Optimised charging scheduling and predictive modelling to reduce emissions from battery electric buses

Padraigh John William Jarvis

Thesis submitted for the degree of Doctor of Philosophy

NATIONAL UNIVERSITY OF IRELAND, CORK

COLLEGE OF SCIENCE, ENGINEERING AND FOOD SCIENCE

SCHOOL OF COMPUTER SCIENCE & INFORMATION TECHNOLOGY
INSIGHT CENTRE FOR DATA ANALYTICS

July 2025

Head of School: Professor Dirk Pesch

Supervisors: Dr. Alejandro Arbelaez
Dr. Laura Climent
Professor. Kenneth Brown

Contents

List of Figures

List of Tables

I, Padraigh John William Jarvis, certify that this thesis is my own work and has not been submitted for another degree at University College Cork or elsewhere.

Padraigh John William Jarvis

To my Dad, Mom, and Grandmother, who have sacrificed for my education.

Acknowledgements

I would first like to thank my Dad, Mom, and Grandmother for the support they have shown me over the years that has helped me become the person I am today. Additionally, I want to thank my extended family.

I would like to extend a special thank you to all of my friends who have been there for me over the years. Especially to Andi, Marlon, Vedat, and Larissa, who have been so supportive throughout my PhD journey.

I would like to thank Andrea Visentin for helping to organise the Computer Science Postgraduate Student Committee. Being able to chair this committee taught me a lot about myself and helped me give back to my fellow students. Furthermore, I would like to thank the members of the committee, Martha, Japneet, Suheyl, and Ege. Without them, projects such as the coffee mornings, games nights, and welcome booklet would have never been finished.

I would also like to thank the Insight administrative staff, Linda, Caitríona, and Eleanor, as well as James Duggan, for their support in all my administrative needs over the years. I also want to give a special thanks to Janet, whose support has been instrumental to my success as a PhD student.

A special thanks to my supervisory team, Alejandro Arbalaez, Ken Brown, and Laura Climent, as well as my advisor, Luis Quesada. They have been a great help over these last few years, and I have learned a lot from them. Without them, this thesis would not have been possible.

Works undertaken in this thesis were initially funded by the Sustainable Energy Authority of Ireland (SEAI) Research, Development & Demonstration under grant number 19/RDD/519. Additionaly financial support was provided by Research Ireland under Grant number 12/RC/2289-P2 which is co-funded under the European Regional Development Fund.

Abstract

Battery Electric Vehicles provide an opportunity to decarbonise the transportation sector, which has proven to be a large contributor to Greenhouse Gases. This decarbonisation proves even more effective for Battery Electric Buses, as public transport services expand worldwide to help mitigate climate change. However, while Battery Electric Buses do not produce emissions themselves, the electricity used to power such vehicles may come from polluting sources, such as diesel or natural gas, resulting in lifecycle emissions. Reduction of these lifecycle emissions can aid in meeting carbon emission targets set by countries across the world.

The Battery Electric Bus Charging Schedule Problem aims to create charging schedules for Battery Electric Bus to enable the continued operation of established bus routes while considering the restrictions tied to such vehicles, such as a lower operational range. The focus of this thesis is the modelling of this problem to prioritise charging of Battery Electric Buses using low-emission electricity, thus lowering lifecycle emissions. Furthermore, the impact on service quality is also considered to reduce any potential delays caused by charging.

The first contribution made by this thesis is the evaluation of Deep Learning predictive model configurations for imbalanced regression problems to estimate the availability of low-emission electricity. Six Deep Learning models are considered, along with three resampling strategies and three loss metrics. The resulting 54 configurations are empirically evaluated using multiple test metrics. It was determined that Convolution Neural Networks trained with a Squared Error Relevance Area loss function performed best when estimating the availability of low-emission energy, such as excess wind-generated electricity. However, the same configuration using an Inverse-Weighted Mean Squared Error loss function appears to perform well in estimating extremes in excess wind-generated electricity.

The second contribution creates a Battery Electric Bus Charging Schedule Problem to reduce the consumption of non-low-emission energy, utilising predictive models. The performance of the models was compared to a naïve scenario, and an ideal scenario with 100% accurate information, low-emission energy information. Use of the predictive models showed improvement of up to 10.04% compared to the naïve scenario. Furthermore, there was only an average discrepancy of 4.89% between the predictive and ideal scenarios.

The final contribution made in this dissertation is the extension of the previously developed Battery Electric Bus Charging Schedule Problem to consider potential negative service impact. The problem aims to reduce the daily operational costs of a bus fleet, considering fuel costs, carbon emissions, and passengers' Value of Time expressed as a financial value. Furthermore, a combination of aspects which have not been modelled before for a Battery Electric Bus Charging Schedule Problem is used to accurately represent the problem. Evaluation on a large dataset with over 1000 buses showed that carbon emissions can be reduced by up to 64%, with savings in fuel costs reaching 50% per day, while only inducing an average delay of 8.8 seconds.

Acronyms

ANN Artificial Neural Network.

ARIMA AutoRegressive Integrated Moving Average.

BEB Battery Electric Bus.

BEB-CSP Battery Electric Bus Charging Scheduling Problem.

BEMUs Battery Electric Multiple Units.

BESS Battery Energy Storage System.

BEV Battery Electric Vehicle.

CC Constant Current.

CNN Convolutional Neural Network.

CO_{2eq} Carbon Dioxide Equivalent.

CO_2 Carbon Dioxide.

COP Constrained Optimisation Problem.

CPU Central Processing Unit.

CSLP Charging Station Location Problem.

CSP Constraint Satisfaction Problem.

CUDA Compute Unified Device Architecture.

cuDNN CUDA Deep Neural Network.

CV Constant Voltage.

DICE Dynamic Integrated model of Climate and the Economy.

DL Deep Learning.

DOD Depth of Discharge.

DT Decision Trees.

E-VRP Electric Vehicle Routing Problem.

E-VSP Electric Vehicle Scheduling Problem.

EoL End-of-Life.

EV Electric Vehicle.

FRVCP-BEV Fixed Route Vehicle Charging Problem for Battery Electric Vehicles.

G-VRP Green Vehicle Routing Problem.

GA Genetic Algorithm.

gCO_{2eq}/km Grams of Carbon Dioxide Equivalent per Kilometre.

gCO_{2eq}/kWh Grams of Carbon Dioxide Equivalent per Kilowatt hour.

gCO_2/kWh grams of Carbon Dioxide per Kilowatt hour.

GDP Gross Domestic Profit.

GHG Greenhouse Gases.

GHz Gigahertz.

GN Gaussian Noise.

GPS Global Positioning System.

GPU Graphics Processing Unit.

GRU Gated Recurrent Unit.

HEV-TSP Hybrid Electric Vehicle Travelling Salesman Problem.

HPO Hyperparameter Optimisation.

ICE Internal Combustion Engine.

ICEB Internal Combustion Engine Bus.

ILP Integer Linear Programming.

IP Integer Programming.

IW-MSE Inverse Weighted Mean Squared Error.

kg/CO_{2eq} Kilograms of Carbon Dioxide Equivalent.

km Kilometre.

km/h Kilometres per hour.

kWh Kilowatt hour.

kWh/km Kilowatt hours per kilometre.

LP Linear Programming.

LS Local Search.

LSTM Long Short-Term Memory.

MAE Mean Absolute Error.

MAPE Mean Absolute Percentage Error.

MD-VSP Multiple Depot Vehicle Scheduling Problem.

MILP Mixed Integer Linear Programming.

MIP Mixed Integer Programming.

MLP Multilayer Perceptrons.

MP Mathematical Programming.

MPM Multi-Period reactive Model.

MSE Mean Squared Error.

MtCO_{2eq} Million Tonnes of Carbon Dioxide Equivalent.

MVT-VSP Multiple Vehicle Type Vehicle Scheduling Problem.

MW MegaWatt.

MWh Megawatt hour.

MWh/km Megawatt hours per Kilometre.

NP Nondeterministic Polynomial time.

OC Opportunity Charging.

PCHIP Piecewise Cubic Hermite Interpolating Polynomials.

PSO Particle Swarm Optimisation.

RAM Random Access Memory.

ReLu Rectified Linear Unit.

RF Random Forest.

RMSE Root Mean Squared Error.

RNN Recurrent Nerual Network.

SARIMA Seasonal AutoRegressive Integrated Moving Average.

SCC Social Cost of Carbon.

SD-VSP Single Depot Vehicle Scheduling Problem.

SDK Software Development Kit.

SEMO Single Energy Market Operator.

SER Squared Error Relevance.

SERA Squared Error Relevance Area.

SMOTE Synthetic Minority Over-Sampling Technique.

SMOTER Synthetic Minority Over-Sampling Technique for Regression.

SOC State of Charge.

SPC Shadow Price of Carbon.

SPM Single-Period reactive Model.

SSE Sum of Squared Error.

STD Standard Deviation.

tCO_{2eq} Tonne of Carbon Dioxide Equivalent.

TOU Time-of-Use.

TRU Temporal Relevance Undersampling.

TSP Travelling Salesman Problem.

TWA-VRP Time Window Assignment Vehicle Routing Problem.

V2G Vehicle-to-Grid.

VOT Value-of-Time.

VRP Vehicle Routing Problem.

VRP-TW Vehicle Routing Problem with Time Windows.

VSP Vehicle Scheduling Problem.

VSP-TW Vehicle Scheduling Problem with Time Windows.

Chapter 1

Introduction

1.1 Motivation

Climate change (also known as global warming) is currently on track to reduce global Gross Domestic Profit (GDP), increase the mortality rates, and, in more catastrophic scenarios, cause the extinction of species due to the increase in temperature worldwide [Pin19]. This worldwide crisis is contributed to by the ongoing release of Greenhouse Gases (GHG) such as Carbon Dioxide (CO_2), methane, and nitrous oxide emissions, all of which can be measured by the metric Carbon Dioxide Equivalent (CO_{2eq}). To combat this global challenge, the Irish government has agreed to several international treaties which set carbon emission targets, such as the Paris Agreement[1], and the European Climate Pact[2].

To achieve these targets, the Irish government has established carbon budgets for various sectors, including agriculture and transportation. In particular, the Irish transportation sector is the second largest contributor to the release of CO_{2eq} in Ireland and is responsible for ≈11.2 Million Tonnes of Carbon Dioxide Equivalent ($MtCO_{2eq}$) in 2023, which is 21.4% of Ireland's total emissions in 2023. Currently, the Irish government has set a total carbon budget of 54 $MtCO_{2eq}$ for the period of 2021–2025 for the transportation sector, with a further budget of 37 $MtCO_{2eq}$ for the 2026–2030 [Dep23a].

To this end, the Irish government, as well as governments across the world, are pushing towards the adoption of low-emission vehicles such as Electric Vehicle (EV), which include Battery Electric Vehicle (BEV) and Hybrid Electric Vehicles.

[1] Paris Agreement: `https://unfccc.int/process-and-meetings/the-paris-agreement`
[2] European Climate Pact: `https://climate-pact.europa.eu/index_en`

For example, the Irish government's 2023 climate action plan has set a target of having 845,000 private EVs on Irish roads by 2030 [Dep23a], with government grants available to promote the purchase of such vehicles [Cit25].

Ireland has a car-centric transportation network. In 2023, private cars were responsible for 39.3% of emissions from the transportation sector [Sus24]. To combat this, the Irish government has committed to reducing the total km travelled by vehicles by 20% by 2030 and increasing public electric transport vehicles, such as Battery Electric Buses (BEBs). With a goal of 1,500 BEBs in operation by 2030 [Dep23a].

This is not a plan unique to Ireland. Several pilot projects are running worldwide to increase the use of BEBs. Such vehicles have also been successfully tested on certain routes in Spain, Germany, Poland, the Netherlands, Switzerland, and the United States. It is expected that more countries will use this technology in the future. This widespread adoption makes the topic of scheduling BEB an attractive prospect, in terms of ensuring service quality is not affected by the use of these vehicles and optimising the emissions of a bus fleet.

While BEBs lack tailpipe emissions, thus offering a cleaner alternative compared to diesel-powered Internal Combustion Engine Bus (ICEB), they can still produce lifecycle emissions. Lifecycle emissions refer to the release of CO_{2eq} during the manufacturing of BEBs (including the mining of resources), as well as the CO_{2eq} produced to create the electricity used to power BEBs. The focus of this thesis is the reduction of lifecycle emissions generated via electricity production. Because of this, the carbon emissions of the energy sector should be considered when scheduling BEBs.

Ireland's energy sector currently consists of a fuel mixture which includes coal, natural gas, oil, as well as renewable energy sources. This sector was responsible for 14.3% of Ireland's carbon emissions in 2023 [Env25]. However, the CO_{2eq} emissions of this sector reduced from 10 $MtCO_{2eq}$ in 2022 to 7.85 $MtCO_{2eq}$ in 2023. This reduction in emissions is indicative of the Irish government's widespread adoption of renewable energy sources, with further plans to increase reliance on such energy in the coming years.

For example, Ireland's capacity for renewable energy generation is expected to increase from 4834 MegaWatt (MW) to 9226 MW between 2021 and 2026 with existing and planned policies from the Irish government. Of which, 4358 MW of the capacity for 2021 originates solely from wind-generated energy, and is planned to increase to 6090 MW by 2026 [Dep24a]. The high capacity for wind-generated energy poses an opportunity to reduce the lifecycle emissions of BEBs. As it has been documented in literature, charging BEBs solely from renewable sources greatly reduces the lifecycle emissions of the fleet [RRHK20].

Recently, Battery Energy Storage System (BESS) have been investigated by the Irish government, with a recent publication explaining the concepts for such a system for both consumers and commercial entities [Sus25]. Explicitly outlining how BESS can be used to store excess energy generated by renewable sources above demand. Furthermore, recent estimations suggest the amount of available wind-generated power has exceeded the national demand for Ireland at least once in February 2025 [Eir25][3]. This is expected to increase along with the further development of both onshore and offshore wind farms, as previously outlined. This provides an interesting use case to use this estimated oversupply of wind energy to charge BEBs, instead of the energy being lost.

Knowledge of the oversupply of low-emission energy is needed to effectively use it for charging BEBs. To this end, advanced prediction technologies can be employed, such as Deep Learning (DL), to estimate how much wind-generated energy is available. However, how these predictions are used when scheduling BEB charging is important to consider when creating the prediction model [IOS12]. This creates an *imbalanced regression* problem, where accurate estimation of a small subset of the data is more important than the rest of the dataset [RM20].

In addition to the difficulties involved in the creation of an imbalanced regression prediction model, several issues arise when scheduling BEBs. These issues come not only from the focus on scheduling charges to minimise lifecycle emissions but also from systemic issues related to the use of BEVs. One such issue is the reduced operational range of BEVs due to the limited capacity of batteries. This means either additional BEBs are required to replace existing ICEB routes, or they must recharge their battery during the day.

While allowing BEBs to charge during the day is an attractive option to reduce capital investment, this may cause delays, as buses may need substantial time to recharge their battery to complete their daily operations. Such a negative impact on service quality can make public transport unattractive compared to travel via private vehicle, which is counterproductive to the Irish government's goal of reducing total km travelled by vehicles. Therefore, these delays caused by charging BEBs should be reduced as much as possible. The study of

[3] System data summary, Ireland, Wind penetration (% of demand)

scheduling charging operations for BEB is called the Battery Electric Bus Charging Scheduling Problem (BEB-CSP) [ZWW+24].

Generally, bus timetables are created months in advance, and require rigorous approval procedures to change. The initial scheduling problem investigated in this thesis uses an online timetable, where the schedule is recalculated every day to reflect changes in the amount of wind-generated power that is available. This schedule is also recalculated at checkpoints throughout the day if a variation in the estimations of available clean energy is detected. While this does not align with the previously outlined process of timetable approval, the arrival times of the new timetable are kept within a range of the original. It is assumed that future bus schedules will incorporate slack into the arrival times, thus allowing for this slight deviation every day. As the deviation from the original timetable is a configurable parameter, the bus operators would be able to explore the trade-off between more slack (represented by a lower deviation) and more clean energy consumption (represented by a higher deviation). Another scheduling problem explored in this dissertation aggregates electricity price and carbon data over one year to create a BEB charging timetable that would generalise for an "average" day of the year. This would align more closely with real-world timetable approval processes, as the timetable would only need to be recalculated yearly.

This work assumes the use of on-route fast charging stations. Contemporary literature has shown that on-route fast charging is both cheaper in terms of capital investment, and gives more opportunity for BEBs to charge from low-emission electricity. Furthermore, several countries, such as Germany, currently use this type of infrastructure to charge BEBs operating in their bus fleet. Because of the use of on-route fast charging stations when creating charging schedules for BEBs, the focus is placed on selecting which stations BEBs charge at and how long they charge for. This necessitates the reevaluation of established bus arrival times at each stop, thus changing the overall schedule for a bus route if operated by a BEB.

While it is known that road gradients impact the energy expenditure of BEBs, such information is not readily available for the datasets investigated. Therefore, it is assumed that the average impact of road gradients (both positive and negative) results in an overall minimal impact on energy expenditure.

The issues of creating an DL model for an imbalanced regression dataset, as well as optimising the BEB-CSP for reducing lifecycle emissions and negative impact on service quality are addressed in this dissertation.

1.2 Thesis statement

> *The lifecycle emissions of battery electric buses can be reduced by using predictive models to forecast low-emission energy availability and by optimising charging schedules, all while maintaining high service quality.*

To prove this statement, three contributions are made within this dissertation.

1.2.1 Contributions

1. A review of contemporary literature shows that BEBs charged purely from renewable sources result in extremely low lifecycle emissions. As a result, an approach is taken to estimate excess wind power above the Irish national grid demand. However, this creates an imbalanced regression dataset. A literature gap is identified regarding the empirical evaluation of resampling methods and loss functions designed for imbalanced regression datasets for DL approaches.

   In Chapter 3, three loss functions and resampling methods for imbalanced regression are tested for six different DL models to determine the best-performing combination. The resulting 54 DL configurations are evaluated with three test metrics based on the examined loss functions. It is determined that Convolutional Neural Network (CNN) trained with either Squared Error Relevance Area (SERA) or Inverse Weighted Mean Squared Error (IW-MSE) loss functions provides the best estimations for excess wind power.

2. A significant gap in the literature addressed in this dissertation is the evaluation of the impact of using a predictive model's estimations of an unknown variable versus 100% accurate historical information when solving a BEB-CSP. Furthermore, using only wind-generated power to charge BEBs has not been considered in contemporary literature.

   In Chapter 4 this gap is filled with the creation of a BEB-CSP to minimise the consumption of non-clean energy, considering the Time-of-Use (TOU) of BEB charging and the availability of wind-generated power throughout the day. Evaluation is performed on the bus systems of three Irish cities with 14 examination days denoting different availability of wind-generated power. Additional sensitivity analysis takes place to emulate the higher availability of wind-generated power. Results showed that the use of predictive models reduces the consumption of non-clean energy compared to a naïve scenario with no knowledge. However, a gap between results using the predictive model and those using 100% accurate historical information can be observed.

Aspects of this contribution are published in:

> *Padraigh Jarvis, Laura Climent, and Alejandro Arbelaez. Smart and sustainable scheduling of charging events for electric buses. Transactions in Operations Research (TOP) 32, 22–56 (2024) https://doi.org/10.1007/s11750-023-00657-5*

3. The final gap addressed by this dissertation pertains to a combination of aspects for the BEB-CSP which has not been seen in contemporary literature. Specifically, the use of non-linear battery charge times, partial battery recharging, limited charger availability, and weight-dependent energy consumption for a mixed fleet of ICEBs and BEBs. Additionally, when service quality is considered, such as passengers' Value-of-Time (VOT), modifiers such as the impact of late arrivals versus early departures are not represented. Works that model emissions as a financial value, such as Social Cost of Carbon (SCC), only examine a single value for the cost of burning one tonne of CO_{2eq}. Furthermore, there is an absence of contemporary literature which focuses on optimising lifecycle emissions and service quality on large datasets.

   In Chapter 5 these gaps are addressed, where a BEB-CSP is created with all previously outlined aspects with a dataset featuring over 1000 buses. Fuel price, emissions, and service quality are modelled as a financial cost using real-world values as weights, and the resulting sum is minimised. Two weight values are explored for SCC, while VOT modifiers are considered for impact on service quality. The three largest cities in Ireland are examined with sensitivity analysis taking place to determine the impact of different weights for SCC, and the impact of decomposition and simplified approaches on solution time and quality. Through these methods, the emissions of the largest bus fleet can be reduced by over 100 Tonne of Carbon Dioxide Equivalent (tCO_{2eq}) per day, while only causing an average delay of $\approx$8 seconds per BEB.

Aspects of this contribution are published in:

> *Padraigh Jarivs, Laura Climent, and Alejandro Arbelaez. Multi-objective mixed bus fleet charging schedule problem with Time-of-Use for real-world datasets. Proceedings of the 32nd Irish Conference on Artificial Intelligence and Cognitive Science (AICS 2024).*

1.3 Thesis outline

The remaining contents of this dissertation are organised into five chapters. The structure of these chapters is as follows:

Chapter 2 outlines the background information needed to understand the concepts, methods, and terminologies used throughout this thesis, as well as contemporary works which are relevant to the thesis statement. The fundamentals of time series analysis are described, as well as sophisticated prediction methods such as Long Short-Term Memory (LSTM), CNN, and Gated Recurrent Unit (GRU) models. The use of such models is discussed concerning predicting low-carbon energy. Issues related to imbalanced regression datasets are analysed, followed by a review of contemporary literature on methods to address these challenges. The fundamentals of Constrained Optimisation Problem (COP) are introduced with illustrative examples, and tools such as Mathematical Programming (MP) are explored. The basics of the Vehicle Scheduling Problem (VSP) are outlined, including important terminology. The Electric Vehicle Scheduling Problem (E-VSP) is then discussed as an extension to the traditional VSP, including the additional issues faced with using EVs. This discussion then extends to the use of BEB, with a comprehensive evaluation of related literature regarding the scheduling of BEB and the BEB-CSP. Additional topics are discussed related to considerations that should be made when creating charging schedules for BEBs.

Chapter 3 presents an evaluation of methods for training DL models based on imbalanced regression datasets. Data from the Irish national grid is collected and analysed, and a dataset for excess wind-generated electricity is created based on the Irish government's future expansion plans. DL approaches such as LSTM, GRU, and CNN models, along with hybrid models, are used to evaluate three resampling techniques and three loss functions. Three test metrics are used to determine the best-performing configuration, which is then extended to predict further horizons. This chapter is tied to Contribution 1.

Chapter 4 outlines a BEB-CSP to increase the usage of energy from renewable sources, specifically excess wind-generated electricity, and is tied to Contribution 2. The two best-performing predictive models from Chapter 3 are used to provide estimations regarding the availability of excess wind-generated energy to the optimisation model developed in this chapter. Where schedules prioritise recharging when clean energy is available, the model is empirically evaluated using data from three major cities in Ireland, with scenarios explored for different battery capacities, and the availability of information regarding clean energy. This chapter also includes considerations on dealing with inaccurate predictions being used as an input to an optimisation problem. Further sensitivity analysis takes place to determine the impact of several key parameters.

Chapter 5 addresses Contribution 3 by extending the optimisation model developed in Chapter 4 and posing it as an objective problem with multiple components. The consideration of clean energy is extended to real carbon emission information from the Irish national grid. Meanwhile, the electricity tariff for the energy used to charge BEBs is also integrated, along with the VOT for public transport users if the new schedule imposes any changes. These objectives are minimised, with weights based on real-world values used to represent each aspect as a financial amount in euros. Several shortcomings of the model used in Chapter 4 are addressed, such as the consideration of non-linear charging of batteries, variable energy consumption based on passenger load information, and mixed fleets of BEBs and ICEB. The outlined model is evaluated on the three largest cities in Ireland, with decomposition and alternative approaches examined to improve performance on larger datasets. Sensitivity analysis is also performed concerning the weight value used for carbon emissions to determine the impact of taking more aggressive carbon-reducing policies.

