive forces	Simulation	Microscopic
Yang et al. (2023)	kWh/100km	temp, HVAC, PCA from speed and accel	PCA, correlation	Macroscopic
Hariharan et al. (2022)	kWh, SOC	temp, resistive forces	Simulation	Microscopic
Abediasl et al. (2023)	kWh/km	temp, resistive forces	Dynamometer	Microscopic
Modi and Bhattacharya (2022)	MJ	temp, speed, accel, grade, HVAC, wind speed	ANN, XGBoost, SVR, kNN, LASSO	Microscopic
Vepsäläinen et al. (2019)	kWh/km	mass, wind speed and direction, temp, resistive forces, battery	Simulation	Microscopic
Xie et al. (2020)	kWh/km	battery, speed	Dynamometer	Microscopic
Felipe et al. (2015)	%SOC/km	speed, road type, pedal, elev	ANN	Mesosopic
Lee and Wu (2015)	SOC	battery, HVAC, resistive forces, driving behaviour	Simulation	Microscopic
Ahmed et al. (2022)	Range	weight, vehicle information	Regression	Macroscopic
Asamer et al. (2016)	kWh/km	resistive forces	Variance cutting	Macroscopic
Jaworski et al. (2023)	kWh; kWh/km	speed, weight	Physics Model	Microscopic
Settey et al. (2021)	kW	speed, weight	Physics Model	Microscopic
Al-Wreikat et al. (2021)	kWh/km	traffic, driving behaviour, temp, distance	Empirical	Macroscopic
Janpoom et al. (2023)	Wh/km	speed, weight, grade, battery	XGBoost, RF, SVR, KRR	Microscopic
Morlock et al. (2020)	kW	battery, traffic	Empirical	Microscopic
Wei et al. (2023)	SOC	battery	Physics Model; LSTM	Microscopic
Madziel and Campisi (2023)	Wh/km	speed, vehicle information	Correlation	Macroscopic
Galvin (2017)	W; kWh/km	speed, acceleration	Physics Model	Microscopic
Mavlonov et al. (2023)	SOC	speed, motor torque	Physics Model	Microscopic
Agrawal et al. (2023)	kW	battery, temp, resistive forces	Physics Model	Microscopic
Miri et al. (2021)	Wh/km	resistive forces	Physics Model	Microscopic
Alhanouti and Gauterin (2024)	kW	speed, battery	Simulation	Microscopic
Mediouni et al. (2022)	W	resistive forces	Physics Model	Microscopic
Rhode et al. (2020)	J	resistive forces	ANN, regression	Microscopic
Sasank et al. (2023)	SOC	battery, DC	Simulation	Microscopic

De Cauwer et al. (2017)	kWh	resistive forces	Regression	Mesosopic
Sun et al. (2019)	SOC, range	speed, temp	Regression tree	Macroscopic
Ekici et al. (2023)	kW	speed, weather, road	SVM, ANN, decision tree, regression	Macroscopic
Marzbnai et al. (2023)	EC	battery	Regression, ANN, SVR, KNN, decision tree, RF	Macroscopic
Sun and Fu (2024)	kWh/100km	resistive forces	XGBoost, LightGBM, genetic	Microscopic

Appendix D. Summary of Driving Behaviour Studies

Study	Number of Groups	Methods	Parameters	Driving/Driver Behaviour
Fafoutellis et al. (2021)	Review	Many	Many	Driving
Ma et al. (2024)	Review	Many	Many	Driving
Faria et al. (2017)	9 (scenario based)	Statistical grouping	VSP	Driving
Chu et al. (2023)	Review	Many	Many	Both
Iclodean et al. (2025)	3	Simulation presets	acceleration, deceleration	Driving
Donkers et al. (2020)	3	Simulation presets	acceleration, deceleration, speed	Driving
Ellison et al. (2015)	None	Simulation presets	speeding, acceleration, braking	Driver mostly
Carmona et al. (2015)	2; additionally split by urban and inter-urban	Data fusion	speed, acceleration, rpm, steering wheel angle, brake frequency, throttle	Driver
Wu et al. (2023)	3	Clustering, SVM	speed, acceleration, jerk	Driving
Yu et al. (2012)	7 (scenario based)	Clustering	speed, gradient	Driving
Hu et al. (2022)	8 (scenario based)	Clustering	speed, acceleration, deceleration, throttle, brake pedal	Driving
Liu et al. (2022)	N/A	Clustering	speed, acceleration	Driving
Ping et al. (2019)	3	Clustering	speed, acceleration	Driving
Bar et al. (2011)	5	Fuzzy-rule based	speed, acceleration, deceleration	Driving
Yang et al. (2018)	6	Kernel density	speed, acceleration, throttle	Driving
Sun et al. (2017)	3	Markov model	speed, acceleration	Driving
Cheng et al. (2018)	3	ANN	speed, acceleration, throttle	Driving
Huang et al. (2020)	4	Clustering	speed, throttle, brake	Driving
Zhang et al. (2020A)	4	Threshold	acceleration, deceleration	Driving
Xing et al. (2020)	3	Empirical	speed, acceleration, jerk	Driving
Al-Wreikat et al. (2021)	3	Threshold	RPA	Driving
Nan et al. (2022)	N/A	LSTM	speed, acceleration, throttle, brake	Driving

Appendix E. Summary of Sensitivity Analysis Studies

Study	GSA or LSA	OATSA?	Method	Sampling Method
Tsvetkova and Ouarda (2021)	Review	-	Many	Many
Razavi and Gupta (2016)	Review	-	Many	Star-based
Razavi et al. (2021)	Review	-	Many	Many
Braband et al. (2022)	GSA	Yes	Morris	MC
Wei (2013)	Review	-	Many	LHS
Ricotto and Zio (1999)	LSA	Yes	ANN	MC
Abburu et al. (2024)	GSA	No	Morris, variance	MC,LHS,Sobol'
Ben-Chaim et al. (2013)	GSA	No	Derivative	N/A
Kioutsioukis et al. (2004)	GSA	No	Morris	LHS
Owen (2014)	GSA	No	Sobol', Shapley	Sobol'
Song et al. (2016)	GSA	No	Shapley	MC
Pianosi and Wagener (2015)	GSA	No	PAWN	Sobol'
Botshekan et al. (2019)	GSA	No	Sobol'	MC
Lai et al. (2021)	GSA	No	Sobol'	MC
Vepsäläinen et al. (2019)	GSA	No	Sobol'	Sobol'
Gawlikowski et al. (2023)	Review	No	Many	Many
Guimarães et al. (2019)	LSA	Yes	Comparison	Chosen Samples
Iora and Tribioli (2019)	LSA	Yes	Comparison	Chosen Samples
Ji et al. (2022)	GSA	Yes	Comparison	Chosen Samples
Miri et al. (2021)	LSA	Yes	Comparison	Chosen Samples
Yi and Bauer (2017)	GSA	Yes	Comparison	Chosen Samples
Schäfers et al. (2024)	GSA	No	Spearman	MC
Huang et al. (2022)	GSA	No	Genetic Algorithm	LHS
Pal et al. (2017)	GSA	No	Comparison	LHS
Tran et al. (2024)	GSA	Yes	Comparison	Chosen Samples
Asamer et al. (2016)	GSA	Yes	Sobol'	Chosen Samples