Chapter 6 concludes the dissertation, with a discussion regarding how each contribution ties back to the thesis statement, along with possible directions for future work in the area of BEB-CSP and reducing lifecycle emissions.

Chapter 2

Background

2.1 Introduction

In this chapter, the principles and concepts used throughout the thesis are introduced, such as time series analysis, Deep Learning (DL) methods, Mathematical Programming (MP), and the Vehicle Scheduling Problem (VSP). These topics are examined in detail to impart a base level of understanding to the reader, so no additional background knowledge is required for the rest of the thesis. Additionally, contemporary literature is examined and gaps in related research areas are identified.

In Section 2.2 some initial basic definitions are outlined. Section 2.3 discusses the basics of time series analysis, along with DL approaches, and contemporary literature around imbalanced regression. Section 2.4 describes some related literature on the topic of predicting electricity demand, as well as estimating the availability of renewable energy. In Section 2.5 the basics of constrained optimisations are explained, such as Constraint Satisfaction Problem (CSP), Constrained Optimisation Problem (COP), MP, and problem complexity.

Section 2.6 outlines the basics of the VSP. In Section 2.7 the extension of VSP to include Electric Vehicle (EV) is outlined. In addition, the Battery Electric Bus Charging Scheduling Problem (BEB-CSP) is defined and an extensive review of contemporary literature is done. Section 2.8 discusses the topic of combining predictive models with optimisation models, while Sections 2.9 and 2.10 expand on details for carbon emissions and passenger impact discussed earlier in the chapter. Finally, Section 2.11 concludes the chapter by enumerating the literature gaps and their relationship with the thesis statement and contributions discussed in Section 1.2.

2.2 Basic definitions

To understand the research area related to the thesis statement, some definitions are required that are used throughout this dissertation. The first is the Battery Electric Bus (BEB):

Definition 2.2.1 (Battery Electric Bus). *Battery Electric Bus is a public transport bus which is powered entirely by electricity and lacks an internal combustion engine. These vehicles store electricity in an onboard battery and lack tailpipe emissions.*

As stated, BEBs are entirely powered by electricity. As a result, they lack tailpipe emissions. This leads to the next definition:

Definition 2.2.2 (Tailpipe emissions). *Tailpipe emissions of a vehicle are the Carbon Dioxide Equivalent (CO_{2eq}) emissions released by a vehicle during its operation.*

While BEBs lack tailpipe emissions, there may be a misconception that they are 100% carbon-free. This is not the case. BEBs are powered by electricity, which itself may come from sources such as coal-powered generators or renewable energy sources. Emissions released by electricity generation methods are counted towards the vehicle's lifecycle emissions, which leads to the definition:

Definition 2.2.3 (Lifecycle emissions). *Lifecycle emissions of a vehicle are the total amount of CO_{2eq} released throughout a vehicle's entire life from its operation. Which includes CO_{2eq} released when generating the electricity used to power EV.*

Therefore, to reduce $\mathrm{CO_{2eq}}$ emissions, knowledge of the carbon emissions of the source of the electricity used is required. Information such as this is often represented as time series data.

2.3 Time series analysis

A time series is a collection of data measurements over a period of time, interspersed by a temporal resolution (i.e., a period of time between data points). A large variety of data can be expressed as time series values, such as flight passenger numbers, stock prices, and music. Time series data can be studied

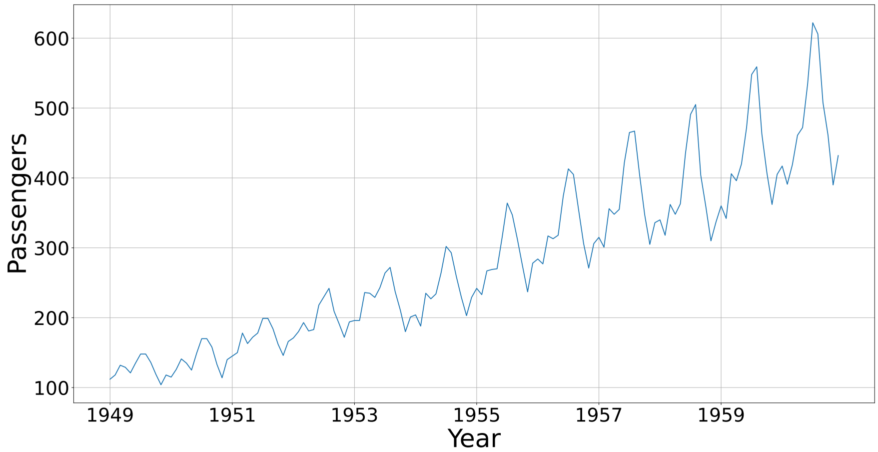


Figure 2.1: Airline passenger dataset for 1949–1960

to determine patterns and trends that occur. This is called time series analysis and has existed as a field for decades. One common use of time series analysis is the estimation of future data points. For example, given a time series of length x predict the time series value at $x + 1$ [BJR13]. In general, the accuracy of time series methods varies considerably with different forecasting horizons [SDMB+21].

Figure 2.1 shows an example of a time series dataset, the monthly number of air passengers from 1949 to 1960[1]. The pattern observed in the data is commonly known as a trend in time series analysis and indicates that the dataset exhibits an increasing pattern change over time. Insights such as this can be used to determine future actions to improve customer satisfaction or profits, as it can be determined that the peak number of flight passengers for subsequent years will be higher than prior years. Therefore, an increased number of flights can be scheduled to benefit from this additional demand.

Over the years, several tools and methods for time series analysis have been developed. AutoRegressive Integrated Moving Average (ARIMA) is a well-known statistical model used for time series analysis, where patterns are learned from historical data and predictions are made for subsequent values [SSS00]. Seasonal AutoRegressive Integrated Moving Average (SARIMA) extends the consequences of ARIMA with seasonality components. As a result, it shows more promising results for time series data which contain some seasonal trends [BJR13].

Compared to traditional methods, several recent works have evaluated time series forecasting using DL approaches such as Long Short-Term Memory (LSTM)

[1] dataset available at https://www.kaggle.com/rakannimer/air-passengers

models across different domains. These works have found that DL approaches can outperform those of ARIMA and SARIMA for several different time series problems [KDKGD+21, KKIM21, SBA22, SB22].

2.3.1 Deep Learning time series prediction

DL approaches are powerful prediction tools, with methods developed for domains such as image classification, natural language processing, and time series analysis. One main advantage of these approaches compared to traditional ARIMA and SARIMA is the ability to deal with particularly large datasets. Making them potentially more accurate and scalable [THS+21].

DL models, such as Artificial Neural Network (ANN), consist of *layers* which contain one or more neurons [RN16]. Each neuron has associated weights, biases, and an activation function. The input of each neuron is subjected to its associated weight, while the bias is used to modify the inputs based on the chosen activation function. Examples of activation functions include hyperbolic tangent (Tahn), sigmoid, and linear. During the training process, the biases and weights of each neuron are adjusted based on the performance or loss of the model. The final layer of a DL model represents the prediction that is produced. For example, a final layer consists of a single neuron with a softmax activation function representing the probability for a particular class, which is useful for binary classification problems (i.e., classifying a tumour as benign or malignant). In contrast, a final layer with n neurons and a linear activation function will produce n predictions of real values, which is commonly used for regression problems (i.e., estimating house prices).

DL models are trained over several epochs to optimise a loss function by adjusting the weights and biases assigned to each neuron. For regression problems, traditional loss functions used include Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and coefficient of determination (R^2). However, more specialised loss functions also exist for specific problems.

Ordinarily, DL approaches work by training on a dataset split into features (X) and labels (y). Where given X, the predictive model estimates that the label is ($\hat{y}$). This differs from time series predictions, where the preceding values for y before the prediction should have a strong correlation. To adapt DL approaches for time series data, several approaches have been developed. One such approach is called one-to-one (also known as iterative) is one such approach where the preceding time series value is provided to the model and a prediction of one time-step ahead is generated. This prediction is then provided as an input to the model to predict the next time-steps. This approach has the strength of being usable for an indefinite number of time-steps in the future and operates similarly to ARIMA and SARIMA models. However, there are shortcomings with this approach, as small errors in predictions can be exacerbated after several iterations [LZ21].

Many-to-many (also known as direct) approach makes use of multiple historical inputs to the model to predict multiple time-steps in the future. Inputs are gathered into a sequence of length k and provided to the model to estimate the next n time-steps. This method overcomes the issue with the one-to-one approach, as the DL model will be trained to account for the error over the multiple time-steps it will predict. However, this method is not without its shortcomings, as the prediction model will only be able to predict a fixed number of time-steps ahead [SD20].

There also exists an approach called many-to-one, which hybridises the two prior approaches by using multiple historical inputs to predict one time-step in the future. This has the benefit of optimising the loss for one value at a time. However, the use cases of such a model may be limited if information about events is needed ahead of time, such as long-term forecasting.

Some publications have questioned DL efficacy for time series analysis, with particular focus on comparing a simple linear model (DLinear) with transformers, arguing the attention mechanisms present in transformers harm time series predictions [ZCZX23]. However, more recent literature has shown both transformers and other DL methods can outperform such simple linear models when dealing with complex non-linear datasets such as wind speed prediction [CZZ24].

2.3.2 Recurrent Neural Networks

Recurrent Nerual Network (RNN) are an example of a DL model based on the previously mentioned many-to-many approach. Figure 2.2 shows an example of the core component of RNNs, memory cells for a many-to-many model with an input sequence (k) of length 2, and a prediction horizon (n) of 2. Memory cells represent the internal memory of the network, which is capable of

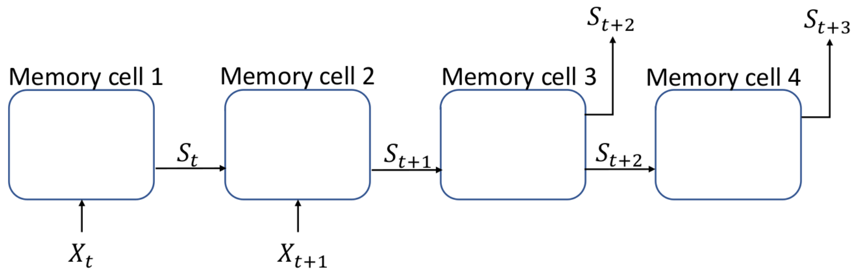


Figure 2.2: Many-to-many RNN architecture

remembering previous values from the input sequence X [THS+21]. As seen, the first memory cell in the architecture takes a single input parameter (i.e., the first value in the input sequence X_t). Every subsequent memory unit after this takes two inputs, the next value in the sequence (i.e., X_{t+1}) and the output of the prior cell (S_t), with each cell subjecting their inputs to a tanh activation function. The introduction of the output of the previous cell allows the RNN to remember information about the cell throughout the training process, and recognise the relationship between successive values in the sequence. The output of the final two cells becomes the prediction of the RNN.

However, RNNs face issues when dealing with patterns in long-term data. As a result, RNNs are considered to have short-term memory and are unable to notice long-term trends. Over recent years, these shortcomings have been addressed with the creation of LSTM models and Gated Recurrent Unit (GRU).

2.3.3 Long Short Term Memory

LSTM models are a popular deep learning architecture proven to be effective at energy forecasting [LZ21]. They are based on RNNs and address the shortcomings of short-term memory by introducing an additional parameter, C_{t-1}, to pass between cells in the network. This parameter is provided to the cell in addition to S_{t-1} and X_t and regulates the memory function of the network [LZ21, THS+21].

Figure 2.3 depicts the internal structure of each memory cell in an LSTM network. Each cell contains three gates, the input, update, and output gates. The input gate determines what information is discarded from previous cells, using a sigmoid activation function and element-wise multiplication.

The update gate decides if the value of the input gate is modified, which is done with a Tanh and Sigmoid activation function and an element-wise multiplica-

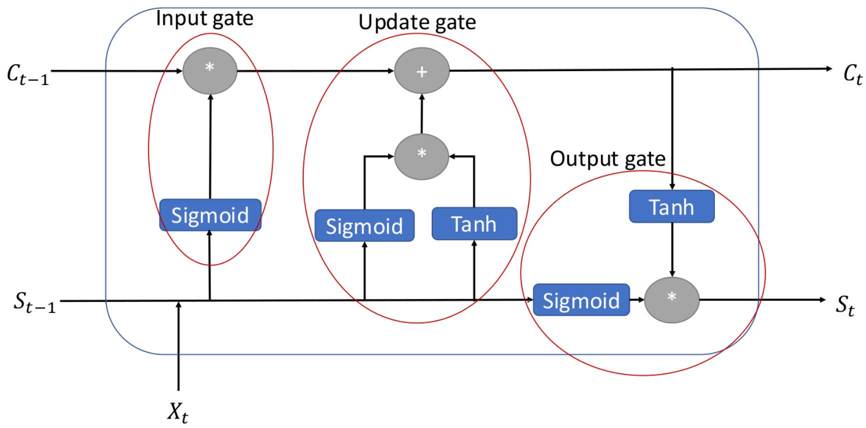


Figure 2.3: LSTM memory cell composition

tion. The resulting value is used as the cell state for time t and is passed into the next cell in the series. The output gate decides the prediction value at time t using the cell state from the update gate subject to a Tanh activation function, and the values for S_{t-1} and X_t subject to a Sigmoid activation function. An element-wise multiplication takes place, and the resulting value (S_t) is passed to the next cell.

LSTMs have been shown to outperform traditional RNNs when it comes to time series problems. However, the inclusion of cell states results in increased training times [CGCB14].

2.3.4 Gated Recurrent Units

GRU models are another type of RNN that aims to solve the issue of remembering long-term sequences similarly to LSTM. GRUs also make use of gates. However, they feature reset and update gates but lack a dedicated output gate. One major difference between GRUs and LSTM cells is the lack of the additional cell state [DS17].

Figure 2.4 shows an example of the internal structure of a GRU. As can be seen, both gates subject the inputs X_t and S_{t-1} to a Sigmoid activation function. In the case of the reset gate, the resulting value is used in an element-wise multiplication with the input S_{t-1}, this is then passed to a Tanh activation function along with the input X_t to generate the candidate state. The output of the update gate is subject to an element-wise multiplication with the value of the candidate state. The update gate's output is also reversed to a negative value and used in an element-wise multiplication with the input of the previous cell

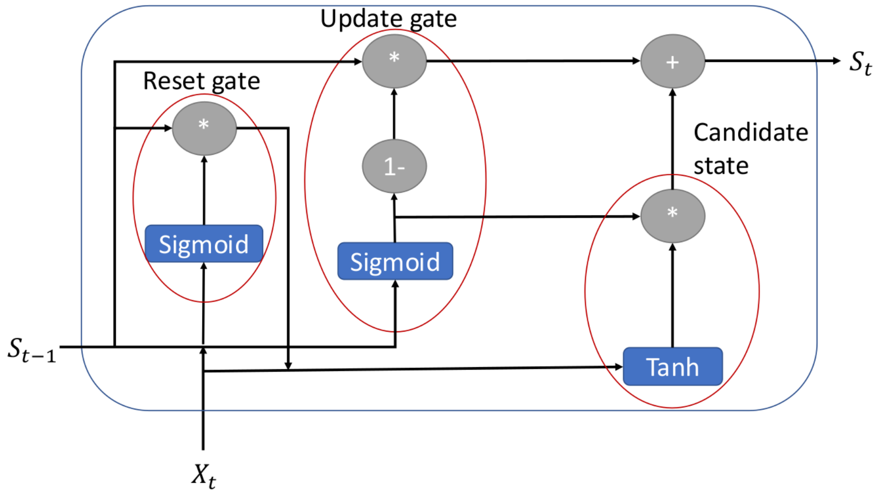


Figure 2.4: GRU memory cell composition

S_{t-1}. This value and the modified value of the candidate state are subjected to an addition operation and provided as the output of the current cell.

GRUs have been shown to generate similar results to LSTMs while having shortened training time thanks to their reduced number of trainable parameters. However, the lack of an output gate means all information from the cell is exposed to the next cell in the sequence [CGCB14, WLC18].

2.3.5 Convolutional Neural Networks

Convolutional Neural Network (CNN) are a type of DL method originally developed for image classification problems [ON15]. However, in recent years it has seen popularity in use for time series problems both as a stand-alone model and when combined with RNN approaches such as LSTMs [LPP20, MS22, SZS+24, LZ21, SVG+17]. Such models consist of convolution layers followed by pooling layers.

In the convolution layer, filters and kernels are used to learn patterns present in the data. These layers use Rectified Linear Unit (ReLu) activation functions. The pooling layer then acts to reduce the dimensionality of the convolution layer using a pooling size. The numbers of filters, kernels, and pooling size are all important hyperparameters to optimise. A filter size that is too small results in the network being unable to represent features, while too large a pooling size will result in a significant amount of information being lost from the convolution layer [ZLC+17, THS+21].

2.3.6 Imbalanced Regression

As previously outlined, predictive models such as DL methods learn from historical data to recognise patterns, and accurately predict either classes for classification problems or values for regression problems. However, in some cases, more importance is placed on accurately predicting extreme minorities of the dataset. For example, when classifying diseases, some may be rarer than others and are under-represented in the data, resulting in a class imbalance. As a result, the accuracy in predicting these classes can suffer [BTR16].

This definition of imbalanced data can also be extended to time series regression datasets, which have received more attention in recent years. Where, for example, a user may place more importance on predicting values that are statistical outliers for anomaly detection. However, regression and classification data differ, as the former lacks a clear definition of what constitutes a rare case. To this end, several works have adapted methods for dealing with imbalanced classification datasets for regression datasets or developed new methods entirely.

To determine if a regression value is considered a part of a rare case, several works have made use of relevance functions $\phi(y)$ to determine how relevant a regression value is. Relevance values can be calculated in several ways. One such approach is using Piecewise Cubic Hermite Interpolating Polynomials (PCHIP) interpolation with control points derived from the lower adjacency, median, and upper adjacent values used in a Tukey box plot [RM20, BTR19, ACC24, SSHSP23, MBT17]. In such relevance functions, samples close to the median are assigned a relevance value of 0, while those closer to statistical outliers are assigned relevance values up to 1. In general, values closer to 1 are considered rarer cases, and more significance is given to them for resampling and evaluation methods.

When dealing with imbalanced classification, resampling has proven to be a popular approach to represent rare classes [KSA$^+$21]. Such methods have been adapted for use in imbalanced regression problems. Common examples of resampling are random undersampling and random oversampling. In the former case, samples that belong to a normal (i.e., non-rare) class are randomly sampled. Random oversampling involves duplicating rare samples, in both over and under sampling, the process is generally carried out until the rare cases represent a given percentage of the total dataset, or both the rare and normal cases are approximately equal.

Adopting these methods for use in imbalanced regression problems involves the use of a relevance threshold t_R, which represents a returned value from the previously mentioned relevance function that represents a rare case. For example, if a t_R value of 0.8 is provided, any regression value which has a relevance value greater than or equal to 0.8 is considered a rare case [BTR19, ACC24].

Another example of a resampling strategy for imbalanced datasets is the Synthetic Minority Over-Sampling Technique (SMOTE), which removes samples of low relevance and artificially generates more relevant ones via interpolation with k nearest neighbours of the under-represented samples [CBHK02]. While SMOTE was created for classification problems, an adaptation called Synthetic Minority Over-Sampling Technique for Regression (SMOTER) has been made [TRPB13]. Gaussian Noise (GN) resampling has also proven to be effective, where over-represented samples are removed using random undersampling, and new samples are generated by taking rare cases and adding GN [ACC24]. Additional resampling strategies have been proposed for imbalanced time series regression problems, which places importance on when a sample occurs. Examples of this include temporal bias over and undersampling, where samples that occur later in the time series are given a higher weight of being sampled in the case of under-sampling or being duplicated in the case of over-sampling [MBT17]. This can be further extended to temporal relevance bias, where samples with higher relevance are given more importance when resampling in addition to when the sample occurs.

Another issue faced with imbalanced regression is the evaluation metrics used. MSE and RMSE provide overall error values for the dataset. However, these metrics do not evaluate how accurately rare cases are predicted. Precision and recall are one such evaluation metric for classification problems that has been adapted for regression datasets [TR09]. New metrics have also been proposed, such as Squared Error Relevance Area (SERA) [RM20] and Inverse Weighted Mean Squared Error (IW-MSE) [SSHSP23], which place higher weighting on the error of values with a relevance closer to 1.

While several methods have been adapted for imbalanced regression, the majority of these approaches have been used with learning algorithms such as Random Forest (RF) and Decision Trees (DT). Work has also been done to adapt neural networks such as Multilayer Perceptrons (MLP) for use in imbalanced classification [AGS+07, Oh11]. Some other works aim to improve the performance of DL approaches for imbalanced regression problems by adding feature distribution smoothing, label distribution smoothing [YZC+21], or proposing new models with attention mechanisms [HWCF21]. Another work looks at adopting evaluation metrics used in imbalanced regression for use as loss functions when predicting wind speeds [SSHSP23].

2.4 Power forecasting

The topic of estimating power demand has existed within the literature for decades, with accurate prediction of demand from time series data existing before contemporary DL approaches [SIMC13]. The accurate estimation of electricity demand is vitally important for domains such as the energy trading market [OCPV24], where changes in demand due to seasonality can cause fluctuations in electricity prices. Knowledge of electricity demand is also critical for setting renewable energy targets to help decarbonise the energy sector [SS12].

In more recent years, DL approaches have been applied to the estimation of electricity demand. This includes the use of LSTMs to predict the power demand in countries such as Poland and India, which has been shown to outperform more traditional machine learning algorithms [Man20, BT19]. GRUs have also been employed to predict electricity demand in Taiwan, outperforming traditional time series analysis tools such as ARIMA and other DL approaches such as LSTMs [YCW+22].

In parallel to the estimation of electricity demand, the prediction of renewable energy has also gained attention in recent years. With growing concerns about the long-term impact of global warming, countries have pledged to reduce carbon emissions by signing up for the Paris Agreement[2]. To meet carbon emission targets, several countries have looked to adopt more renewable energies into their national power grid. With wind proving to be a prominent choice with the national grid of countries such as Ireland and Scotland able to supply a significant portion of their national demand with wind power for a period of time [Iri24, Sco21].

As a part of decarbonising the energy sector, accurate estimation of renewable energy such as wind power is required. As previously outlined, the accuracy of time series values can depend heavily on the prediction horizon. This is especially the case for wind power prediction due to its stochastic nature. As

[2]https://unfccc.int/process-and-meetings/the-paris-agreement

such research in this area can be divided into several categories. Such as immediate prediction (several hours ahead), short-term prediction (a day ahead), and long-term prediction (several days ahead) [WGH11]. However, the actual definition of forecast horizon varies between different works.

Contemporary work on this topic has focused heavily on the use of DL methods and RNN models such as LSTMs and GRUs as these methods demonstrate their usefulness when forecasting chaotic time series data [SD20]. These methods have proven to outperform more traditional predictive models, with hybrid models such as CNN-LSTM and CNN-GRU proving to outperform base LSTM and GRU models when predicting wind power [WZL+21]. Both 1D and 2D CNNs have been employed in hybrid models, where it has been found that 1D-CNNs have proven superior at feature extraction for use in wind power time series and feature less complexity than their 2D counterparts. Several works have also shown to make use of different multiple features with a strong correlation to wind power, such as wind speed, temperature, and wind direction. Such works have also commonly been created using only one year's worth of historical data[HLLL20].

Extended works on the topic of DL methods show that GRU outperforms LSTM in both accuracy and training time (due to their simplified nature) [KLL21, FDAH+22]. However, in some cases, LSTMs have been shown to outperform GRUs, especially when paired with CNN and attention mechanisms [SZS+24]. Research on the topic of imbalanced wind power time series data also exists, with DL methods such as CNN and GRUs showing deteriorating error scores as the number of samples for extreme cases of wind power generation decreases [MXY+23].

One aspect of estimating wind power is the prediction of wind speed, as there is a higher correlation between the two. To this end, several works have looked at accurately estimating wind speed using traditional time series models and DL approaches [SDMB+21]. However, some works highlight the need for a large amount of data for DL methods such as GRU and LSTMs to outperform traditional time series prediction methods such as SARIMA [LLF21]. Other works focus on the estimation of wind gusts or extreme wind speeds, which constitutes estimating rare events and is an example of imbalanced regression [ZZZZ24, SSHSP23].

2.5 Constrained optimisation

To create a schedule for BEBs charging, information about the problem must be known and modelled.

2.5.1 Constraint Satisfaction Problem and optimisation

One common way of problem formulation is the Constraint Satisfaction Problem (CSP). A CSP defines a set of variables, each associated with a domain of possible values to satisfy one or more constraints [RN16]. This can be represented as:

- $X = \{X_1, ..., X_i\}$
- $D = \{D_1, ..., D_m\}$
- $C = \{C_1, ..., C_l\}$

Where X represents a set of variables, D_i represents the domain of variable X_i, and C represents a set of constraints. A solution to a CSP is an assignment of values to all variables such that all constraints C are satisfied. For example, consider the following simple case:

$$x \leq 3 \tag{2.1}$$

$$x > y \tag{2.2}$$

Where x and y are variables with domains $x \in \{1, 2, 3, 4\}$ and $y \in \{1, 2, 3\}$, and Equations 2.1 and 2.2 represent constraints. There exists three possible assignments for each variable to satisfy Constraints 2.1 and 2.2. $(x = 2, y = 1)$, $(x = 3, y = 2)$, and $(x = 3, y = 1)$. For CSPs, all valid solutions are output; however, often there is a preferred or optimal solution to a problem which can optimise the use of a resource such as time or money. As such, a CSP can be turned into a Constrained Optimisation Problem (COP) with the addition of an objective function [RVBW06].

An objective function is a mathematical equation that is used to evaluate a solution to a COP. The goal of the COP is to find a solution that minimises or maximises this mathematical equation to find the optimal solution. This optimal solution can be referred to as a global minima for minimisation problems, a global maxima for maximisation problems, or a global optima in either case. For example, an objective function can be added to the previously outlined problem to find an optimal solution.

$$maximise : x - y \tag{2.3}$$

If the objective function expressed in Equation 2.3 is used for Constraints 2.1 and 2.2, then the solution $(x = 3, y = 1)$ will be returned as the optimal solution. It is important to note that $(x = 3, y = 2)$ and $(x = 2, y = 1)$ are still considered valid solutions. As such, if the optimal solution cannot be found (due to external factors such as a timeout), a suboptimal solution may still be provided so long as one has been found and it is valid.

2.5.2 Mathematical Programming

In the previously outlined example, all variables may only be assigned an integer value (i.e., the domain for all variables is discrete). These types of problems can be called Integer Programming (IP); however, if all variables are continuous and feature linear constraints, then it is a Linear Programming (LP) problem. Linear constraints feature strict inequalities and avoid the multiplication of variables by other variables. Mixed Integer Programming (MIP) is another type of problem, where some variables have a discrete domain, while others are continuous. IP, LP, and MIP all belong to a class of programming called Mathematical Programming (MP) [Wol20]. IP and MIP can be hard problems to solve, however, LP problems can be solved in polynomial time. IP and MIP problems, which contain linear constraints, are called Integer Linear Programming (ILP) and Mixed Integer Linear Programming (MILP) problems, respectively, and can be approximated into LP problems using LP relaxation.

To understand LP relaxation and its importance, some knowledge of problem complexity is required. An algorithm's effectiveness is often evaluated based on how quickly it executes. To evaluate algorithms in a way agnostic to the particular software it is developed in, or the hardware it is run on, a method called O() notation is used (sometimes also referred to as Big O notation). This symbolises the number of operations which must take place for the algorithm to complete. Take a $max()$ algorithm for example, which, given an input sequence s of numeric values, returns the element with the highest value. If the sequence s is of length n, then n operations are required to find the highest value. Therefore, the algorithm used in $max()$ has a complexity of O(n).

If a problem has a deterministic algorithm whose complexity can be expressed in Big O notation, it is classified as a polynomial-time problem. This class of problems are generally considered relatively easy compared to other classifications. Another classification for problems is Nondeterministic Polynomial time (NP) problems. Such problems have an algorithm which, provided with a solution, can prove if the solution is valid within polynomial time. While validating solutions can be done in polynomial time, finding solutions is much more difficult. Other classifications related to NP also exist, such as NP-hard and NP-complete. NP problems can be classified as NP-hard if a polynomial-time reduction for the problem exists. However, not all NP-hard problems are NP in nature. NP-complete problems are considered the hardest type of NP problem, such type of problem is considered both NP and NP-hard. An example of an NP-hard problem is the popular Travelling Salesman Problem (TSP), while an example of NP-complete is the graph colouring problem.

As previously outlined, IP and MIP problems can be approximated to LP problems using LP relaxation. Ordinarily, an integer value may only take a value from a finite domain (i.e., $x \in \{1, 2, 3, 4\}$). However, this can be relaxed into a linear constraint ($1 \leq x \leq 4$). This relaxed version of the problem is significantly easier, as LP problems can be solved in polynomial time. The solutions generated via this relaxation are used to inform variable assignments in the full version of the problem. An integrality value is used to set a maximum allowed amount the value for a variable with the optimal assignment is allowed to violate the bounds of the same variable in the unrelaxed problem. Effectively, this means the optimal variable assignment for the LP relaxation is very close to the domain of the variable in the full problem.

One technique used often when modelling MIP problems is Big-M, which allows for the modelling of logical *if* statements in a linear fashion. This technique is comprised of a binary variable (i.e., an integer variable with a domain of 0 or 1) to denote the occurrence of an event. Big-M is a relatively large constant that affects a constraint depending on the value of the binary variable. For example, assume a problem where variable x must be larger than variable y if event z occurs. This can be modelled as seen in constraint 2.4.

$$x \geq y \cdot z \tag{2.4}$$

If event z occurs (i.e., binary variable z has a value of 1), x will be greater than

y. If event z does not occur, the lower bound on x will be removed (as the constraint equates to $x \geq 0$). However, the multiplication of two variables is not linear. This modelling can be reformulated using big-M to keep the model linear as seen in constraint 2.5.

$$x \geq y - M(1 - z) \tag{2.5}$$

Where M is a value that is at least as large as the highest domain value of y. This equation is linear in formulation, as it avoids variable multiplication, and still enforces the desired logical if statement. While this is a useful tool, the use of Big-M comes with some risks. As previously outlined, due to LP relaxation and integrality values, integer variables can take values close to, but not exactly within, their domain. This can cause an issue called trickle-flow, where binary variables can be assigned a value close to 0, such as $1x10^{-5}$, and still be considered 0. This may cause M to have a non-zero value in a constraint when it should not. To mitigate this, it is important to carefully choose a value for M based on the constraints it is used in.

A search space contains all possible solutions, both feasible and infeasible, for a problem. As constraints are added to the model, the size of the search space, which contains feasible solutions, reduces. When looking for a solution to a problem, this search space is explored, with the objective function for the problem used to evaluate the quality of the solutions found. The optimal solution in the search space is called the global maxima or minima, depending on whether the problem uses a maximisation or minimisation objective function.

Algorithms such as branch and cut used by CPLEX (a popular MP tool) are employed to explore this search space and find the optimal solution to a problem [IBM]. To apply this algorithm, the search space of the problem is represented as a search tree, which consists of nodes and branches. Nodes in a search tree may represent different things depending on the problem. For MIP problems, they represent a solution to the LP relaxation of the original MIP problem. If the solved LP problem does not have any integer values that violate the maximum allowed integrality value, then it is considered the optimal solution to the MIP problem. If this is not the case, the branch aspect of the algorithm occurs where two child nodes are created from the current one. Child nodes are identical to their parent, except the bounds for one relaxed variable will change. Often this means that for one child the upper bound is reduced, while the other child features a higher lower bound. The variable which is subjected to this change in its bounds is referred to as a branched variable. The cut aspect of the algorithm is used to reduce the total amount of branching that is needed to find the optimal solution. This is done by the inclusion of additional constraints to the LP relaxed problem to reduce the domain size of the linear variables used, while not excluding valid integer assignments.

2.5.3 Metaheuristic approaches

Alternatives to MP exist when solving COPs, the most popular of which are Metaheuristic approaches. Metaheuristic approaches encompass several methods. One of the most popular ones is Local Search (LS) [RN16]. LS approaches begin with a starting state (i.e., an assignment of values to variables) which is often randomly generated, and may not be a valid assignment. Immediate neighbouring solutions are generated from this state, where minor changes are made, such as changing the value of the variable. Each neighbouring solution is evaluated based on a heuristic. A basic example of a heuristic is greedy search, where the neighbouring solution with the highest objective score is selected as the new current state. However, heuristics such as this may not find the optimal solution, as it may become stuck in a local optima, where all neighbouring solutions have a worse objective score, but a more optimal solution may exist somewhere else in the search space. To address this, a large number of metaheuristic approaches have been developed, such as tabu and large neighbourhood search. Methods outside of neighbourhood generation also exist. For example, Genetic Algorithm (GA) features a population of solutions [RN16]. Each iteration of the GA algorithm generates a new population by selecting two parent solutions and generating two child solutions by selecting variable assignments from the different parents and applying some level of genetic mutation (i.e., randomly changing a variable assignment).

2.6 Vehicle Scheduling Problem

The Vehicle Scheduling Problem (VSP) is a well-studied derivative of a scheduling COP, which assigns trips with fixed arrival and departure times to a vehicle, to create a *block* [BK09]. These are often based on operational and time constraints, which include each trip being serviced at least once, as well as deadheads. Deadheads are trips between a point of origin (such as a depot) to the start of the block, from the end of the block back to the point of origin, as well as between subsequent trips in the block.

The VSP commonly aims to minimise delays or overall costs. One of the primary use-cases for VSP is the scheduling of bus fleets. In such cases, additional aspects must be taken into account such as headway(i.e., the frequency of buses along particular routes or lines). This is of particular concern as some bus companies are required to maintain certain headways where inconsistent headway is considered undesirable for public transport users [vOBvN12].

There exist multiple variations of VSP. For example, Single Depot Vehicle Scheduling Problem (SD-VSP) focuses on VSPs where buses originate from a single depot, which can be solved in polynomial time if modelled in certain ways. A contrasting variation is the Multiple Depot Vehicle Scheduling Problem (MD-VSP), which allows vehicles to start at any one of many depots and is proven to be NP-hard [OF20].

Of particular interest are VSP variations, such as the Multiple Vehicle Type Vehicle Scheduling Problem (MVT-VSP) where different vehicles in the fleet may have other associated costs or operational restrictions. This problem has been proven to be NP-hard [LK81]. This topic of scheduling multiple vehicles can go by alternative names, such as mixed-fleet or heterogeneous fleet. Henceforth, the term mixed-fleet will be used with the following definition:

Definition 2.6.1 (Mixed-fleet). *A mixed-fleet scheduling problem involves the creation of a schedule for a group of vehicles which are not uniform in their specifications or requirements, such that additional considerations are required for each vehicle type in the COP.*

Another relevant variation of VSP is the Vehicle Scheduling Problem with Time Windows (VSP-TW), which defines a time window for each trip that denotes the earliest and latest possible departure times [BK09]. This problem is considered to be NP-complete. VSPs and their extensions have shown to be solved quite well using MP approaches [BK09].

A related problem is the Vehicle Routing Problem (VRP), which focuses on the generation of vehicle delivery routes. This problem assigns an optimal set of trips to a vehicle's route so every customer is visited, and the total distance travelled or route cost is minimised. Additional constraints are also considered, such as ensuring the vehicle does not exceed its cargo capacity [BRVN16]. A popular use-case of the VRP is the previously mentioned TSP [MSM10]. Much like VSP, extensions for VRP also exist. For example, the Vehicle Routing Problem with Time Windows (VRP-TW) adds a requirement that vehicles must visit a set of customers within certain predefined time periods (e.g., outlined by the customers or local governments). This adds additional complications to VRPs as a vehicle arriving early to a destination might be required to wait, while arriving late may invalidate the solution [DDS92]. This has also spawned variations such as Time Window Assignment Vehicle Routing Problem (TWA-VRP) focusing on assigning time windows to deliveries before the demand is known [SG15].

In these problems, assumptions may be made to simplify the solving process. Much like the scheduling problems previously outlined, time is made into a discrete variable. Another pertinent simplification is the ability for vehicles to complete all trips without needing to refuel. Variations of VSPs which feature limited fuel capacity have gained more attention in recent years, with Electric Vehicle Scheduling Problem (E-VSP) gaining a particular focus.

2.7 Electric Vehicle Scheduling Problem

The prevalence of Battery Electric Vehicle (BEV) has created sub-problems of VSP and VRP called Electric Vehicle Scheduling Problem (E-VSP) and Electric Vehicle Routing Problem (E-VRP). These problems feature additional challenges such as the reduced fuel capacity of BEV compared to traditional Internal Combustion Engine (ICE) vehicles, and limited BEV charging infrastructure compared to traditional gas station infrastructure [WS17]. This is exacerbated for BEBs as their higher energy consumption compared to personal BEVs and fixed schedules may lead to depleting the battery before the end of the operational day [JBZW22]. As a result, budgeting constraints are considered when transitioning from Internal Combustion Engine Bus (ICEB) to BEBs, to encapsulate the limited distance they can travel, which has been proven NP-hard [SO17]. This may result in the need for additional BEBs to meet customer demand, or a charging schedule so vehicles can recharge during the operational day.

Like traditional VSP and VRP, numerous variations exist [ECLR19]. The Green Vehicle Routing Problem (G-VRP) features vehicles with alternative fuel sources, such as electricity and hydrogen. The Fixed Route Vehicle Charging Problem for Battery Electric Vehicles (FRVCP-BEV), where routes are fixed and charging of BEVs is optimised. A version of the popular Travelling Salesman Problem (TSP) also exists, called the Hybrid Electric Vehicle Travelling Salesman Problem (HEV-TSP), where a decision must be made on whether a route is driven in diesel or electric mode. Additional work has also focused on identifying the best route to take for solar-powered EVs based on cloud coverage [GC23]

Alternative solution approaches to those previously discussed have also been explored in the literature. One such example of this is the use of DL to estimate what the next best destination should be if given an input sequence of bus stops for a E-VRP [FAC23]. However, this requires a large number of E-VRP solutions to train the DL model used. Furthermore, the authors of the work outlined that pure DL heuristics do not generalise well to larger problem sizes, requiring instances with similar sizes to be used as samples in the training process. It was also noted that the solutions generated by this approach are often suboptimal and featured long solve times. However, more recent research into this area has shown some promise with the use of Reinforcement Deep Learning for generating solutions to E-VRP which cannot be solved within a timeout by more traditional solution approaches [WWHG24]. Furthermore, this work also showed the potential of using this solution approach for instances of a size larger than those seen in the training dataset.

2.7.1 Electric public transport

Focusing on electric public transport, there are several different charging technologies which impact the planning and operation of urban transit [HCEQ19]. Overnight charging is a slow-rate charging technology which takes place at depot stations and relies on BEBs with high battery capacity to complete blocks without a need to recharge. Continuous charging allows BEBs to constantly charge due to either a hybrid motor (like in the HEV-TSP) or overhead bus charging lines. Battery swapping allows BEBs to swap out their batteries with a low State of Charge (SOC) with a fully charged one. Where the SOC of a battery can be defined:

Definition 2.7.1 (State of Charge). *The State of Charge of a battery is the current amount of energy remaining in a battery, expressed as a percentage value of the maximum battery capacity.*

Finally, Opportunity Charging (OC), also known as fast charging, can be defined as:

Definition 2.7.2 (Opportunity Charging)**.** *Opportunity charging is the use of high-rate charging stations (i.e., charging stations with a high current) to rapidly charge BEBs while passengers embark and disembark, or at specified charging points, such as depots.*

The planning process for public transport networks can be split into four phases. The first involves the network route and design, that is, the development of bus routes, stops, and the identification of depot locations. This phase must take into account all the mentioned charging technologies. For instance, the placement of fast charging stations for OC must be considered when creating bus stops. Overnight charging needs to calculate the amount of time and energy needed for the BEBs to undertake all trips in their block and perform the necessary deadheads. For battery swapping, work in the area has highlighted the need for specialised infrastructure for different techniques. Therefore, sufficient room for these machines is necessary [AAAS20]. This is exacerbated when dealing with mixed-fleets, which may require more than one type of battery swapping technique. Finally, the system needs to define potential bus routes for overhead bus charging lines for continuous charging.

The second phase involves developing timetables. Here, the use of OC and battery swapping must be considered. Any potential detours to charging or battery swapping stations must be considered, as well as increased idle times if recharging takes place during passenger loading and off-loading. Which would increase the total operational time of the BEB and the required number of deadhead trips. The third phase assigns blocks to BEBs where the use of overnight, opportunity, and continuous charging technologies should be considered. BEBs suited for overnight charging often feature higher battery capacities to complete their operational day without needing to recharge. However, they can be significantly more expensive than BEBs with smaller capacities. BEBs which use OC would need to be assigned to blocks with charging stations on their route. While BEBs using continuous charging would require infrastructure such as overhead bus charging lines to be on their route.

The final phase involves the assignment of buses to drivers, and this phase is not expected to be impacted by the usage of any charging technologies.

Each charging technology has its benefits and issues. Both continuous charging and battery swapping offer fast solutions to charging BEBs. However, as previously outlined, they rely upon specialised hardware. For these reasons, the majority of works related to BEB scheduling focus on the use of overnight and OC [PLL22]. Overnight versus OC has been a topic of debate for scheduling BEBs, with some work examining the effects of each type of technology on the lifecycle costs of a bus fleet over a prolonged period [Laj18]. Where depot charging, OC at bus stops, and OC at trip endpoints are examined. While depot charging has the lowest investment in charging infrastructure, it also features the highest cost associated with bus capital and battery replacements. This is because BEBs with higher battery capacities cost significantly more than lower-capacity batteries such as those employed for OC. While OC appears to be the overall cheapest option, effective use depends on BEB trips sharing the same end-points to minimise total investment in charging infrastructure.

The use of OC and optimal placement of fast-charging stations engenders a new problem, the Charging Station Location Problem (CSLP), which, as previously outlined, needs to be considered during the first stage of planning a public transport network [KB21]. OC and battery swapping technology also give rise to the problem of creating optimal schedules to charge BEBs, which is called the battery electric bus charging schedule problem (BEB-CSP).

2.7.2 Battery Electric Bus Charging Schedule Problem

While traditional EVSP generates blocks for vehicles to be assigned to, subject to range constraints, the use of OC and battery swapping allows BEBs to operate for longer periods. However, the decision of when these vehicles charge or change batteries is subject to some optimisation, as there may be costs associated with charging at certain points of the day, or time-related delays. Optimising a charging schedule for BEB is called the Battery Electric Bus Charging Schedule Problem (BEB-CSP), which is defined as:

Definition 2.7.3 (Battery Electric Bus Charging Schedule Problem)**.** *The Battery Electric Bus Charging Schedule Problem involves the creation of a schedule which denotes which times BEBs should recharge to optimise an objective function.*

For BEB-CSP, there exist two variations, the first is integrated BEB-CSP, where the charging schedule problem is integrated into E-VSP, where both trips and charging times are assigned to vehicles. The second is pure BEB-CSP, where blocks (i.e., bus routes) are fixed, and the problem becomes assigning charging times subject to maintaining the original bus schedule. The benefit of pure BEB-CSP is that it allows easier adoption of BEBs as they can be directly swapped with ICEBs, and as such is the primary focus of this dissertation [RCS+21, ZWW+24].

As previously mentioned, in some VSP mixed-fleets can be employed where not all vehicles are identical. Some BEB-CSPs utilise mixed-fleets, as blocks that cannot be serviced with BEBs are assumed to be serviced by ICEBs instead [WLOKF18, HLS23]. However, some problems consider different BEBs, where larger blocks are assigned BEBs with bigger battery capacities, and the battery size is optimised. These can also be classified as mixed-fleet Problems. However, most of these works assume that all BEBs start the operational day with full battery capacity from overnight charging [YY21, ZLZ+22, ZWWL22].

The availability of charging stations for OC also varies in the literature. Some implementations of the BEB-CSP have limitations placed on how many BEBs can use a charger at once [AC20, LYP+24, WS17]. While others assume that there are enough chargers available to service the entire fleet at any time [HLS23, ZMOW24]. Additionally, some works assume that when charging, a BEB regains a fixed amount of energy or fully recharges during a single charge event [GIK23, WLLJ22, RvdHLS18]. While others allow partial recharging to maximise the use of idle times at bus stops, or optimise electricity costs [ZXZL20, WS17, HDLP22, IK19].

2.7.3 Energy expenditure of Battery Electric Buses

Creating feasible solutions to the BEB-CSP requires knowledge of the operational ranges of BEBs. This is informed by the battery capacity and energy expenditure of the BEB, measured in Kilowatt hour (kWh) and Kilowatt hours per kilometre (kWh/km), respectively. Given the critical role of accurate energy modelling in BEB scheduling, recent literature has increasingly focused on estimating real-world energy expenditure under diverse operating conditions [JBZW22, GMH18, LMH+23]. Several factors have been shown to impact energy consumption, such as ambient temperature, trip time, distance, BEB mass, the frontal area of the BEB, rolling resistance, and auxiliary power draw such as lights and air conditioning. While all of these factors influence energy consumption, the passenger load and total mass of the BEB have been proven to be the most critical when estimating energy consumption. With predictive models' accuracy improving by up to 51.4% when such information is included [JBZW22]. The same work also shows a linear positive relationship between energy expenditure and the total mass of the BEB.

The relationship between energy expenditure and weight is further enforced by another work which uses longitudinal dynamic models to calculate the energy expenditure of BEBs in Singapore [GMH18]. Single-decker BEBs feature an energy consumption of around 1.1-1.2 kWh/km, while larger double-decker and articulated BEBs feature a higher energy expenditure of between 1.6-3.2 kWh/km. Compared to ICEBs, it was found that BEBs are two to three times more energy efficient than diesel vehicles. Finally, a survey reviewed several works related to estimating the energy consumption of BEBs [LMH$^+$23]. It was found that varying mass and road gradients play a large role in energy consumption. Despite this, only 25% of papers reviewed by this survey considered passenger load information.

2.7.4 Non-linear charging

The process of recharging BEBs is also a topic of consideration, as the charging of batteries is non-linear. Batteries are charged using a Constant Current (CC) - Constant Voltage (CV) schedule. In the CC phase, battery charging is linear with time. However, as the SOC nears 100, the terminal voltage of the battery is reached. When this occurs, the CV phase begins, and the relationship between charge time and energy gained is no longer linear [OK20, ZWQ21].

The omission of this relationship can lead to infeasible schedules when charging too high SOC values. However, allowing BEBs to do so can result in a high variance in SOC if the battery has a high Depth of Discharge (DOD) during the day. Where DOD represents the difference between the amount of energy in a battery and the maximum capacity of the battery expressed as a percentage. Repeated charging and discharging cycles, along with high variation in SOC, result in damage to a battery's long-term health. This reduces the overall capacity of the battery over time and leads to a battery reaching its End-of-Life (EoL) when capacity has diminished to 70-80% of its original capacity [ZWQ21, ZMO22].

When modelling non-linear charging of batteries, works have implemented a piecewise linear charging approximation, where two or more sections are created. Each section of the approximation features a linear relationship between charge time and charge rate, with the value for charge rate changing depending on the section [Kar21, LAC20]. However, it should be noted that if capacity is maintained below a certain level (approximately a SOC of 80), then it can be assumed charging is linear with time as the terminal voltage of the battery is not reached [WS17, ZWQ22].

2.7.5 Optimising the Battery Electric Bus Charging Schedule Problem

Objective functions for pure and integrated BEB-CSP are similar to those used for traditional VSP, such as minimising total fleet size, operational costs, operational time, or total distance travelled. However, additional aspects can also be taken into account, such as the fuel price, which includes the fluctuating price of electricity, carbon emissions of the fleet, battery replacement costs, and charging installation price [PLL22, LMH+23, ZWW+24].

Objective functions for the BEB-CSP can generally be split into one of two categories. The first is capital investment, which includes objective functions such as minimising total bus cost, battery sizes, and charger placement. The second is operational costs, which involve minimising fuel costs, passenger waiting times, carbon emissions, and battery replacement costs. Other objective functions can also be found in literature, such as minimising electrical grid fluctuations which may occur during charging operations and peak electrical loads.

Table 2.1: Assumptions made in BEB-CSP literature, such as Mixed-Fleet (MF), Partial Charging (PC), Charger Limit (CL), Pure BEB-CSP (P-BEB-CSP), irregular Energy Consumption (EC), and the Objective function used by each work. Objective functions are categorised as Capital Investment (CI), Operational Costs (OC), and other (O). Works that contain elements, such as the inclusion of renewable energy (R), carbon emissions (C), and passenger waiting times (P) are also denoted.

Work	MF	PC	CL	N-LC	P-BEB-CSP	EC	CI	OC	O	R	C	P
[CX13]			X	X			X		X			
[Li14]			X					X				
[AM17]							X					
[RKW15]	X					X	X	X				
[WLR+16]		X					X	X				
[WS17]		X	X	X				X				
[RvdHLS18]	X		X			X	X	X				
[YLLY20]						X	X	X				
[LAC20]		X			X	X	X					
[JK19]		X	X				X					
[LLX19]	X		X				X	X			X	X
[RPDV20]	X		X					X				
[XYL23]								X	X			
[YY21]	X	X			X	X	X	X				

Continued on next page

Table 2.1 – continued from previous page

Work	MF	PC	CL	N-LC	P-BEB-CSP	EC	CI	OC	O	R	C	P
[ZLZ$^+$22]	X	X			X	X	X	X				
[PDH$^+$21]								X				
[LWL$^+$20]		X					X	X				
[OK20]		X		X		X		X				
[OKW22]	X					X	X					
[ZXZL20]	X	X						X			X	
[TLH19]			X			X		X				X
[KLCF20]			X		X	X	X	X		X		
[CME21]			X				X					
[EMSB22]			X		X		X					
[RRHK20]		X			X			X		X	X	
[MTHA22]								X				
[AAK21]	X	X	X			X	X	X				
[WLL22]		X	X				X					
[ZWQ21]	X	X	X	X				X				
[JZZ21]		X	X					X				
[WLLJ22]			X					X	X			
[ZWQ22]			X	X	X	X		X				
[BHG21]		X				X		X				X
[SCC$^+$22]	X		X				X	X				
[ZAdAR22]		X	X		X			X		X		
[HDLP22]		X	X		X		X	X				X
[LGL$^+$21]		X	X		X			X				
[LSPS21]		X	X		X		X	X				
[STLS22]	X					X		X			X	X
[AC20]		X	X		X		X	X	X			X
[IK19]		X	X				X	X				X
[WLOKF18]	X		X		X		X					
[ZMO22]		X		X	X		X	X				
[ZMOW24]	X	X				X	X	X				
[LYP$^+$24]		X	X		X	X		X		X	X	
[ZWWL22]	X	X	X		X	X	X	X				
[HLS22]		X			X	X	X	X				
[HLS23]	X	X					X	X				
[GN23]		X	X		X	X	X	X				
[WSXW23]		X	X		X		X	X				
[FAMES23]	X	X	X		X		X	X			X	
[GIK23]			X					X				
[JZ22]		X	X				X	X				
[BJWQ21]		X				X	X	X				
[JES19b]		X			X	X			X			

Continued on next page

Table 2.1 – continued from previous page

Work	MF	PC	CL	N-LC	P-BEB-CSP	EC	CI	OC	O	R	C	P
[ALS$^+$20]	X	X			X		X	X		X		
[RNI$^+$22]					X	X		X		X		
[AWAL20]	X						X					
[ETFM19]		X	X			X	X					
[MTZ$^+$20]	X	X	X					X		X		
[JES19a]		X	X		X	X			X			
[JZZ$^+$18]		X	X					X				
[SMGS19]			X		X			X	X	X		

Table 2.1 lists 64 works from contemporary literature for the BEB-CSP along with their assumptions. Where mixed-fleets, partial charging, charger limits, and non-linear charging are represented by MF, PC, CL, and N-LC, respectively. EC represents variable energy consumption, while BEB-CSP that is not integrated with the E-VSP (i.e., pure BEB-CSP) is expressed as P-BEB-CSP.

Objective functions used in these works are classified into Capital Investment (CI), Operational costs (OC), and other (O). All of these functions are minimised, except for two works which maximise some features such as profit or charger occupancy [XYL23, SMGS19]. Additionally, some objective functions included in the other category include profit, charger occupancy, optimising peak load, and total number of charges. Objectives classified as *operational costs* are further divided into those which consider the availability of renewable energy (R), if they consider carbon emissions (C), and additional passenger waiting time considered (P).

From this table, common trends in the literature are more apparent. For example, consideration for some features is divided, with partial charging and charger limits being featured in ≈63% and 61% of papers, respectively. Meanwhile, in less than half of the works examined, operational constraints such as mixed-fleets, pure BEB-CSP, and irregular energy consumption are considered (≈31%, ≈39%, and ≈37% respectively). Where irregular energy consumption represents changes in BEB's energy consumption based on factors such as passenger load or ambient temperatures. The least considered feature of BEB-CSP is the non-linear charging of batteries, with only six out of 64 works modelling it.

Of the works examined, the majority of them (approximately 79%) minimise an objective that can be classified as operational costs, while ≈35% consider both operational costs and capital investment. Of those that consider operational costs, eight use renewable energy generated from sources such as photovoltaic cells. Six works model carbon emissions of the fleet as a part of their objective function, while seven consider passenger impact from potential schedule changes.

The majority of works (≈93%) make use of OC in some form, either charging at bus stops or allowing BEBs to return to the depot to recharge between trips. It was also found that 40 out of 51 works consider electricity tariffs when optimising operational costs, 28 of which use a non-static value for electricity cost. Of these 28 works, 23 represent electricity price time series data with a temporal resolution of one hour or more. These works calculate the Time-of-Use (TOU) of the electricity and use the price corresponding to the time the electricity was consumed.

Of the 50 works that minimise operational costs, 74% use MP to solve the problem in full or in part, or as a comparison to other methods. Works that deal with large instances often use metaheuristic approaches, such as GA, to solve the problem. However, only 40% of the examined works make use of metaheuristic approaches in any way. Therefore, MP and tools such as CPLEX and Gurobi (the most popular choices of MP solvers in literature) are the most common methods of solving the BEB-CSP.

Recalling the statement for this dissertation, the focus is on using renewable energy, minimising carbon emissions, and considering passenger impact when optimising the BEB-CSP.

2.7.6 Renewable energy

As previously outlined, only eight of the examined 64 works explicitly consider the use of renewable energy when solving the BEB-CSP [KLCF20, RRHK20, ZAdAR22, LYP+24, ALS+20, RNI+22, MTZ+20, SMGS19]. Of these, six works focus on photovoltaic (solar power), while the remaining two examine solar and wind power [KLCF20, RRHK20].

The first uses battery swapping technology to create a battery charging and discharging schedule to allow reselling of electricity back to the grid while maintaining operational constraints of a fleet of BEBs [KLCF20]. This work uses a fixed bus schedule discretised into 15-minute increments, with a fully BEB bus fleet. At the same time, non-linear charging was not considered, and battery capacity was capped at 80%, thus avoiding the CV phase of battery charging.

Actual readings from photovoltaic and wind power are considered when evaluating the demand response, which disallows the charging and discharging of batteries when demand is above or below the daily average by a set amount. Using a GA to solve the problem, it was found that reselling electricity back to the grid increased overall costs, as additional batteries were needed. However, if the electricity reselling value was increased by 200%, the system was able to sell power at a profit.

The other work focuses on creating a charging schedule for one BEB operating on a fixed schedule, to minimise electricity costs and carbon emissions, modelled as an LP problem and solved using Gurobi [RRHK20]. It is assumed that the bus begins its operational day fully charged, and explores three scenarios. The first allows the BEB to recharge at any point in the day, the second only allows overnight charging between 20:00–5:30, and the final scenario uses OC after a round trip. Electricity tariff and CO_{2eq} emissions information are separated into 15-minute increments, while other bus schedule information is discretised into 5-minute segments. Results showed that the first scenario obtained the best results in terms of carbon emissions and electricity prices. Overnight charging recorded better electricity prices compared to OC. However, the use of OC showed a larger reduction in carbon emissions. Further sensitivity analysis showed that charging the BEB exclusively from wind and solar power dramatically reduces emissions, at a cost comparable to charging based on a fixed electricity tariff price. This highlights the potential of using renewable sources to power BEBs. The authors conclude the work by stating that precise forecasting of weather and electrical demand is important to create low-emission solutions to the BEB-CSP.

Furthermore, one examined work uses ANNs to estimate several parameters for the BEB-CSP, such as on-site photovoltaic power generation to minimise electricity cost and battery degradation while allowing reselling of electricity to the grid [RNI+22]. Another work models a two-stage optimisation problem, the first of which maximises the occupancy of chargers, while the second minimises grid fluctuations, accounting for solar-generated power [SMGS19]. An alternative objective function for stage two was also examined, which minimises electricity cost and battery degradation. One work solves an integrated BEB-CSP while minimising daily electricity costs [MTZ+20]. Power generated by on-site solar panels is prioritised, the availability of which is estimated via historical solar data, and it is noted that a rolling horizon approach could be investigated to deal with stochastic solar generation. An additional work creates a MILP model with a CPLEX solver to maximise the profit of a bus depot operator by selling electricity generated via on-site solar panels to BEB owners [ALS+20]. Where the projected salvage costs of batteries and photovoltaic units are considered.

Historical solar irradiance data have been used to calculate solar-generated power, where the focus is placed on minimising the sum of electricity costs, carbon emissions, and energy storage costs for a pure BEB-CSP [LYP+24]. OC at depots is allowed within a 30-minute charging block, with variable charging power delivered to the BEB. Electricity price TOU information is split into six different sections throughout the day to represent *peak* and *off-peak* prices. A MILP model is solved using Gurobi with a dataset of 50 BEBs in 1.05 seconds. While accommodations are made for a mixed-fleet of BEBs with different battery capacities, it is considered an input parameter and not a decision variable. The energy consumption of BEBs is calculated using a longitudinal dynamics model, using historical data from Global Positioning System (GPS) trajectory points and the speed of the vehicle.

2.7.7 Carbon emissions

Only a handful of works in contemporary literature consider the lifecycle emissions of a BEB fleet when solving the BEB-CSP [FAMES23, LYP+24, STLS22, RRHK20, LLX19, ZXZL20]. Two of which have already been discussed in the section prior.

The third work solves a pure BEB-CSP using a fleet of only BEBs, with the battery capacity of each bus modelled as a variable [FAMES23]. This work minimises the total annual cost, taking into account capital costs (such as charging station infrastructure, BEBs with their batteries, and transformer costs), maintenance costs, cost of electricity, and cost of the fleet's lifecycle emissions. Electricity costs and CO_{2eq} emissions are aggregated into an hourly temporal resolution, with the bus schedule discretised into time slots of 2-minutes. The cost of emissions is modelled as \$65 per Tonne of Carbon Dioxide Equivalent (tCO_{2eq}). The resulting model was evaluated with a dataset of 55 BEBs and 506 bus stops, and the modelled ILP was solved using Gurobi, with one section of the problem solved using a metaheuristic approach called Particle Swarm Optimisation (PSO). It was highlighted that considering electricity costs and GHG emissions leads to significant reductions in operating costs and a reduction of 9058.25 tCO_{2eq} per year after fully transitioning to a BEB fleet. Further sensitivity analysis showed that considering the TOU of electricity and emissions results in 13.12% lower costs compared to using a fixed average value. It is highlighted that while energy consumption at a trip level was considered, variations due to passenger load were not. The modelling of CO_{2eq} as a monetary value is further explored in Section 2.9.

The next work creates an ILP solved with CPLEX for an integrated BEB-CSP with a mixed-fleet of BEBs and ICEBs [LLX19]. Lifecycle emissions and potential impact on passengers are taken into account, with lifecycle emissions of BEB modelled as a static value per kWh. Additional waiting time passengers spend is considered in the objective function, with a passenger Value-of-Time (VOT) assigned a value of $0.3 per 30 minutes. If passengers are left waiting for too long, it is assumed they find alternative transportation, penalising the system by $10 per passenger. Time is divided into 30-minute segments. Meanwhile, battery capacity is also discretised, with different values explored. Modelling battery capacity as a discrete domain leads to reductions in computation time at the cost of overall solution quality. Furthermore, schedules are created for a horizon, the longest of which is 12-hours in length. An empirical evaluation of a dataset from Hong Kong revealed that the bus fleet size could be reduced by 24%, resulting in a 8% decrease in operational costs and a 24% reduction in carbon emissions. Passenger VOT is further examined in Section 2.10.

Tailpipe emissions and electricity costs are considered in another work solving an integrated BEB-CSP with a mixed bus fleet of BEBs and ICEBs [ZXZL20]. A dataset containing 575 trips, originally serviced by 79 ICEBs, had half of the buses replaced with 52 BEBs, bringing the fleet total to 92. For electricity costs, TOU information was split into *peak*, *off-peak*, and *mid-peak* times. Only the emissions of ICEBs were considered (as BEB lack tailpipe emissions), however, both objectives were optimised separately. Considering TOU of electricity led to a reduction of up to 13% in charging costs. Meanwhile, if carbon emissions were optimised, the distance travelled by BEBs increased between 36% and 46%.

The final work solves an integrated BEB-CSP to minimise carbon emissions and operational costs, where operational costs take into account fuel cost and travel cost of passengers [STLS22]. Additional passenger waiting time due to BEB usage is modelled along with their VOT. Furthermore, the energy expenditure of BEBs is estimated based on the weight of the passengers onboard. Both carbon emissions and fuel costs are modelled as static values for each type of bus examined, and the optimal assignment of BEBs to replace ICEBs is determined.

The proposed model is used to solve an instance with 114 bus lines and 838 stations, and the trade-off between reducing carbon emissions and operational costs is explored. It was determined that a carbon emission reduction of up to 77.04% was possible (207.15 to 47.56 tCO_{2eq}), however, this results in increased operational costs.

2.7.8 Passenger impact

Of the works examined, seven explore the impact on service quality due to additional constraints associated with BEBs [STLS22, LLX19, IK19, HDLP22, BHG21, TLH19, AC20]. Two of these works have been previously discussed in Section 2.7.7 [LLX19, STLS22].

One work models an integrated BEB-CSP with charging infrastructure placement [IK19]. Where passenger cost and operator cost are minimised and solved using PSO. Charging occurs while passengers embark and disembark from the BEB. If charging time exceeds this, a cost penalty is applied considering the passenger's VOT (which is assigned as 5$/hour). When balancing the two objectives in a generalised cost function, the operator costs were prioritised over passenger costs. As a result, it was outlined that the selection of parameter values, such as VOT, greatly affects the results of the BEB-CSP solution. Another work solves a similar problem of minimising charger infrastructure while accounting for passenger delays [HDLP22]. This work also considers the TOU of electricity to solve a pure BEB-CSP with a fully electric fleet. A MILP formulation is used with Gurobi to solve an instance with three bus routes and 111 stops. Electricity TOU is split into *peak*, *mid-peak*, and *off-peak costs*, and passenger consideration is modelled as a penalty on increased waiting time, subject to weights. It was found that charging times did not exceed passenger boarding times in most cases. It was also noted that the model could be extended to consider mixed-fleets and allow dynamic adjustments to the charging schedule.

A separate work optimises the weighted sum of passenger costs, electricity costs, and BEB depreciation [BHG21]. Passenger costs are modelled by the VOT of passengers and any extra waiting time at a bus stop, where embarkation and disembarkation times were also accounted for. Electricity costs are split into *peak*, *off-peak*, and *mid-peak* prices. An empirical evaluation of a single bus route with 25 stops is done, with a VOT value of 0.21 Chinese Yuan per minute. Results showed that overall passenger travel time cost was reduced by 15.1%, while also achieving an electricity price reduction of 19.5%. The final work represents stochastic traffic conditions and minimises operational costs [TLH19]. Each day is split into periods, where the operational costs of the current period are optimised, as well as the estimated cost of any following period. Any delayed trip start times caused by the changes in traffic conditions are penalised by a quadratic function. Evaluation of a dataset with up to 96 trips shows that if sufficient idle times are added between each trip, then any delays caused by stochastic traffic conditions can be absorbed.

The final work solves a pure BEB-CSP for a fleet consisting of only BEBs, where the total number of charges, number of charging stations, and deviation from the original schedule is minimised [AC20]. Using an MIP model, the total number of chargers was optimised separately from the number of charges and deviation. Weights were then used for the remaining components to create a combined cost objective function. A weight of 10^7 was used for timetable deviation, while a weight of 1 was used for minimising the total number of charges, as a result, deviation from the original timetable was heavily penalised. Interestingly, time is not a discretised variable, and is continuous instead, representing an hour decimal format (i.e., a timetable deviation of 1-minute is represented as 0.01667). Therefore, a 1-minute delay is the equivalent of 166,700 charges. As a result, the average delay per stop was reported as 2–5 minutes depending on the dataset used.

2.8 Predicting and optimising

While COPs aim to optimise an objective function, it is often the case that some aspect of the problem is unknown or stochastic. Works previously described in Sections 2.7.6, 2.7.7, and 2.7.8 generally make use of historical information when generating schedules to know the optimal time to schedule charging events. However, such information as electricity TOU or CO_{2eq} is not known ahead of a scheduled time and must be predicted.

To resolve this, prediction models can be used to estimate values for these unknown variables and pass them to a COP. An example of this has already been discussed where ANNs were used to estimate available solar power and electricity costs for a BEB-CSP [RNI+22]. This combination of predictive and optimisation models has also been applied to other domains, such as estimating weights in the weighted knapsack problem [MDSG20] and scheduling of combinatorial problems with uncertain duration times [DAD18]. These approaches can be classified as *decoupled* methods.

However, in recent years, the paradigm of *Predict and Optimise* has gained attention, where instead of creating a prediction and COP model separately, aspects of the COP model are used in the training of the prediction model [EG22]. Approaches which combine the training of prediction models with the COP estimations are used in can be classified as *combined* methods [ALKS+25].

The reason for this distinction is that an increase in accuracy for the predictive model may not improve the quality of a COP solution. In some cases, regression predictive models with better MSE scores result in COP solutions with worse optimas compared to the same problem when provided with the output of a predictive model traditionally considered *worse* [IOS12]. The work that highlights this focuses on predicting electricity TOU for a scheduling problem to optimise total energy costs. After this finding, an alternative approach is proposed where the predictive model is used to estimate the classification of energy prices as *peak* or *off-peak*, with *peak* prices being classified as any value above the 66^{th} percentile. This, in essence, creates a classification problem (similar to those discussed in Section 2.3.6) where it was found that improving the loss of the classification models does improve the optima for the scheduling problem.

The previously discussed work was further expanded to investigate potential ways of improving the solution quality of COPs while keeping a regression prediction model to optimise the TOU of electricity tariffs [GIOS14]. Evaluation with two COPs showed that the inclusion of price rankings (such as Spearman's rank correlation) along with traditional regression metrics, such as MAE, when training predictive models, helps to boost the solution quality of the optimisation problems examined.

These works showcase the primary issue with *decoupled* approaches. However, combined approaches are not without issues. While incorporating the COP into the training of the predictive model allows the estimations to optimise the optimisation problem, its inclusion may increase computation time, especially if said problem is NP-hard [ALKS+25, EG22]. Therefore, *decoupled* approaches such as the ones previously described may be more practical. As consideration of how the prediction model's output is used in the COP, reformatting the input data, and choosing a loss function in an informed way, would be less complex than adding the COP into the training process for the prediction model. This would also potentially allow the reuse of the prediction model in analogous

COPs.

2.9 Carbon intensity and Social Cost of Carbon

As outlined in Section 2.7.7 the minimisation of carbon emissions from a fleet of BEBs is an objective that has been considered by some works. Generally, works that optimise carbon emissions of BEB fleets fall into one of three categories. The first category does not model the lifecycle emissions of BEBs and compares the tailpipe emissions of ICE vehicles vs BEBs [ZXZL20, SIKF24]. The second does models lifecycle emissions of BEBs as a constant variable that does not fluctuate throughout the schedule [LLX19, STLS22]. The final category models lifecycle emissions of EVs as a variable that can change throughout the day via TOU constraints such as the ones employed for fluctuating electricity prices [RRHK20, DBEB20, FAMES23].

However, such TOU constraints are rarely applied to fleets of BEBs to reduce CO_{2eq}. When modelled for BEBs, emissions are often represented as Grams of Carbon Dioxide Equivalent per Kilometre (gCO_{2eq}/km) [STLS22, MBEH22]. However, the use of this metric may lead to issues when including other objectives in an objective function, as it requires the introduction of weights for each objective. These weights would then have to be subjected to parameter optimisation. To resolve this, some works model the carbon emissions of EVs as a financial cost [LLX19, FAMES23], which can also be referred to as the Social Cost of Carbon (SCC) [HMMY21]. This can be defined as:

Definition 2.9.1 (Social Cost of Carbon). *The Social Cost of Carbon is the measurement of costs related to the burning of carbon and is modelled as the financial cost per tCO_{2eq} burned.*

SCC may also be referred to as the Shadow Price of Carbon (SPC) and is regularly used to inform policies such as carbon taxes. It is often generated via expert opinion and economic growth models such as the Dynamic Integrated model of Climate and the Economy (DICE) model [Pin19, BN24]. This is reflective of the estimated loss in Gross Domestic Profit (GDP) that can come about from the displacement of people and rising sea levels.

Some governments publish reports outlining the SCC they plan to use in the future for forthcoming policies [Dep24c]. This has been adopted by some works in literature to model carbon emissions of BEBs [LLX19, IL19].

2.10 Public transport user's Value of Time

A majority of works that focus on scheduling fleets of BEBs do not take the impact of customer service into account when generating new schedules. Generally, works set a fixed start time for trips which buses in the fleet must service, ensuring that the service quality is no worse than any prior schedule [ZXZL20, SCC+22]. However, some works, such as those discussed in Section 2.7.8, aim to optimise service quality by modelling it as a part of the objective function [LLX19, HDLP22, STLS22, IK19, BHG21, TLH19, AC20].

There are two methods of modelling service quality. The first is to model passenger wait time, in-vehicle time, or timetable deviation as a part of the objective and use an associated weight to balance it with other objectives such as operational costs [HDLP22, AC20]. However, such weights are highly influential to the solution, and values must be chosen in an informed manner to avoid overpowering other aspects of the objective function. An alternative approach is to model passenger VOT, which represents the cost that passengers are willing to pay for a service [JCCL12, STLS22]. Such metrics are important parameters used to inform government policies, such as the development of new transportation routes [Dep24c]. This, in effect, can be used as a weight informed by real-world practices that can be used to balance service quality with other objectives such as operational costs.

Works that solve the BEB-CSP considering VOT base the value on national GDP or residents' incomes [IK19, BHG21, LLX19], however, these do not take into account passengers' perception of delays. When new bus schedules are created, buses may arrive at some stops early or later than their original timetable. This impacts service quality and passengers' VOT, as public transport users will spend more or less time on the bus compared to the original schedule. However, such changes to scheduled arrival times do not have a symmetrical impact on passengers' VOT, and require consideration of VOT modifiers. For example, if a bus is scheduled to arrive one minute later than the original schedule at a destination stop, the penalty on passengers' VOT will be higher compared to if the bus arrives one minute early to an origin stop. Even though in both scenarios, passengers are spending one additional minute on public transport [WCdJ16].

Therefore, to fully account for how changes in bus schedules impact passengers, both VOT and modifiers should be taken into account. This is especially important when modelling a problem which uses a combined cost objective function.

2.11 Conclusion

This chapter has outlined several definitions, and basic knowledge of the methods used has been shared. In addition, contemporary literature has been discussed, and as a result, some gaps in the literature can be identified which are filled with works outlined by the contributions of this dissertation outlined in Chapter 1 Section 1.2.1 (Page 5).

Contribution 1 relates to the creation of prediction models to be used in conjunction with the BEB-CSP to optimise carbon emissions. As outlined by one work, the lifecycle emissions of BEBs are lowest when powered solely from renewable energy [RRHK20]. The largest source of renewable energy in Ireland is from wind. Therefore, a prediction model that was able to estimate available wind power would be useful for Contribution 2. However, as noted in Section 2.8 (Page 41) an improvement in accuracy in a prediction model may not result in improved solution score for a COP. Therefore, aspects of the COP used for Contribution 2 should be considered for the prediction model of Contribution 1. This can be done by reformulating the training data of the prediction model to represent when large amounts of wind energy are available (i.e., a wind power excess). However, this turns the problem into an imbalanced regression such as discussed in Section 2.3.6 (Page 17). While methods have been outlined to resolve issues with imbalanced datasets, certain combinations of methods have not been evaluated for use with DL models. As such an empirical evaluation of these methods with DL models would be a novel contribution. The implementation of this is outlined in Chapter 3.

Contribution 2 relates to modelling a pure BEB-CSP to reduce carbon emissions. As previously outlined, renewable energy can be used to keep carbon emissions of a fleet of BEBs low, therefore the objective of minimising non-renewable energy used both reduces the amount of CO_{2eq} and is not explored by any of the 64 works outlined in Table 2.1 (Page 33). Furthermore, only one examined work uses predictive models to estimate values of unknown variables for use in the BEB-CSP [RNI+22]. The aforementioned work does not do a comparison between the BEB-CSP solution using the predictions vs using actual values. Therefore, an empirical evaluation of the effectiveness of prediction models for this domain would be novel. These topics are further discussed in Chapter 4.

Finally, Contribution 3 involves the extension of the MILP model created in Chapter 4 to model a mixed-fleet pure BEB-CSP with non-linear charging, partial charging, charger limits, and non-linear charging with consideration for passenger load's effect on BEB energy consumption. Furthermore, both passenger VOT, electricity, and CO_{2eq} TOU are considered for the objective function of the problem. As seen from Table 2.1 (Page 33), no other work has modelled all 7 aspects previously mentioned. Additionally, no works have considered both passenger VOT in conjunction with its modifiers and works that model CO_{2eq} emissions using SCC only explore one value. Furthermore, works that do explore the consideration of lifecycle emissions and passenger impact do not explore large datasets. The implementation for Contribution 3 empirically evaluates a dataset with 65 bus routes, 1169 buses, and 4279 unique bus stops. As such, a pure BEB-CSP modelled with the outlined aspects, using an objective function considering VOT and modifiers with electricity and carbon emission TOU, evaluated on a large dataset, would be a novel contribution. This contribution and its implementation are outlined in Chapter 5.

Chapter 3

Prediction of Imbalanced Resource Availability for Scheduling Constraint Optimisation Problems

3.1 Introduction

As outlined in Chapter 2 Section 2.7.6 (Page 36), the lifecycle emissions of a Battery Electric Bus (BEB) are lowest when charged solely from renewable energy. Therefore, a predictive model trained to estimate the amount of wind-generated power available would be useful when formulating Contribution 2. However, as discussed in Chapter 2 Section 2.8 (Page 41), a prediction model trained without knowledge of how its estimations are used can result in poor solutions for a Constrained Optimisation Problem (COP) such as the Battery Electric Bus Charging Scheduling Problem (BEB-CSP). To avoid this, a prediction model should be trained considering aspects of the BEB-CSP. To this end, instead of estimating wind-generated power, the amount of wind power excess is predicted. This is done as the BEB-CSP implemented for Contribution 2 prioritises the charging of BEBs when there is an excess of wind-generated power.

Data from the Irish national grid is examined from 2013 to 2022. It should be noted that the amount of wind-generated energy does not exceed national demand. Therefore, there is no wind power excess. However, wind-generated power is Ireland's largest source of renewable energy. Wind power capacity is expected to increase by $\approx$40% between 2021 and 2026 with the development of additional wind farms [Dep24a][1]. If wind-generated power data is scaled up by 40%, a wind power excess is possible. This raises a new problem, as the resulting wind power excess dataset is an example of an imbalanced regression dataset as discussed in Chapter 2 Section 2.3.6 (Page 17).

While methods exist to improve the accuracy of models trained with imbalanced regression datasets, there is a gap in the literature regarding an empirical evaluation. In particular, imbalanced regression evaluation metrics and resampling methods for Deep Learning (DL) approaches, such as Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), and Convolutional Neural Network (CNN) models, as well as their hybrids. To this end, Contribution 1 as detailed in this chapter performs such an evaluation, filling this gap in the literature.

This chapter is divided into the following sections. First, Section 3.2 discusses the dataset for electricity demand and wind-generated power, along with supplementary datasets. The upscaled wind-generated power dataset is defined, the wind power excess dataset is created and its imbalanced nature is discussed. A relevance function for the imbalanced regression problem is also created. Section 3.3 explains the resampling methods investigated, while Section 3.4 discusses the two imbalanced regression loss metrics used.

In Section 3.5 the DL configurations examined are described along with the hyperparameter optimisation which takes place. Then prediction of the imbalanced regression dataset is done, and an empirical evaluation of the configurations, evaluation metrics, and resampling methods is done. Finally, Section 3.6 concludes the chapter and ties the work contained within back to the thesis statement and Contribution 1.

3.2 Demand and wind datasets

Data for electrical demand (Demand) and wind-generated power (Wind) of the Irish national grid are publicly available online[2]. In addition to the Demand and Wind datasets, data for electricity generation (Generation), Carbon Dioxide Equivalent (CO_{2eq}) intensity (Intensity) and CO_{2eq} emissions (Emissions) are gathered. Demand, Generation, and Wind are all measured in MegaWatt (MW), while Intensity is represented as Grams of Carbon Dioxide Equivalent

[1] Table 11 *Trajectories by renewable energy technology With Additional Measures* of Ireland's integrated National Energy and Climate Plan 2021-2030. The increase in wind energy generation between 2021 and 2026 is 39.7%, rounded up to 40%.

[2] `https://www.smartgriddashboard.com/`

per Kilowatt hour (gCO_{2eq}/kWh), and Emissions are Tonne of Carbon Dioxide Equivalent (tCO_{2eq}) per hour. Data is represented in a univariate time series format with a temporal resolution of 15-minutes, and collected between November 2013 and October 2022 inclusive. After cleaning the data and interpolating missing values, there are a total of 313,046 data points for each dataset.

Table 3.1: Mean, median, standard deviation (STD), min, and max values for the datasets used.

Dataset	Mean	Median	STD	Min	Max
Demand (MW)	3245.42	3297.0	634.28	1664.0	5363.0
Generation (MW)	3237.36	3200.0	380.39	1494.0	5993.0
Wind (MW)	930.83	754.0	741.82	0.0	3604.0
Intensity (gCO_{2eq}/kWh)	352.17	349.0	102.62	125.0	700.0
Emission (tCO_{2eq}/hour)	1149.66	1138.0	336.62	450.0	2430.0

Table 3.1 shows the metrics for each dataset, including the mean, median, Standard Deviation (STD), minimum, and maximum values. From this table, a strong relationship between Demand and Generation can be seen, with little variation in the mean, median, minimum, and maximum values. Note that the maximum value for Generation is larger than the maximum value for Demand. Events such as this occur 155,986 times in the dataset and indicate Ireland's energy exportation to Wales via the East West Interconnector and the Greenlink. However, exporting of energy is not the topic of this chapter and will not be discussed further.

The metrics for Wind show it has low mean and median values compared to Demand and Generation, with the lowest reading showing a complete absence of wind-generated energy. However, this table also shows the dataset has the highest STD of all those examined. Additionally, the maximum value of the dataset is 67.2% of the largest value for Demand. This shows that Wind is highly variable, with a STD 79.69% of the mean value of the dataset, and may be difficult to predict. Both Intensity and Emissions showcase lower STD values, both approximately 29% of the mean value for each dataset. Intensity, Emissions, and Wind should have a close relationship. Therefore, the inclusion of carbon intensity and emissions should aid with estimating wind power. This will be discussed in more detail later.

All of these datasets are aggregated into a single multivariate dataset, in addition, the wind-generated power excess time series data (Excess) is calculated. This dataset is generated by determining the difference between Wind and Demand at each time-step for all data points. The resulting time series has a mean of -2314.58, a median of -3240.0, an STD of 842.82, and minimum and maximum values of -5115.0 and -60.0, respectively. This shows that wind-generated energy can come close to meeting demand, with the lowest difference between the two being 60 MW. Furthermore, the STD of this time series is much lower compared to Wind (36.4% and 79.6% of the respective mean value).

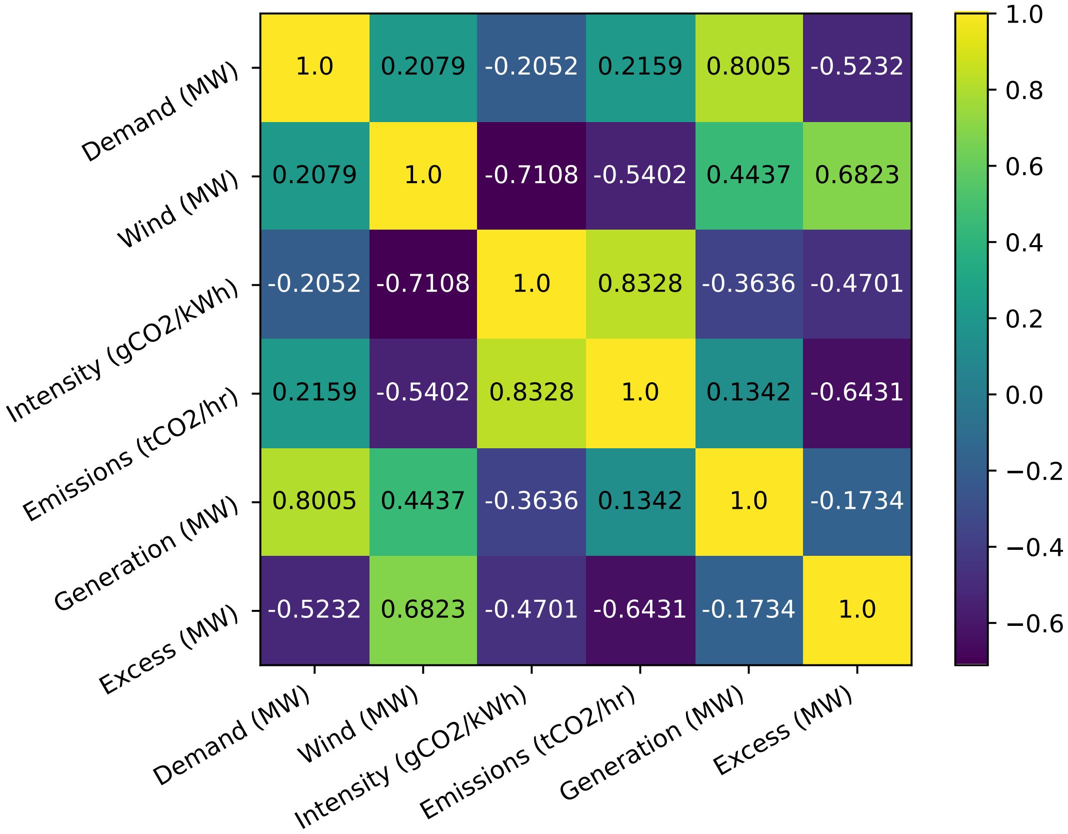


Figure 3.1: Spearman's rank correlation coefficient matrix of dataset features

Figure 3.1 shows the Spearman's rank correlation coefficient matrix of features for the multivariate dataset, where a value closer to 1 shows a strong positive monotonic relationship, while a value closer to -1 shows a strong negative monotonic relationship between two features of the multivariate dataset. As shown, Generation and Demand have a strong relationship as expected. Similarly, Wind has a strong negative correlation with Intensity and Emissions, as the more wind power in the system, the less CO_{2eq} the national grid will produce. Meanwhile, Excess shows a strongly positive relationship with Wind, and strong negative correlations with Demand and Emissions. Intensity also indicates a moderate relationship with Excess. However, the relationship between Generation and Excess is weak. These correlations motivate the inclusion of

Intensity and Emissions to assist in estimating wind-generated power excess.

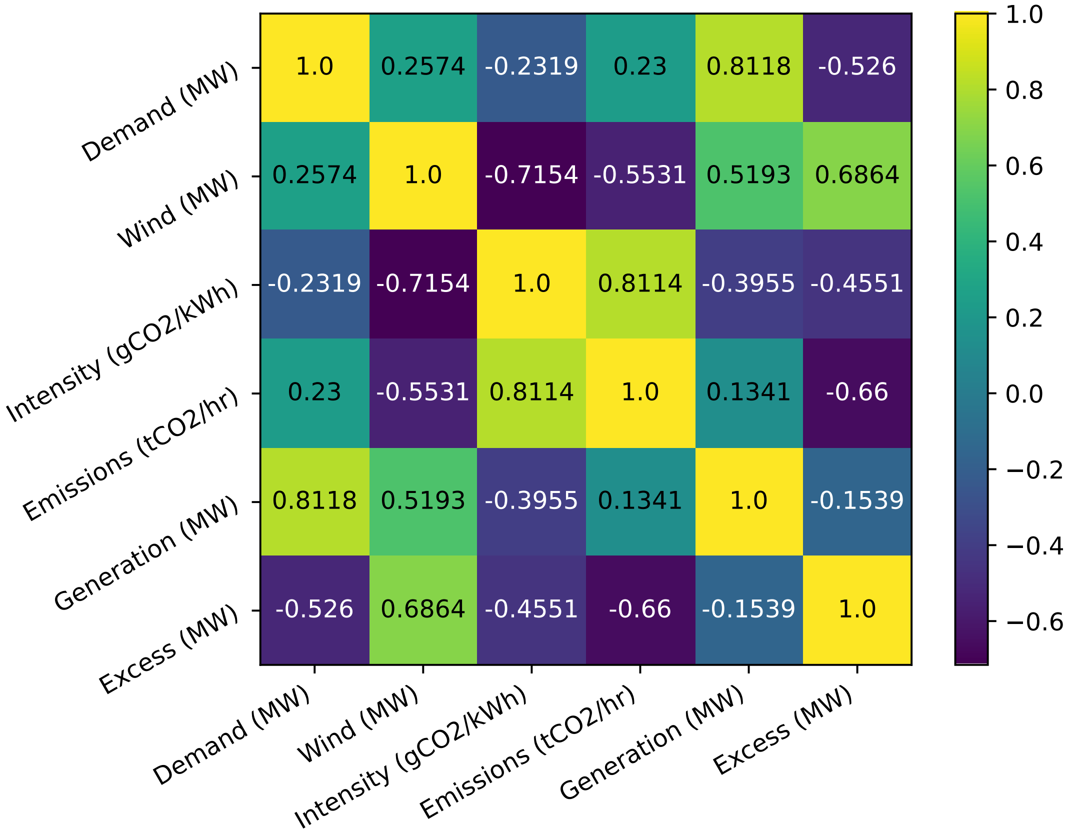


Figure 3.2: Pearson's correlation matrix of dataset features

Figure 3.2 shows the Pearson's correlation for the multivariate dataset, where a value closer to 1 or -1 represents a strongly positive or negative linear relationship between features, respectively. Here, it can be seen that the target variable, Excess, does not have any strong linear correlations with the other features in the dataset. Most features showcase only a moderate linear correlation, with Generation maintaining a potentially non-linear relationship with Excess. This lack of strongly linear relationships between features indicates that non-linear modelling approaches (such as DL) may perform better for a predictive model.

Attention now shifts to the upscaling of Wind based on Ireland's future wind generation expansion plans. Between the years 2021 and 2026, the Irish national grid is expected to increase the capacity for wind-generated power by ≈40% with the construction of both onshore and offshore wind farms [Iri24]. This has provided key motivation to explore the upscaling of the available wind-generated power. To this end, all data points in the Wind time series are multiplied by a value of 1.4 to generate an Upscaled Wind dataset, thus emulating the 40% increase in wind-generated power capacity. The Upscaled Wind values are then used to calculate the Upscaled Excess using the same method as the generation of the Excess dataset.

Table 3.2: Mean, median, STD, min, and max values for Upscaled Wind and Upscaled Excess.

Dataset	Mean	Median	STD	Min	Max
Upscaled Wind (MW)	1303.17	1055.6	1038.55	0.0	5045.6
Upscaled Excess (MW)	-1942.24	-2027.2	1068.51	-5097.0	1074.4

Table 3.2 shows metrics such as the mean, median, STD, minimum, and maximum values for the Upscaled Wind and Upscaled Excess datasets. The change in metrics for Upscaled Wind is unremarkable, as all values are an increase of 40% from the values found in Table 3.1. However, the values for Upscaled Excess are notable, while the mean and median remain firmly below 0 (i.e., representing a wind-generated energy deficit). The maximum value for the dataset showcases a significant excess in wind-generated energy. The STD of Upscaled Excess has also shown an increase from 36.4% before the wind modifier is applied, to 55.01%, showcasing larger deviations for the mean value of the dataset.

Figure 3.3 shows the Spearman's rank correlation coefficient matrix for upscaled Wind and Upscaled Excess and the other features of the multivariable dataset. Here, it can be seen that while the monotonic relationship between Upscaled Excess and Upscaled Wind, Intensity, and Excess is stronger compared to those featured in Figure 3.1, the relationship between Demand and Generation worsens. While Demand maintains some negative monotonic relationship with Upscaled Excess, the correlation between Generation and Upscaled Excess is negligible.

Figure 3.4 shows the Pearson's correlation for the upscaled dataset. Here it can be seen that the linear relationship between Upscaled Excess and Upscaled Wind is much stronger compared to Excess and Wind from Figure 3.2. However, the linear relationship between Demand and Upscaled Excess is much weaker compared to Demand and Excess. It should also be noted that there is no significant linear relationship between Generation and Upscaled Excess. Despite this, initial feature selection analysis shows that the inclusion of this feature improves results. Therefore, all five features (Demand, Upscaled Wind, Intensity, Emissions, and Generation) are used to estimate the available upscaled wind power excess (Upscaled Excess).

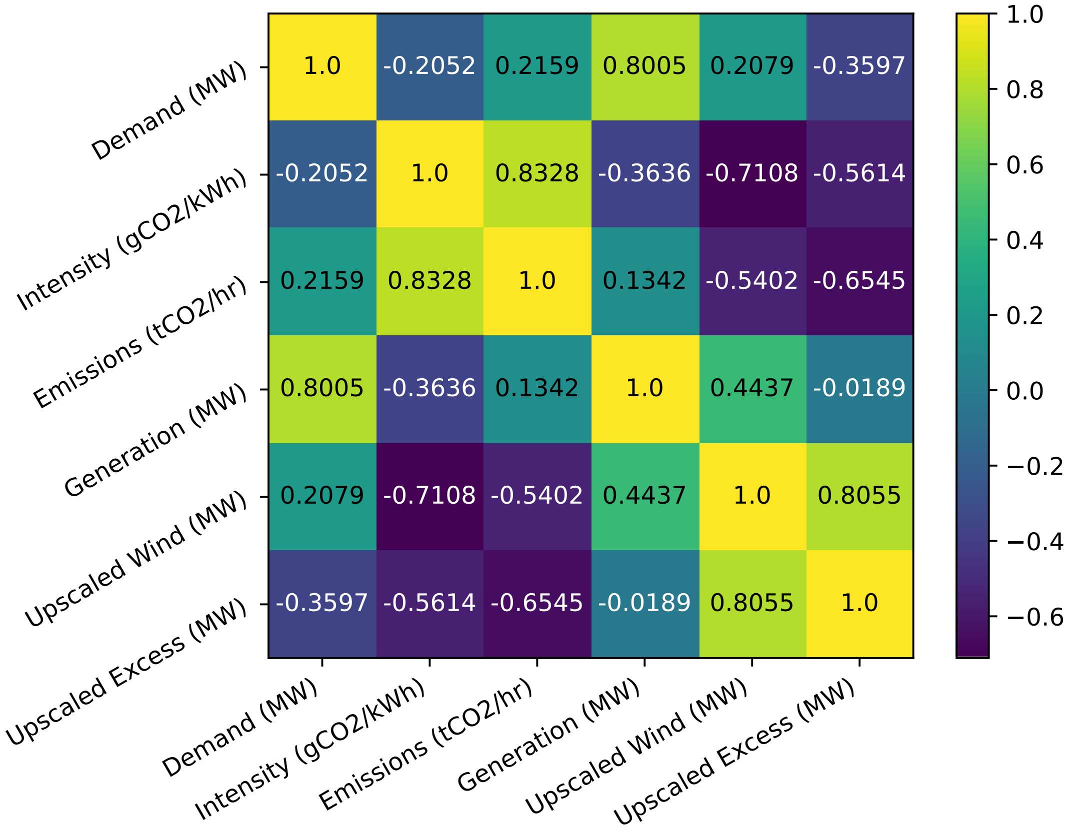


Figure 3.3: Spearman's rank correlation coefficient matrix with Upscaled dataset features

Ultimately, the estimations of the prediction model are intended to be given to a COP to denote when an excess of wind-generated power is available. Representing the estimation target as Upscaled Excess instead of Wind or normal Excess allows the loss function of the prediction model to optimise the model's predictions considering aspects of the COP. However, only Upscaled Excess values above 0 represent a wind-generated power excess, while values below 0 denote a deficit. Therefore, for the prediction model to truly be trained based on how the COP uses its estimations, more importance must be placed on accurately predicting values above 0.

Figure 3.5 shows the probability density of the outlined dataset, where a vertical line shows the cut-off point of values which are passed to the COP. Here, it can be seen that a new problem arises, as only 12,722 data points out of 313,046 exhibit an Upscaled Excess value above 0, representing an imbalance where more importance is placed on values between the 95.93$^{\text{th}}$ and 100$^{\text{th}}$ percentile of the dataset. This creates a problem of an imbalanced regression dataset. Assuming the final year's worth of data (November 2021–October 2022) is reserved as a

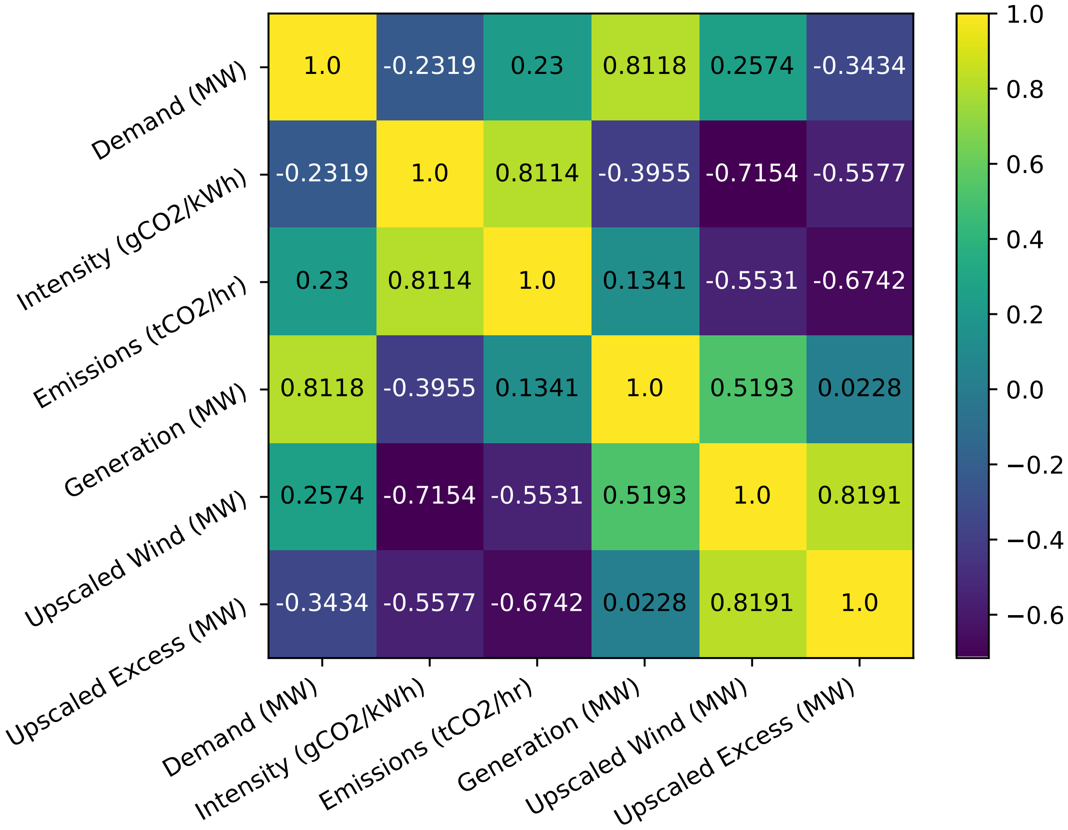


Figure 3.4: Pearson's correlation matrix with Upscaled dataset features

test dataset, the training dataset contains 278,002 data points. Of these, 9602 datapoints represent Upscaled Excess values above 0, constituting the 96.54$^{\text{th}}$ percentile of the training dataset and above.

Figure 3.6 shows the Pearson's correlation matrix for all features in the upscaled dataset where Upscaled Excess is above 0. Here, it can be seen that the linear relationships between some features have disappeared. The highest value indicates only a moderately linear relationship at best (Upscaled Wind and Upscaled Excess), suggesting that the relationship between other features and Upscaled Excess for the target values above 0 is non-linear. Thus, non-linear models, such as DL, are more appropriate compared to linear models.

3.2.1 Relevance and imbalanced regression

Imbalanced regression datasets and methods on how to improve the error of predictions of lesser represented values are discussed in Chapter 2 Section 2.3.6 (Page 17). Highlighted is the use of relevance functions $\phi(y)$, which, given a regression value y, returns how relevant the value is. In general, values closer

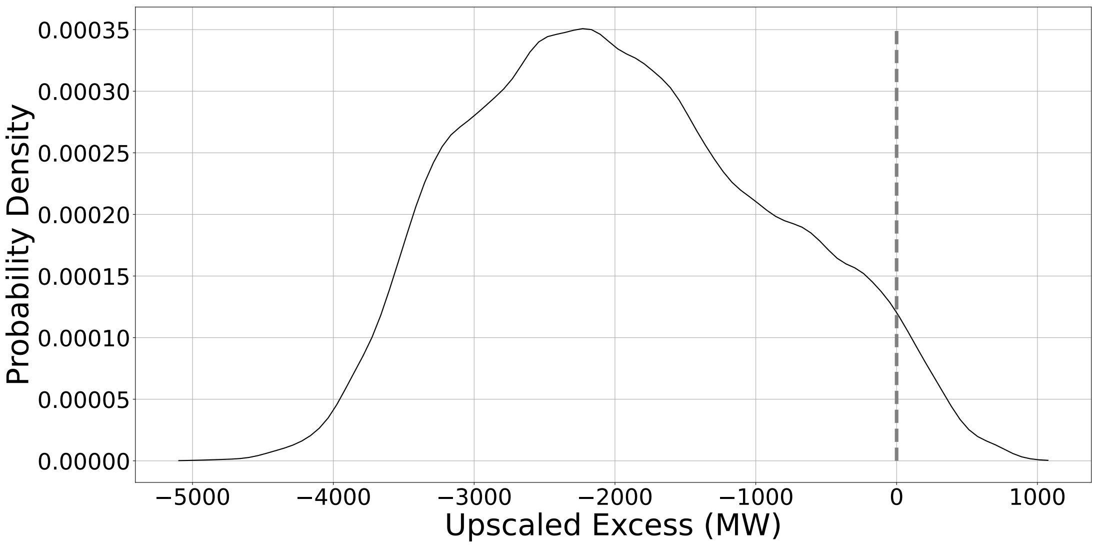


Figure 3.5: Probability Density Function of Upscaled Excess (MW). Where the vertical line represents values of interest to the COP

to statistical outliers are assigned a relevance value closer to 1, while values around the median are assigned relevance values near 0. One of the most popular methods of generating a relevance function is Piecewise Cubic Hermite Interpolating Polynomials (PCHIP) [RM20]. PCHIP assigns a relevance value to all data points using the relevance scale (i.e., the minimum and maximum relevance values), the median data point value, and upper and lower adjacent values.

Figure 3.7 shows the relevance function for the training dataset Upscaled Excess time feature using scipy's *pchip_interpolate* function[3]. It should be noted that only regression values closer to the upper adjacent are given higher relevance values, while those closer to the lower adjacency have a value of 0. This is not normally the case for relevance functions and represents the problem to be solved (i.e., the estimation of regression values above 0). Another noteworthy aspect of this relevance function is the scaling of the relevance value. Ordinary relevance values are of some range $\phi(y) : y \to \{0, 0.01, \ldots, 1\}$, however in this figure it is $\phi(y) : y \to \{0, 1, \ldots, 100\}$. This is an intended change of the function's ordinarily floating point representation of relevance values to an integer where values are rounded down to the nearest whole number. This aids the implementation of DL loss functions explained later in this chapter.

As can be seen from this figure, regression samples with values at the median

[3]`https://docs.scipy.org/doc/scipy/reference/generated/scipy.interpolate.pchip_interpolate.html`

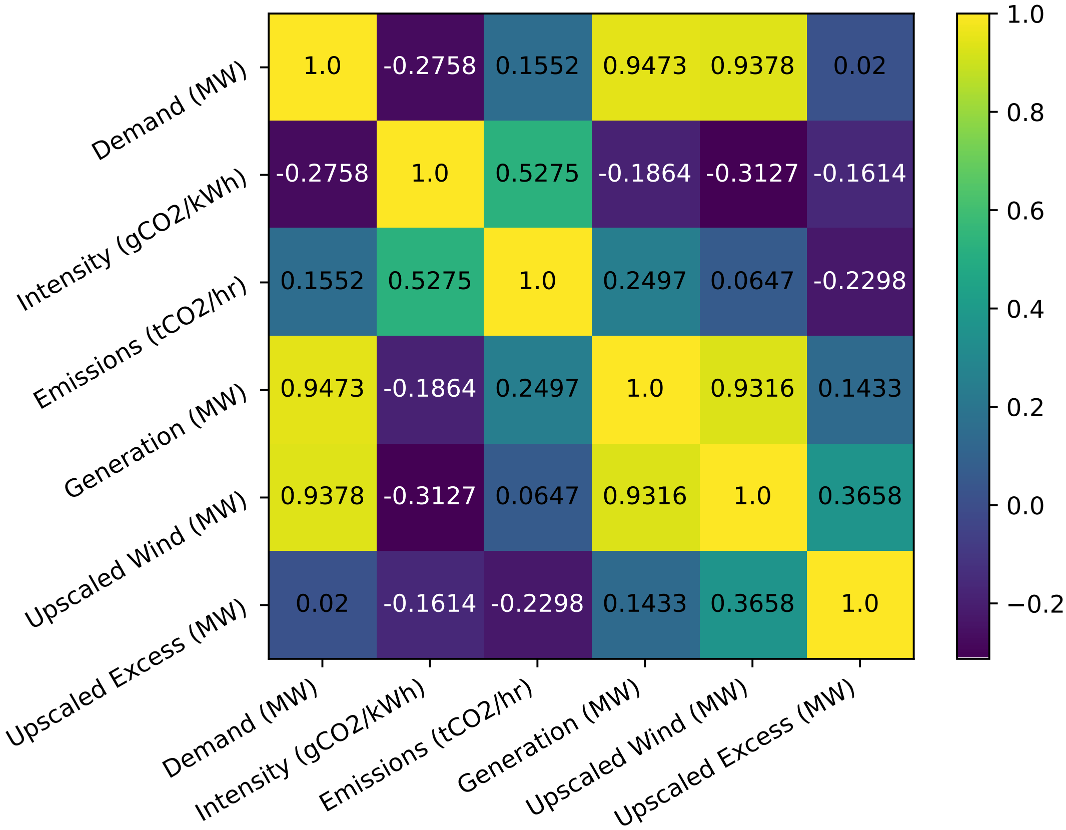


Figure 3.6: Pearson's correlation matrix with Upscaled dataset features for Upscaled Excess values above 0

and lower adjacency are assigned a relevance of 0, while those around 0 are assigned the value ≈40. This relevance function is used in conjunction with loss metrics and resampling strategies used throughout the test of this chapter to place more importance on the prediction of regression values above 0.

3.3 Data resampling

One method of dealing with imbalanced regression data is to use data resampling. As discussed in Chapter 2 Section 2.3.6 (Page 17), data resampling is originally a classification problem strategy for balancing rare classes with common classes, which has been adapted for use in regression problems using relevance values. For example, if the relevance value of a sample y is above a relevance threshold t_R then it is considered a member of the *rare* class (the word case is henceforth used as a synonym for class). Resampling strategies often work by splitting samples belonging to rare and common cases into separate bins. These bins are referred to as oversample bins for rare cases, and

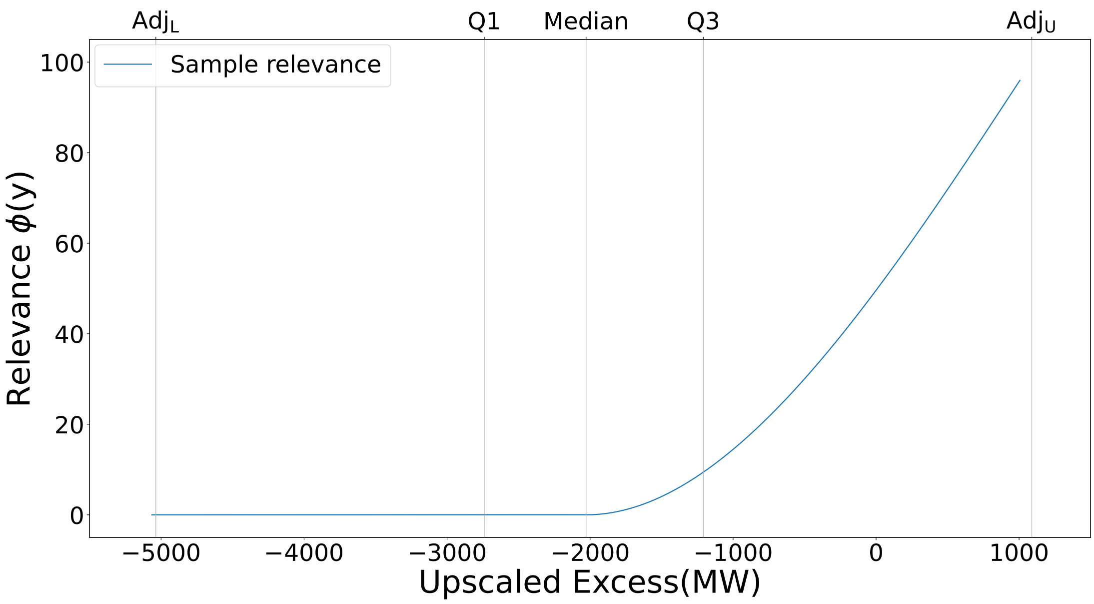


Figure 3.7: Relevance function for Upscaled Excess time series generated using PCHIP

undersample bins for common ones. As the dataset examined is time series in nature (and therefore, the ordering of samples is important), bins are also ordered temporally.

Algorithm 1 CreateBins(D, y, t_R, $\phi()$)

1: D: Examined data set
2: y: Target sample
3: t_R: Relevance threshold such that $t_R \in \{0, 1, \ldots, 100\}$
4: $\phi()$: Relevance function such that $\phi(y) : y \to \{0, 1, \ldots, 100\}$
5: $Bins \leftarrow \emptyset$
6: $b \leftarrow 0$
7: **for** $i \in D$ **do**
8:     $Bins_b \leftarrow Bins_b \cup (x_i, y_i)$
9:     **if** $i < |D|$ **then**
10:         **if** $\phi(y_i) \leq t_R < \phi(y_{i+1}) \vee \phi(y_i) > t_R \geq \phi(y_{i+1})$ **then**
11:             $b \leftarrow b + 1$
12: **return** $Bins$

Algorithm 1 defines the creation of bins, where every bin in the set $Bins$ alternates between common and rare cases, based upon a similar algorithm from literature [MBT17]. Lines 1–4 define the input parameters to the algorithm, while lines 5 and 6 create an empty set for the bins and the bin counter. Lines 7–11 iterate over all samples in the dataset (D). Line 8 assigns samples to the current bin. Line 9–11 increases the number of bins if a change is detected in the relevance value of the next sample (i.e., the next sample's relevance is lower than the relevance threshold if the current sample is above the threshold, or the inverse). As a result, every alternating bin holds either rare or normal cases. This bin-creation algorithm is used for all resampling strategies explored.

In this chapter, two regression resampling strategies are explored that have seen success in literature: Gaussian Noise (GN) and Temporal Relevance Undersampling (TRU).

3.3.1 Gaussian Noise

GN resampling stratagem is comprised of two parts, the first part undersamples the normal cases. The second uses the existing rare cases to create additional samples with the introduction of GN. To facilitate the introduction of GN to the dataset, a perturbation value is used (δ), such that $\delta \in [0, 1]$.

Algorithm 2 shows the *AddGaussianNoise* function, which has been modified from literature [ACC24] to account for the temporal nature of the datasets. Lines 1–5 describe the input parameters of the algorithm. Line 6 denotes the creation of Bins based on Algorithm 1, while line 7 identifies bins which contain normal cases (i.e., holds samples with relevance values below the threshold t_R) as the bins to be undersampled. Line 8 defines the number of samples each bin should contain. In undersample bins, the total number of samples is reduced to this value, while in oversample bins, they are increased. Lines 9 and 10 instantiate the new dataset and samples as empty sets. Lines 11–13 reduce the total number of normal samples using random undersampling. Line 11 iterates over all bins containing normal cases. Line 12 performs random undersampling, where the indices of $bintarget$ samples are selected to be kept from the current undersample bin. Here $sample()$ is a function which randomly selects indices from a list, and $sorted()$ sorts the indices from lowest to highest. The sorting of indices is important to ensure the samples kept are in the same time-based ordering as the original dataset. The randomly sampled indices are used to define the new contents of the current undersampled bin on line 13.

Lines 14–25 define the loop responsible for the creation of additional rare samples. Line 14 denotes the loop for all rare case bins while line 15 defines the number of new samples per existing sample in the current bin κ. The use of κ ensures new samples are based on an even distribution of existing rare samples. Lines 16 and 17 iterate through the contents of the oversample bin and create κ new samples respectively. Line 18 declares the new sample μ as an

Algorithm 2 AddGaussianNoise(D, y, t_R, $\phi()$, δ)

1: D: Examined data set
2: y: Target sample
3: t_R: Relevance threshold such that $t_R \in \{0, 1, \ldots, 100\}$
4: $\phi()$: Relevance function such that $\phi(y) : y \to \{0, 1, \ldots, 100\}$
5: δ: Perturbation value such that $\delta \in [0, 1]$
6: $Bins \leftarrow CreateBins(D, y, t_R, \phi())$
7: $Bins^{-} \leftarrow \{Bins_b \in Bins : \forall(x, y) \in Bins_b | \phi(y) < t_R\}$ // Undersample bins
8: $bintarget \leftarrow \frac{|D|}{|Bins|}$
9: $D_{new} \leftarrow \emptyset$
10: $\mu Bins \leftarrow \emptyset$
11: **for** $b \in Bins^{-}$ **do** // Random undersample common cases
12: $\quad v \leftarrow sorted(sample(|Bins_b|, bintarget))$
13: $\quad Bins_b \leftarrow Bins_{bv}$
14: **for** $b \in Bins - Bins^{-}$ **do** // Oversample rare cases
15: $\quad \kappa \leftarrow \frac{bintarget - |Bin_b|}{|Bin_b|}$
16: $\quad$ **for** $c \in Bin_b$ **do**
17: $\quad\quad$ **for** κ **do**
18: $\quad\quad\quad \mu \leftarrow \emptyset$
19: $\quad\quad\quad$ **for** $x \in c$ **do**
20: $\quad\quad\quad\quad \Delta x \leftarrow norm(0, \delta \cdot STD(x))$
21: $\quad\quad\quad\quad \mu_x \leftarrow c_x + \Delta x$
22: $\quad\quad\quad \Delta c_y \leftarrow norm(0, \delta \cdot STD(c_y))$
23: $\quad\quad\quad \mu_y \leftarrow c_y + \Delta c_y$
24: $\quad\quad\quad \mu Bins_b \leftarrow \mu Bins_b \cup \mu$
25: $\quad Bins_b \leftarrow Bins_b \cup \mu Bins_b$
26: $D_{new} \leftarrow Bins$
27: **return** D_{new}

empty set, while line 19 loops through all features of the current sample. Line 20 calculates the change that is applied to attribute x of sample c to create the new sample. Where $norm()$ is a function that selects a value from a normal distribution, where the centre of the distribution is 0, and the standard deviation is the STD of attribute x multiplied by the perturbation parameter δ. This value is added to the original value to create the new value for the attribute for the new sample in line 21. Lines 22 and 23 repeat this process for the target value of sample c. Line 24 then adds the new sample to a new bin, while line 25 adds the newly created samples to the current oversample bin b. Line 26 defines the new dataset based on the modified bins in lines 11–25 and returns it.

It is important to note that when the new sample featured in μBin_b is combined with the existing ones in Bin_b, new samples are inserted immediately after the sample used in its creation. This is done to ensure the temporal consistency.

3.3.2 Temporal Relevance Undersampling

Random undersampling is a traditional resampling method which randomly reduces the total number of normal samples to be balanced with rare ones, an implementation of this is seen in Algorithm 2. Extensions of random undersampling have been proposed to select samples in an informed way for time series datasets. One of these is resampling with a temporal bias, which gives a preference for samples which occur later in the time series (as these are often more relevant to estimations). This is further built upon by including relevance bias for imbalanced regression problems, which prioritise samples with higher relevance, even if that relevance value is below t_R. These extensions have been combined into temporal relevance bias undersampling, which has been shown to improve accuracy compared to resampling techniques which do not take into account these aspects [MBT17].

Algorithm 3 TemporalRelevanceUndersampling(D, y, t_R, $\phi()$)

1: D: Examined data set
2: y: Target sample
3: t_R: Relevance threshold such that $t_R \in \{0, 1, \dots, 100\}$
4: $\phi()$: Relevance function such that $\phi(y) : y \to \{0, 1, \dots, 100\}$
5: $D_{new} \leftarrow \emptyset$
6: $Bins \leftarrow CreateBins(D, y, t_R, \phi())$
7: $Bins^{-} \leftarrow \{Bins_b \in Bins : \forall(x, y) \in Bins_b | \phi(y) < t_R\}$ // Undersample bins
8: $bintarget \leftarrow \frac{|D| - \sum_{i=0}^{|Bins^{-}|} |Bins_i^{-}|}{|Bins^{-}|}$
9: **for** $b \in Bins^{-}$ **do**
10: $\quad p \leftarrow [\frac{1}{|Bins_b^{-}|} \cdot \phi(y_1), \dots, \phi(y_{|Bins_b^{-}|})]$
11: $\quad v \leftarrow sorted(sample(|Bins^{-}|, bintarget, p))$
12: $\quad Bins_b \leftarrow Bins_{bv}^{-}$
13: $D_{new} \leftarrow Bins$
14: **return** D_{new}

Algorithm 3 describes the implementation of the *TemporalRelevanceUndersampling* resampling method adapted from literature [MBT17]. Lines 1–4 define the input parameters for the algorithm, while lines 5–7 describe the creation of the empty dataset, bins, and identification of undersample bins in a similar way described for Algorithm 2. However, the definition of bin targets on line 8 differs, as the difference between the total number of rare and normal samples is used to determine how many samples each undersample bin should have (i.e., the total number of rare and normal samples should be balanced). Lines 9–12 define the undersample loop, where every undersample bin is iterated over. Line 10 calculates the selection preference for all samples in the current undersample bin, where the time at which the sample occurs (i.e., its index in the bin, as all bins are ordered in a temporal way as seen in Algorithm 1) multiplied by the relevance value for that sample is its preference. Line 11 selects the indices of samples in order of preference, or randomly if ties occur. These indices are then ordered from lowest to highest to maintain the temporal ordering. In line 12, the selected indices are used to define the new contents of the undersample bin b. Lines 13 and 14 show the assignment of the new dataset to the modified contents of all bins and returned.

3.4 Evaluation metrics

As discussed in Chapter 2 Section 2.3.6 (Page 17), using traditional regression metrics such as Mean Squared Error (MSE) have shown worse performance on imbalanced regression datasets compared to more specialised metrics. The performance of MSE as an evaluation metric and loss function is explored in this chapter, along with two specialised metrics from contemporary literature, Squared Error Relevance Area (SERA), and Inverse Weighted Mean Squared Error (IW-MSE). Both place more importance on samples with higher relevance or larger values when calculating metrics.

3.4.1 Squared Error-Relevance Area

SERA is a regression metric which uses a relevance function $\phi(y)$ to calculate the Sum of Squared Error (SSE) for different subsets of the dataset D.

$$SER_t = \sum_{i \in D^t} (\hat{y}_i - y_i)^2 \tag{3.1}$$

Equation (3.1) shows the Squared Error Relevance (SER) of each group, where D^t is a subset of the dataset D where $D^t = \{(x_i, y_i) \in D | \phi(y_i) \geq t\}$. In effect, t acts as a relevance threshold value similar to t_R, which divides the dataset into k groups with increasing relevance values. Each value for SER is plotted as a point on a curve, where the area under the curve is taken as the metric to be reduced.

$$SERA = \int_0^{100} \sum_{i \in D^t} (\hat{y}_i - y_i)^2 \, dt \tag{3.2}$$

Equation (3.2) shows the calculation of the SERA metric, which is the area under the curve denoted by SER_t, where t is the variable of the integration which takes whole numbers between 0 and 100. As t begins with a value of 0, all samples in dataset D are considered in the error (as all samples have a relevance value of at least 0). Therefore, the value of SER where $t = 0$ is the same as the SSE for the entire dataset. As the value for t increases, the number of samples in the calculation decreases, resulting in a decreasing SSE. Because of this, the squared error of samples with the highest relevance (i.e., $\phi(y) = 100$) is counted k times.

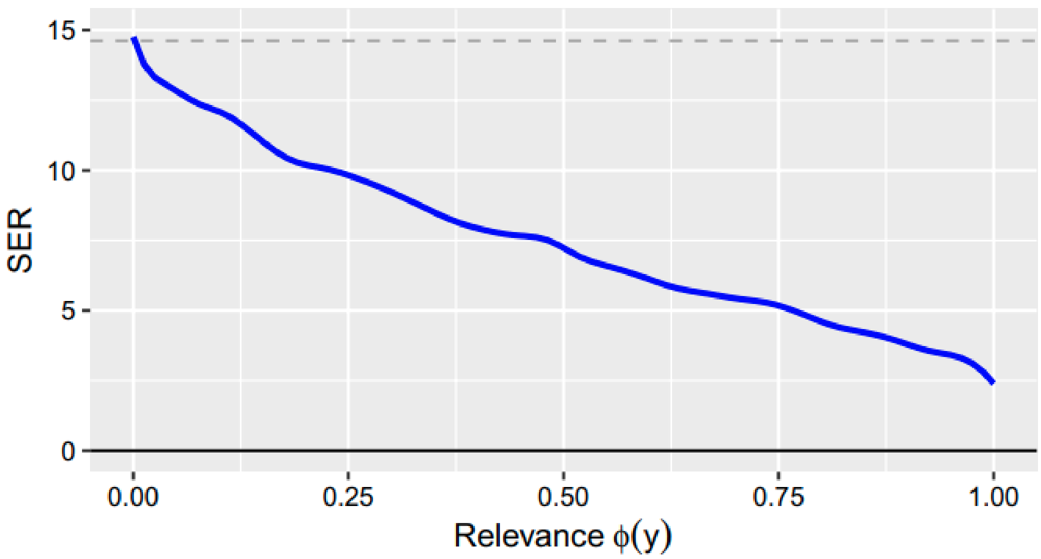


Figure 3.8: Example of SER curve from literature (Figure 8 from [RM20])

Figure 3.8 shows an example of a SER curve from the work which proposed the metric (Figure 8 from [RM20]). The y-axis represents the SER score, while the x-axis describes the value for t (i.e., the subset of the dataset D), and the grey dashed line represents the SSE of all predictions. As the area under the curve is minimised, reducing the SER of a subset with larger relevance values has the effect of reducing the SER for all preceding subsets.

3.4.2 Inverse Weighted Mean Squred Error

IW-MSE is a weighted MSE metric which assigns weights based on the percentile of the samples examined.

Algorithm 4 CreateWeights(D)

```
1: D: Examined data set
2: W ← {1 : ∀(x, y) ∈ D}
3: for i ← 50 to 100 do
4:     v ← P(D, i, i + 1)
5:     W_v ← 50 · 1/(100−i)
6: v ← P(D, 100)
7: W_v ← 51
8: return W
```

Algorithm 4 describes the assignment of weights to samples, which is adapted from literature [SSHSP23]. Line 1 denotes the dataset input for the algorithm, while line 2 sets all samples to a weight of 1 to initialise the set. Lines 3–5 denote a loop which assigns weights to samples between the 50$^{\text{th}}$ and 99$^{\text{th}}$ percentile inclusive. Line 4 obtains the index of all samples between the i^{th} and $i + 1^{\text{th}}$ percentile. Where P is a function which returns the index of all samples in dataset D at the i^{th} percentile, or between the i^{th} and $i+1^{\text{th}}$ percentile. Line 5 assigns all selected samples a weight that ranges between 1 and 50, depending on the percentile. Line 6 obtains the index of the sample at the 100$^{\text{th}}$ percentile, and line 7 assigns it a weight of 51.

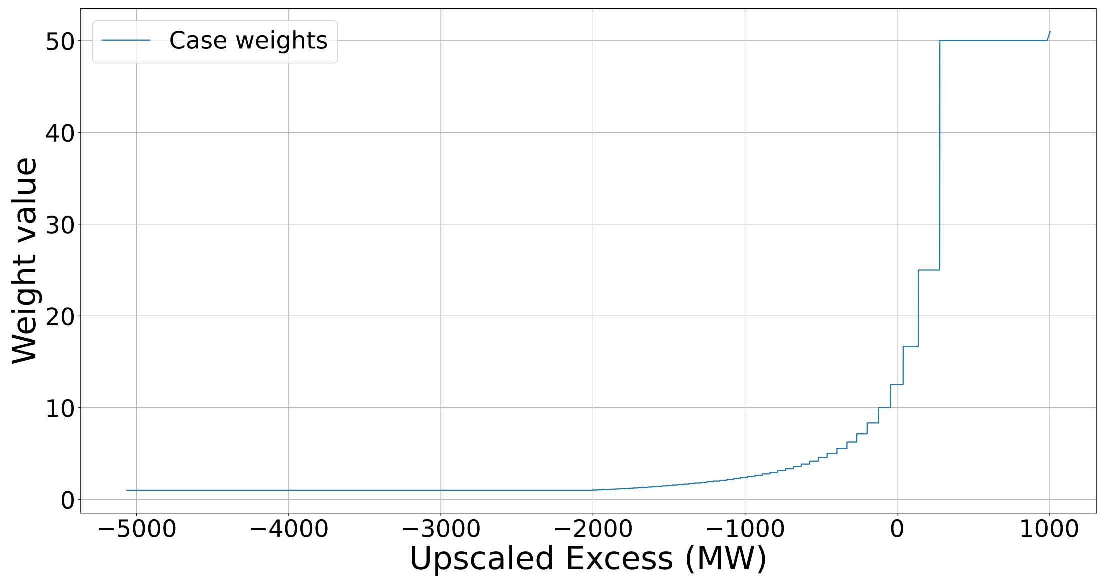


Figure 3.9: Weights assigned to cases

Figure 3.9 shows that the increase of sample weight is exponential with the percentile assigned to it. Each percentile between the 50$^{\text{th}}$ and 100$^{\text{th}}$ contains between 2770 and 2792 samples. From the figure, it can be seen that the range of values between percentiles increases as the 100$^{\text{th}}$ percentile is approached.

For example, the range of values between the 50th and 51st percentile is 28.4 (-2028.0 to -1999.6), while the range of values between the 99th and 100th percentile is 725.41 (280.39 to 1005.8). This explains the large weight *plateaus* which occur at higher percentiles.

It should be noted that both SERA and IW-MSE have been empirically evaluated with each other, the results of which show that both metrics perform well in different scenarios [SSHSP23]. They have, however, only been examined with convolutional LSTMs, and no other DL method has been explored.

3.5 Model composition and evaluation

Before beginning experimentation and the generation of weights and relevance values, the dataset is prepared. The data is first split into a training set and a test set. Where the test set is comprised of one year's worth of data, from November 2021 to October 2022 inclusive and contains 26,044 samples, 3120 of which feature an energy excess value above 0, representing 24.52% of all rare cases in the entire dataset. If no resampling takes place and the loss function is MSE, then the data is made stationary by applying a differencing operation to each subsequent value (i.e., the difference between two adjacent values is used rather than the actual value itself). After this, all experiments scale the data between a value of -1 and 1 using a MinMaxScaler[4], where a separate MinMaxScaler is used for the input data and prediction data.

Data is then organised into input sequences and prediction sequences. As discussed in Chapter 2 Section 2.3 (Page 10), a time series of length x is used to estimate the next n values. A value of 96 is used for x, as it represents 24 hours of input information prior to estimates. Meanwhile, a n value of 24 is used for initial experiments, representing predictions of 6 hours in the future. Additional values for n are explored later in this chapter. This results in an input sequence shape of (277881, 96, 6) and a prediction sequence length of (277881, 24, 1). A batch size of 96 is used (to represent 24 hours of information).

Implementation was done using Python 3.9, with Keras and TensorFlow version 2.10, and TensorFlow-probability library version 0.18. Furthermore, Compute Unified Device Architecture (CUDA) toolkit 10.1 and CUDA Deep Neural Network (cuDNN) Software Development Kit (SDK) 8.0.5 are used. Experiments

[4]scikit-learn's MinMaxScaler was employed for this operation `https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.MinMaxScaler.html`

Table 3.3: Parameters for each model to be optimised along with their explored value range

Model	Parameter	Value range
LSTM	# Layers	{1, 2}
	# Neurons	{50, 75, . . . , 200}
	Dropout rate	{0.0, 0.1, 0.2}
	Learning rate	[0.0001, 0.001]
CNN	# Layers	{1, 2}
	# Filters	{32, 42, . . . , 62}
	Kernel size	{2, 3, 4}
	Pool size	{2, 3, 4}
	Dropout rate	{0.0, 0.1, 0.2}
	Learning rate	[0.0001, 0.001]
GRU	# Layers	{1, 2}
	# Neurons	{50, 75, . . . , 200}
	Dropout rate	{0.0, 0.1, 0.2}
	Learning rate	[0.0001, 0.001]

are conducted on a computer with a 3.6 Gigahertz (GHz) AMD Ryzen 7 3700X Central Processing Unit (CPU), 16GB of DDR4 Random Access Memory (RAM), and an RTX 2070 Super Graphics Processing Unit (GPU).

To examine the efficacy of the described imbalanced regression methods, 6 different DL models are examined. These include LSTMs, GRUs, CNNs, and hybrid methods like LSTM-CNN, GRU-CNN, and LSTM-GRU. These models are selected as they have proven to be effective at estimating wind power energy as detained in Chapter 2 Section 2.4 (Page 19). Furthermore, details about how these models work can be found in Chapter 2 Section 2.3.1 (Page 12). These hybrid models are make by placing layers of different models types in sequential order. For example, the architecture for CNN-LSTM consists of CNN layers (along with a maxpooling layer per CNN layer) followed by a dropout layer. The output of the dropout layer is then passed to the LSTM layers. This process is followed for the other hybrid models, where the first model denoted in the name denotes the type of model placed at the start of the sequential ordering. To ensure the best possible results, Hyperparameter Optimisation (HPO) took place using Keras Tuner for each configuration [OBL+19].

Table 3.3 lists all parameters to be optimised for each DL configuration, along with the range of values explored. The dropout rate is optimised per model, irrespective of the total number of layers, with a step value of 0.1. The number of neurons for LSTMs and GRUs is explored with a step value of 25, starting from a base value of 50 neurons, which has been used in literature [LLF21], with a maximum of 200 neurons per layer explored to determine the impact of models with higher learning capacities. Meanwhile, the number of filters for CNN models is explored with a step size of 10, kernel sizes of 3x3 have previously been used in literature [SSHSP23]; therefore, a kernel size of 3 and adjacent values are explored. Additional works feature the use of 32 filters for CNN models, with a pooling size of 2 [SZS+24]. Larger values for these parameters are tested to determine the impact of more complex CNN models. The learning rate is also optimised, with log-based sampling, while the number of neurons, filters, kernel size, and pool size is optimised per model layer. The range of values for learning rate is between 0.0001 and 0.001, both of which have been used in literature [YCW+22, ZZZZ24, FDAH+22]. The Keras Tuner *BayesianOptimisation* class is used, with 10 trials per configuration and an objective of minimising validation loss. Each trial is executed for 15 epochs with 33% of the training dataset reserved for validation. The best-performing configuration is selected and trained for 200 epochs on the entire training dataset. A callback stops the training process if the loss has not improved within 25 epochs.

It should also be noted that simple linear models have been shown to outperform complex DL methods such as transformers for time series analysis [ZCZX23]. Specifically, the argument is made that attention mechanisms may be harmful to time series prediction due to their unordered nature. To this end, the application of attention mechanisms for the aforementioned DL models is not explored. However, more recent literature has shown DL methods do outperform linear models such as DLinear for complex problems that feature non-linear data, such as wind speed prediction [CZZ24]. As previously shown, the subset of the dataset which is the target of the imbalanced regression problem has non-linear relationships with other features of the dataset. Because of this, and the consensus in literature that non-linear models perform better with non-linear data compared to linear models [Ros24, HZ24, TCW25], a baseline of a linear model is not considered and is out of scope for the contribution of this chapter and this thesis. However, the exploration of linear models such as DLinear and transformers for imbalanced regression on a non-linear dataset is an interesting topic, which can be explored in the future.

The loss functions for SERA and IW-MSE are created using the Keras frontend library, along with the TensorFlow probability library. Algorithm 5 shows the implementation for the SERA loss metric. Lines 1–3 describe the input parameters for the loss function, while line 4 creates a set to contain all SSE values.

Algorithm 5 SERALoss(Y, $\hat{Y}$, ϕ)

1: Y: set of all actual values y to determine the loss of
2: $\hat{Y}$: set of all prediction $\hat{y}$ to determine the loss of
3: ϕ: set of relevance values for Y
4: $SSE \leftarrow \emptyset$
5: **for** $t \in 100 : -1 : 1$ **do**
6: $\quad v \leftarrow equal(\phi, t)$
7: $\quad SSE_t \leftarrow \sum_{i \in Y_v} (\hat{y}_i - y_i)^2$
8: $\quad$ **if** $t < 100$ **then**
9: $\quad\quad SSE_t \leftarrow SSE_t + SSE_{t+1}$
10: $SERA \leftarrow trapz(SSE)$
11: **return** $SERA$

Lines 5–9 denote a loop to calculate the SSE for the predictions $\hat{y}$, where t is assigned a value of 100, and decremented each loop to a value of 1. The function $equal()$ returns the indices of all Y values equal to the relevance value denoted by t on line 6. Line 7 calculates the SSE for all samples with relevance t. Lines 8 and 9 add the previous SSE value to the current one, this is done to reduce the size of summation operations needed to calculate SERA and is the reason why a relevance function $\phi(y) \in \{0, 1, \ldots, 100\}$ is used instead of $\phi(y) \in \{0, 0.01, \ldots, 1\}$. Finally, line 10 uses the TensorFlow probability $trapz()$ function to perform a trapezoidal integration with the SSE values and step values for t to determine the area under the SER curve.

$$\frac{1}{|Y|} \sum_{i \in Y} (\hat{y}_i - y_i)^2 \cdot W_i \tag{3.3}$$

Equation (3.3) shows the implementation of the IW-MSE metric, which is the same as the calculation for MSE, however, the multiplication of weights by the square error is done before the mean operation. Furthermore, for experiments that use resampling methods, t_R is assigned a value of 49. This is the relevance value for $\approx$-14 in the Upscaled Excess feature and is the closest whole relevance value to an Upscaled Excess value of 0. This value is selected so that all samples of value 0 or above are a part of the rare case, while samples below -14 are considered a part of the normal case to be undersampled. In addition to TRU and GN, the absence of any resampling method is also examined, denoted by *None*.

Table 3.4: Minimum, maximum, and mean trainable parameters for HPO chosen model architectures, along with mean dropout value for all configurations per model

Model	Min. params	Max. params	Mean params	Mean dropout
CNN	16204	71138	43598.8888	0.0777
LSTM	12624	173824	84909.7142	0.0571
GRU	56424	149424	91157.3333	0.0888
CNN-GRU	29397	140997	76903.4444	0.1388
CNN-LSTM	38022	307072	116581.5555	0.1166
GRU-LSTM	65799	448299	311977.1250	0.1250

3.5.1 Hyperparameter optimisation results

Before discussing the results of the configuration experiments, the results of the hyperparameter optimisation are explored. Table 3.4 shows the minimum, maximum, and mean number of trainable parameters, along with the average dropout value across all configurations for the best performing architecture for each type of model. Here, it can be seen that the hybrid models, such as GRU-LSTM, feature a very high number of trainable parameters. Both the GRU-LSTM models trained with IW-MSE and SERA using the resampling strategy GN have 448,299 trainable parameters, yielding a 1.61:1 ratio for trainable parameters to training dataset data points. Furthermore, except for CNN-GRU, the hybrid models feature a large number of trainable parameters on average across the nine configurations examined per model. This would suggest that the hybrid models are susceptible to overfitting. This is further reinforced by the mean dropout values, with all explored hybrid models featuring a notably higher average dropout compared to the more simplistic models. This suggests more aggressive dropout values could be explored, along with more simplistic architectures for these models, if their performance is poor.

The more non-hybrid DL models feature a more balanced ratio for training parameter to training data point, with LSTM SERA None featuring the highest number of trainable parameters of all of the single type models at 173,824. Despite this, the best performing LSTM architectures featured the lowest average dropout across all models, with LSTM SERA None having a dropout rate of 0. This would imply that, despite a trainable parameter-to-datapoint ratio of 0.62, it is unlikely that overfitting is occurring. Also of interest are the results for CNN models, which featured the lowest average number of trainable parameters, with around 0.15 trainable parameters per datapoint. As previously outlined, an early stopping mechanism was used during training. However,

Table 3.5: Number of times each resampling method provides the best results for each model across all loss functions examined.

Model	None	TRU	GN
CNN	3	0	0
LSTM	2	0	1
GRU	2	1	0
CNN-GRU	2	0	1
CNN-LSTM	2	0	1
GRU-LSTM	1	2	0
Total	**12**	3	3

Table 3.6: Number of times each resampling method provides the best results for each loss function across all models examined.

Loss	None	TRU	GN
MSE	6	0	0
SERA	1	2	3
IW-MSE	5	1	0
Total	**12**	3	3

only two configurations stopped before reaching the maximum training epoch of 200, those being CNN SERA None and GRU SERA GN, which stopped at epochs 137 and 150, respectively.

3.5.2 Results

Table 3.5 shows the total number of times each resampling method generated a model with the best loss score across all three loss metrics for each model type examined, where all rows must add up to three. From this table, it is clear that both imbalanced regression resampling strategies do not perform well compared to the absence of resampling methods. One exception to this is GRU-LSTM models, which see better performance with TRU resampling overall. This may imply that more complex, RNN models see better performance with fewer samples, implying overfitting may be occurring as previously speculated.

Table 3.6 shows the number of times each resampling method generated the best results for each loss function across all six models, where each row must add up to six. Here, it can be seen that configurations using MSE or IW-MSE favour the use of no resampling strategy. Interestingly, when using SERA loss function, both TRU and GN perform notably better. This is discussed in more detail later. More importantly, the reason for TRU and GN's poor performance compared to using no resampling method in Table 3.5 becomes clearer. All

Table 3.7: Number of times each resampling method provides the best results for each model across all loss functions examined for all three test metrics.

Test metric	Model	None	TRU	GN
MSE $y \geq 0$	CNN	3	0	0
	LSTM	2	1	0
	GRU	2	1	0
	CNN-GRU	2	0	1
	CNN-LSTM	2	0	1
	GRU-LSTM	1	2	0
	Total	**12**	4	2
SERA $\phi(y) \geq 49$	CNN	3	0	0
	LSTM	2	0	1
	GRU	2	1	0
	CNN-GRU	2	0	1
	CNN-LSTM	2	0	1
	GRU-LSTM	1	2	0
	Total	**12**	3	3
IW-MSE P^{96}	CNN	3	0	0
	LSTM	2	1	0
	GRU	2	1	0
	CNN-GRU	2	0	1
	CNN-LSTM	2	0	1
	GRU-LSTM	1	2	0
	Total	**12**	4	2

examined models trained using MSE and IW-MSE loss functions favour the use of no resampling method over TRU and GN in terms of the test metric. This is reflective of the fact that the inclusion of resampling methods increases the representation of rarer cases. So while rarer cases may be more accurately predicted, the overall accuracy will be lower due to the existence of normal cases in the test set, which is counterproductive to the goal of imbalanced regression. Therefore, to properly examine the performance of the methods employed in this evaluation, the error of samples above 0 (i.e., the rare cases) should be evaluated instead of MSE, SERA, or IW-MSE.

Therefore, three test metrics are introduced to evaluate resampling methods, loss functions, and models. The first is MSE $y \geq 0$, where only the MSE of samples with a value above 0 are considered. SERA $\phi(y) \geq 49$ where the SERA of samples with a relevance above 49 is considered. Finally, IW-MSE P^{96} considers the IW-MSE of samples above the 96$^{\text{th}}$ percentile of the original dataset, which is the closest percentile to the Upscaled Excess value of 0.

Table 3.7 shows the number of times a resampling method generates the best

Table 3.8: Number of times each resampling method provides the best results for each loss function across all examined models for all three test metrics.

Test metric	Loss	None	TRU	GN
MSE $y \geq 0$	MSE	6	0	0
	SERA	1	3	2
	IW-MSE	5	1	0
	Total	**12**	4	2
SERA $\phi(y) \geq 49$	MSE	6	0	0
	SERA	1	2	3
	IW-MSE	5	1	0
	Total	**12**	3	3
IW-MSE P^{96}	MSE	6	0	0
	SERA	1	3	2
	IW-MSE	5	1	0
	Total	**12**	4	2

score for all three previously outlined test metrics for each model across all examined loss functions. The best-performing resampling method per test metric is highlighted in the **Total** row. As shown, the change in test metric does not notable impact the performance of the three resampling metrics. With the main difference being the slight improvement in the performance of TRU over GN. This could imply that imbalanced regression resampling strategies are not as effective at improving the performance of DL models' ability to estimate less represented values compared to the loss function used.

Table 3.8 showcases the number of times a resampling method outperforms all others for all three test metrics for each loss function across all six models. Again, the only notable difference from the consideration of the new test metric is the slight improvement of TRU in configurations that use SERA as a loss function. This is likely due to TRU's reduced number of samples with a relevance value below 49, this would allow for more of a focus on samples with larger relevance values, as they would proportionally affect the overall SSE more during the training process. Also worth noting is that training with the SERA loss function takes significantly longer compared to the other examined loss functions due to the calculations needed, as seen in Algorithm 5. This is contrasted by the reduced dataset size when using TRU resampling, which would reduce the total time per epoch to train the model. This makes the SERA loss functions' preference for TRU resampling interesting. Even though there is no timeout element of the training process examined for these models, the relationship between the two could be exploited for faster model training with a specialised imbalanced

Table 3.9: Top 10 configurations ranked by test metrics of the MSE of samples above 0, SERA of samples with a relevance value of at least 49, and IW-MSE of samples apart of the 96th percentile and above.

Rank	MSE $y \geq 0$	SERA $\phi(y) \geq 49$	IW-MSE P^{96}
1	CNN SERA None	CNN SERA None	CNN SERA None
2	CNN IW-MSE None	CNN IW-MSE None	CNN IW-MSE None
3	LSTM MSE None	LSTM MSE None	LSTM MSE None
4	CNN SERA GN	GRU SERA TRU	CNN SERA GN
5	GRU SERA TRU	CNN SERA GN	GRU SERA TRU
6	LSTM SERA TRU	CNN MSE None	CNN MSE None
7	CNN SERA TRU	CNN IW-MSE TRU	CNN SERA TRU
8	LSTM SERA GN	CNN SERA TRU	LSTM SERA TRU
9	CNN MSE None	GRU IW-MSE None	LSTM SERA GN
10	CNN IW-MSE TRU	LSTM SERA GN	CNN IW-MSE TRU

regression loss function.

Another finding is that when training with a MSE loss function, there is a clear bias against any resampling method, even though this results in a model with no specialisation for imbalanced regression problems. This is seen in Tables 3.6 and 3.8. Furthermore, this may not be indicative of poor performance of the specialised loss functions or resampling methods, as the times the imbalanced regression methods perform better may produce better results overall.

Table 3.9 shows the top 10 ranking for all 54 configurations based on the three test metrics examined. Where the configuration denoted by rank 1 obtains the best score for the associated test metric column. As can be seen, both CNN models with no resampling methods trained with SERA and IW-MSE loss functions perform very well, taking the 1st and 2nd places respectively for all three test metrics. Interestingly, configurations featuring IW-MSE as a loss function are rare in the top 10, with only three distinct configurations appearing. This may be due to the split of samples across the percentiles of the original training dataset. For example, of the 26,044 samples in the test dataset, 1803 fall between the 99th and 100th percentiles, while 12 exceed the 100th percentile. Meanwhile, only 1557 samples fall between the 96th and 99th percentiles. This leads to a heavy skew as the more numerous samples have a might higher weight (50 for samples between the 99th and 100th percentiles) compared to others. This could result in IW-MSE generalising better for samples with higher Upscaled Excess values, while performing worse when predicting lower, but still positive values found in the 96th to 98th percentiles.

Also of note is the presence of configurations with no resampling method trained with MSE loss function. Despite these configurations not having any mechanism to aid in the prediction of imbalanced regression datasets, two such configurations are placed in the top 10. With LSTM MSE None model appearing as the third best configuration in all three examined test metrics. This is attributed to the fact that models trained without any imbalanced regression mechanisms represent the dataset as stationary. Representation of the dataset as stationary may play a large role in the DL model's ability to generalise on the rare cases in the dataset. Theoretically, such configurations will also be able to generalise better in some scenarios. For example, both SERA and IW-MSE loss functions place more importance on accurately estimating positive values, and less on negative ones. This could lead to a scenario where models trained with MSE can accurately predict "*false positives*" better (i.e., a sample with a negative value being incorrectly predicted as having a positive value). Due to the scarcity of Upscaled Excess energy, it is considered that accurately estimating positive samples is more important than this generalisation in the subsequent COP.

Additionally, TRU and None resampling strategies appear almost evenly split in their representation in the top 10. With configurations featuring TRU appearing 11 times, while configurations that use no resampling methods appear 13 times. This is reflective of the results seen in Table 3.8 regarding the performance of SERA. SERA is shown to be best performing loss function out of those examined, appearing in configurations 5.667 times on average for each test metric. While configurations using GN loss function only appear 6 times, it consistently appears within the top 5 configurations for each test metric. This is counter to the findings of Tables 3.5–3.8, where the resampling method performed poorly. This can also be attributed to the performance of the SERA loss function.

Attention should also be drawn to the fact that no hybrid model (CNN-LSTM, CNN-GRU, and GRU-LSTM) appears in the top 10 for any of the test metrics. Of the models that appear in the top 10, CNN appears to dominate the rankings, appearing 18 times, and consistently appearing in first and second place for all three test metrics. A possible explanation for the hybrid model's poor performance is overfitting, as the increased complexity of such models may lead to poor generalisation for the unseen samples present in the test dataset. This enforces the previous observations regarding the ratio of trainable parameters to datapoints for hybrid models, along with their high dropout rates. While CNN's ability to capture patterns and low ratio of trainable parameters to datapoints

Table 3.10: The training time post HPO for the configurations featured in Table 3.9 along with number of trainable parameters.

Model	Training Time (seconds)	Number Parameters
CNN IW-MSE TRU	404.99	71138
CNN IW-MSE None	1730.0	40012
CNN MSE None	1853.0	36536
LSTM MSE None	3553.0	12624
GRU IW-MSE None	5014.76	129624
CNN SERA TRU	6356.4	16204
LSTM SERA TRU	7528.25	97824
GRU SERA TRU	7928.16	74724
CNN SERA None	15444.21	40012
CNN SERA GN	24370.02	20954
LSTM SERA GN	35148.91	74224

may explain its good performance compared to other models.

Table 3.10 shows the training time post HPO and the total number of trainable parameters for the models featured in Table 3.9. From this, it is immediately obvious that configurations which use SERA for a loss function take significantly longer to train while not necessarily featuring more trainable parameters. This is in contrast to the performance of the loss function previously discussed, and can motivate not to use SERA in scenarios which may require the retraining of prediction models, or where training costs are a concern. This is due to the intensive calculation process for the loss function outlined in Algorithm 5. However, the use of TRU as a resampling method has been shown to decrease training time, due to the reduced number of samples used to train the models.

While the best-performing configurations using SERA in Table 3.9 do not use TRU, the configuration GRU SERA TRU features in the top 5 configurations. This can lead to an interesting discussion around regarding the trade off between training times and model performance for scenarios where models need to be retrained. However, this is not the focus of this chapter. It is also clear that the selected loss function has more impact on training time compared to the number of trainable parameters. For example, CNN IW-MSE None (the second best performing configuration) features the second lowest training time at 1730 seconds while having the same number of trainable parameters as CNN SERA None (the best performing configuration), which takes 8.92 times longer to train. From this, it is clear that the DL model used and its architecture do not impact training costs as much as the loss function and resampling strategy

Table 3.11: Average difference between the ranking of models based on loss function and test metric across all resampling methods.

Test metric	Model	MSE	SERA	IW-MSE
MSE $y \geq 0$	CNN	2.33	2.33	5.0
	LSTM	0.67	2.67	2.67
	GRU	5.0	6.0	3.33
	CNN-GRU	9.0	7.33	4.0
	CNN-LSTM	4.67	2.33	4.0
	GRU-LSTM	6.0	1.33	2.0
	Average	4.61	3.67	**3.5**
SERA$\phi(y) \geq 49$	CNN	1.33	2.33	5.67
	LSTM	0.67	2.67	4.0
	GRU	3.33	7.0	2.67
	CNN-GRU	6.0	11.33	4.67
	CNN-LSTM	2.33	3.33	4.0
	GRU-LSTM	5.0	3.67	2.0
	Average	**3.11**	5.06	3.83
IW-MSE P^{96}	CNN	1.33	2.33	4.67
	LSTM	0.67	2.0	3.0
	GRU	4.67	6.67	3.67
	CNN-GRU	8.0	9.67	4.0
	CNN-LSTM	4.0	3.67	3.67
	GRU-LSTM	5.67	3.0	2.0
	Average	4.06	4.56	**3.5**

selected.

Table 3.11 shows the average difference between the ranking assigned to each model when using each loss function as an evaluation metric, against the ranking assigned for the three test metrics previously discussed, across all examined resampling methods. Here, larger values represent a large difference between the ranking of the model using the loss metric for the column and the ranking of the model using the test metric. To this end, the loss metrics with a large average value can be considered more inaccurate in determining the accuracy of models for the goal of imbalanced regression, which may have an impact on the training process. As can be seen from this table, SERA performs poorly, with configurations ranking 3.67–5.06 positions different on average depending on the test metric examined. This is surprising considering the good performance of the loss function as previously seen, but can be attributed to the way importance is placed on different samples at different relevance values. As SERA compounds the SSE of samples with higher relevance values onto the SSE of lower relevance samples, the resulting SER of samples that are $\phi(y) < 49$ can

Table 3.12: Difference between the ranking of configurations based on loss function and test metric ranking for the top ranked configurations listed in Table 3.9.

Model	MSE $y \geq 0$	SERA $\phi(y) \geq 49$	IW-MSE P^{96}	Average
CNN SERA None	0	0	0	**0.0**
CNN IW-MSE None	1	1	1	1.0
GRU SERA TRU	2	1	2	1.67
LSTM MSE None	2	2	2	2.0
CNN IW-MSE TRU	1	4	1	2.0
CNN SERA GN	2	3	2	2.33
GRU IW-MSE None	5	0	4	3.0
LSTM SERA GN	2	4	3	3.0
LSTM SERA TRU	4	4	2	3.33
CNN SERA TRU	5	4	5	4.67
CNN MSE None	7	4	4	5.0

be significant. This can lower the overall ranking of a configuration when considering all samples.

Meanwhile, IW-MSE displays the lowest average difference for two of the three test metrics, meaning it is the most accurate loss function for ordering configurations in regards to their performance of estimating under-represented values. This is due to the distribution of weights across samples, as the under-represented values present at the highest percentiles are assigned significantly higher weighting compared to lower values. Meaning much more importance is placed on these samples than those found at the percentiles under 96. This also potentially means IW-MSE may generalise better for datasets featuring large amounts of Upscaled Excess, which could occur during storms or certain periods during the year.

Table 3.12 shows the difference in ranking for configurations featured in Table 3.9 and the rankings based on the loss metric used to train that particular model. Here it can be seen that CNN SERA None achieves the lowest difference of 0, with SERA accurately ranking the best-performing model. CNN IW-MSE None is ranked as the second-best model when ordered using IW-MSE, thus having a distance of 1, as it is the best-performing configuration using IW-MSE. Interestingly, the best performing configuration when ordered by IW-MSE is LSTM MSE None. Furthermore, when ranking configurations by MSE LSTM MSE None is also considered the best, thus leading to an average distance of two between the loss functions ranking and the test metrics ranking. This is representative of the lack of an imbalanced regression mechanism from said configuration, which results in reduced performance when only considering samples with a value above 0. When configurations are ranked by SERA, LSTM MSE None is considered the fourth best, which then changes to the third best when ranking based on the test metrics. This may be representative of the configuration's ability to generalise better using a differenced dataset when not using any imbalanced regression methods.

In general, IW-MSE proves the best at ranking the best performing configurations, with an average difference of 2. This is consistent with the previous findings from Table 3.11. However, SERA is shown to outperform MSE when ranking the configurations featured in Table 3.9 with an average difference of 2.5 and 4.5, respectively. This is different from the findings of Table 3.11, and implies that SERA is better at ranking these best performing configurations, while not necessarily generalising well. To this end, a more detailed examination of the top-performing configurations' performance is needed to determine why these differences occur.

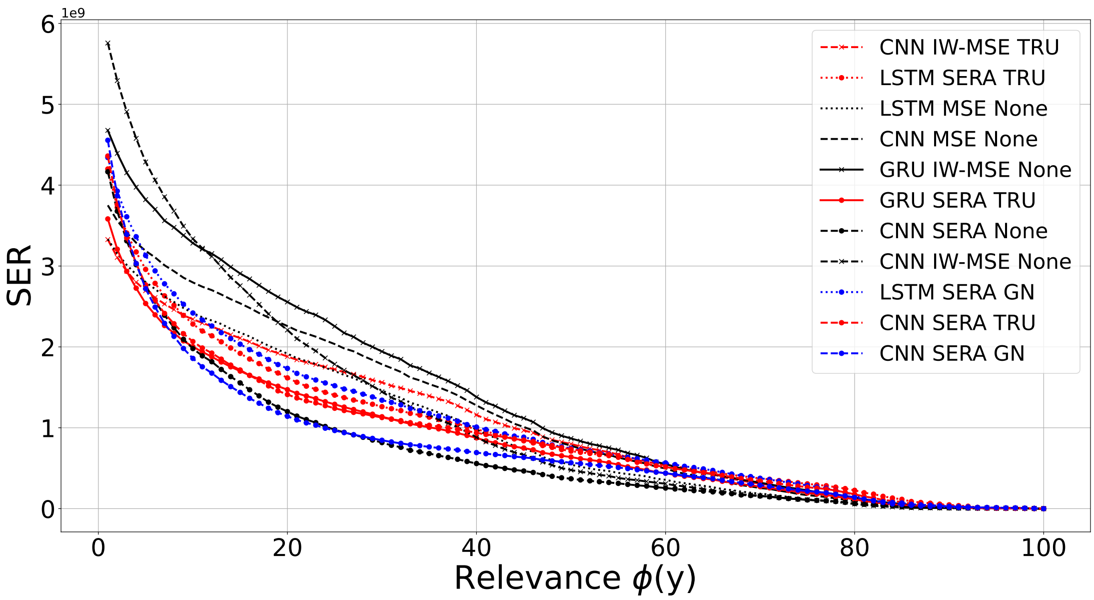


Figure 3.10: SER curve for the configurations listed in Table 3.9 where $\phi(y) \geq 1$

Figure 3.10 shows the SER curve of all configurations listed in Table 3.9 for all samples with a relevance of 1 or greater (i.e., $\phi(y) \geq 1$). From this figure, more insight can be gleaned as to the ranking of the configurations in Table 3.9.

For example, CNN IW-MSE None generalises poorly across the entire dataset when including values below 0. As previously outlined, samples below the median value of -2028 are assigned a relevance value of 0. As seen from this figure, CNN IW-MSE None has the highest SER score for all samples with a relevance value above 0. Similarly, CNN SERA None shows poor results, with LSTM MSE None appearing to perform the best when $\phi(y) \geq 1$. Additionally, CNN SERA TRU was shown to do quite well in Table 3.9, taking fourth place for the SERA based test metric. However, from this figure, it is clear that this configuration only performs well in generalising with samples that have a relevance value of less than $\phi(y) \leq 40$. Up until this point, this configuration outperforms both CNN IW-MSE None and LSTM MSE None. Furthermore, CNN SERA GN featured prominently in Table 3.9, taking fourth place for two of the three test metrics. Here it can be seen that this configuration outperforms the second and third place configurations until $\phi(y) = 45$, while it was also able to outperform CNN SERA None for a subset of relevance values.

As previously outlined, samples with a relevance below $\phi(y) < 49$ are considered normal cases. As a result, the previously mentioned configurations do not generalise well for the rare cases. It would seem from this that configurations utilising resampling strategies perform generalise well with the normal cases. However, this is not necessarily the case.

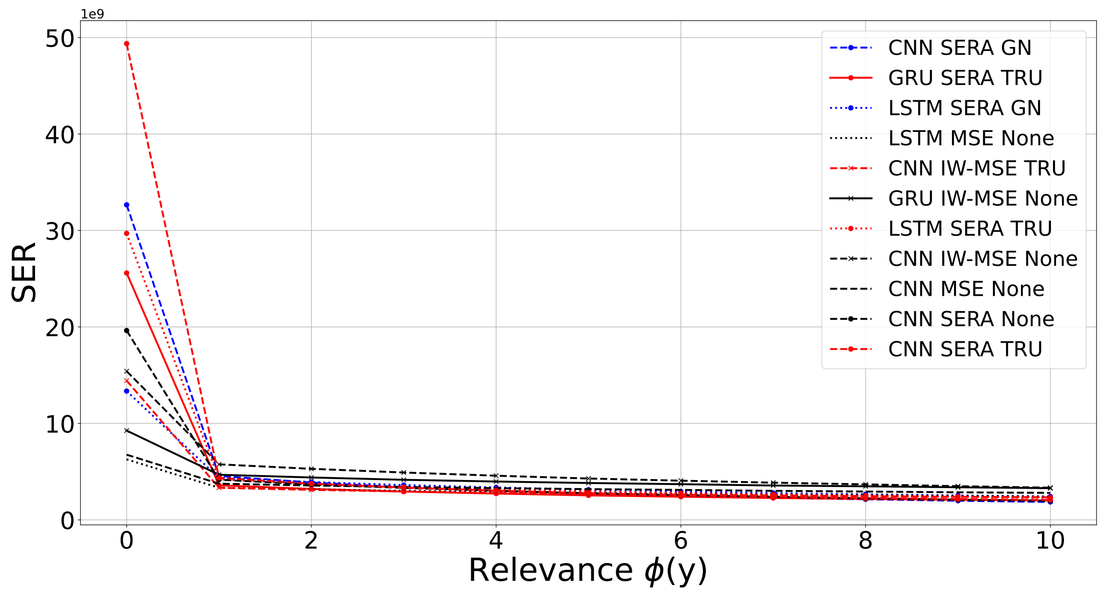


Figure 3.11: SER curve for the configurations listed in Table 3.9 where $\phi(y) \leq 10$

Figure 3.11 shows the SER curve of all configurations listed in Table 3.9 for all samples with a relevance of 10 or less (i.e., $\phi(y) \leq 10$). Of note, the inclusion of $\phi(y) = 0$ means all values in the dataset are evaluated, including those below the median. Such samples were excluded from the previous figure due to scaling reasons. However, from this figure, it is clear that the examined resampling strategies perform poorly when including all samples in the dataset. This is representative of the results seen in Tables 3.5–3.8, where both resampling strategies perform poorly. The addition of new test metrics in the later tables does not resolve this, as such strategies only seem to perform well when relevance is in a particular range (i.e., $1 \leq \phi(y) \leq 45$). However, the goal of this Chapter is to evaluate methods for the estimation of rare cases.

Figure 3.11 also shows that configuration LSTM MSE None achieves the lowest overall SSE score of all examined configurations. This is expected, as no resampling methods are employed. Therefore, the configuration treats the loss of the entire dataset equally. This is the reason why this configuration is placed first if MSE and IW-MSE are used as previously seen. As the loss of all the samples below the median seems to outweigh the loss of the more relevant samples.

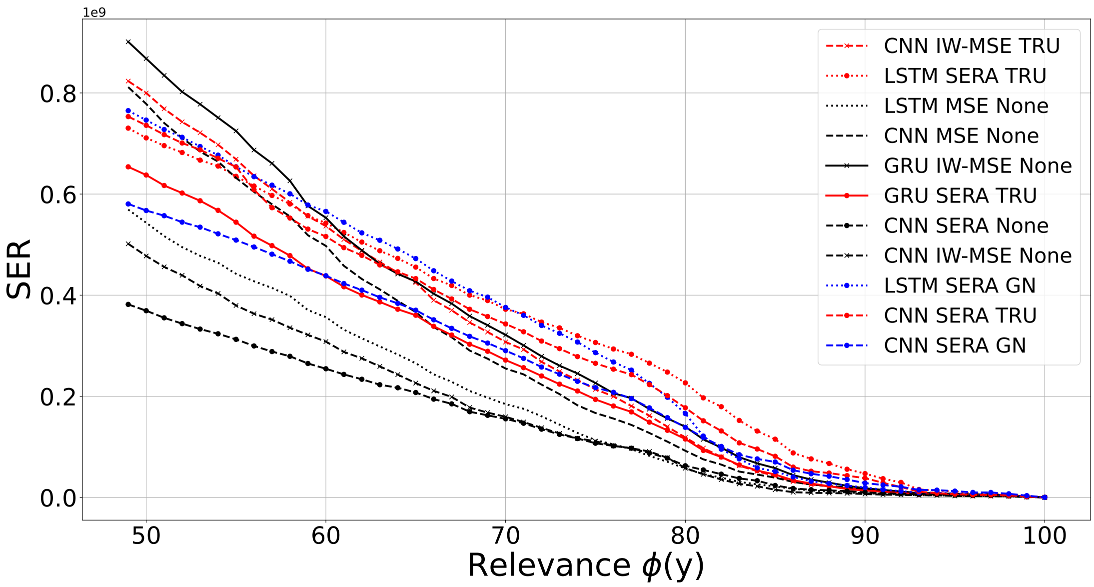


Figure 3.12: SER curve for the configurations listed in Table 3.9 where $\phi(y) \geq 49$

Figure 3.12 shows the SER curve for all configurations listed in Table 3.9 for all samples in the test dataset with a relevance of 49 or greater ($\phi(y) \geq 49$), representing the rare cases of the dataset. Here, a closer look can be seen at each configuration's performance when estimating samples with more importance. From this, it can be seen that the best performing configuration from Table 3.9, CNN SERA None, outperforms all others until around $\phi(y) = 77$. As previously outlined, 3120 samples above 0 exist in the test dataset, which increases to 3192 when considering samples with a relevance of $\phi(y) \geq 49$. Of these, 2562 samples (approximately 80%) lie between the range of relevance values $49 \leq \phi(y) < 77$. This shows that the best performing configuration, CNN

SERA None does a very good job at generalising for the majority of the rare cases examined in the test dataset.

As the relevance value of samples exceeds 76, both CNN IW-MSE None and LSTM MSE None begin to outperform CNN SERA None. This reinforces the argument that loss functions such as IW-MSE generalise better for the least represented portions of the dataset, as only 630 samples have a relevance value above 76. This represents an Upscale Excess value of approximately 609 and above, which is included inside the 99th percentile of the training dataset, which is heavily weighted. Therefore, IW-MSE may prove advantageous to train time series models if there is an expectation of an increasing trend in the data. Which for the examined dataset, there is a consideration of the Irish governments commitment to increasing the capacity of wind-generated energy. For similar reasons, LSTM MSE None also performs well, as the lack of any imbalanced regression mechanism means the dataset is different, resulting in the DL model learning the trend of the data, instead of the values themselves.

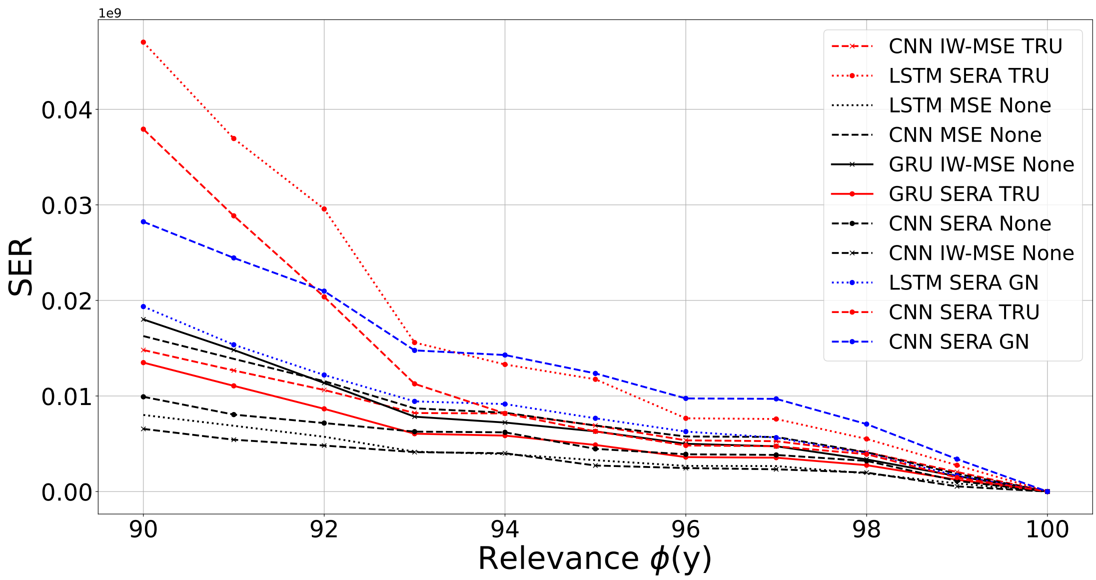


Figure 3.13: SER curve for the configurations listed in Table 3.9 where $\phi(y) \geq 90$

Figure 3.13 shows the SER curve for all configurations listed in Table 3.9 for all samples in the test dataset with a relevance of 90 or greater ($\phi(y) \geq 90$). In total, 48 samples of the test dataset have a relevance value equal to or above 90. While this is a very small portion of the test dataset, due to the way both SERA and IW-MSE work, a high level of importance is placed on these samples. Note, the SER for $\phi(y) = 100$ is zero as there is no sample in the test dataset with a value equivalent to $\phi(y) = 100$ in the training dataset.

As can be seen, CNN IW-MSE None outperforms all other configurations for this subset of samples, except for LSTM MSE None, which achieves similar SER scores at certain subsets (i.e. $\phi(y) = 94$). These results show that CNN IW-MSE None can more accurately predict the most under-represented values of an imbalanced regression dataset when estimating available Upscaled Excess. This can be useful in certain scenarios, such as estimating the additional energy generated during storms or other times when wind-generated power is high. However, CNN SERA None has shown that can generalise better for approximately 80% of the rare cases in the examined test dataset.

3.5.3 Further horizons

Previous discussion has shown that both CNN SERA None and CNN IW-MSE None perform well in the examined imbalanced regression dataset for predicting Upscaled Excess of the Irish national grid. The former has proven to generalise well with 80% of the rare cases in the examined test dataset, while the latter can more accurately predict the extreme cases which are present in the 99$^{\text{th}}$ percentile of the training dataset. Both of these configurations have merits, and to determine their impact, both models will be used in Chapter 4. To this end, they must be extended to predict further horizons. This is done by increasing n from 24 to 48, 72, and 96 so estimations 12, 18, and 24 hours into the future can be made. The resulting models will work with the others of the same configuration as will be detailed in Chapter 4. Therefore, these models specialise in estimating the final 24 time-steps of their respective n value. For example, the 12-hour model, which estimates 48 time-steps ahead, will only consider the loss of estimations 25–48 of each sample.

Table 3.13: The test metrics for all extended prediction horizon CNN models trained with SERA and IW-MSE along with the MAPE for all samples above 0

Loss	n	MSE $y \geq 0$	SERA $\phi(y) \geq 49$	IW-MSE P^{96}	MAPE $y \geq 0$
SERA	24	1.2x10^5	6.97x10^9	4.79x10^6	322.34
	48	7.34x10^5	48.99x10^9	31.09x10^6	496.06
	72	12.55x10^5	83.09x10^9	52.98x10^6	673.39
	96	19.47x10^5	137.74x10^9	83.47x10^6	838.33
IW-MSE	24	1.56x10^5	7.99x10^9	6.0x10^6	529.25
	48	5.18x10^5	29.22x10^9	20.27x10^6	603.52
	72	18.08x10^5	98.59x10^9	69.47x10^6	1166.97
	96	33.87x10^5	210.46x10^9	138.96x10^6	1605.97

Table 3.13 shows the test metrics MSE $y \geq 0$, SERA $\phi(y) \geq 49$, IW-MSE P^{96},

and Mean Absolute Percentage Error (MAPE) of samples above 0 (MAPE $y \geq 0$) for all CNN models trained with SERA and IW-MSE with no resampling method for four different prediction horizons. Interestingly, while the metrics for CNNs trained with n=24, 72, and 96 favour the SERA loss function, IW-MSE performs better when n=48.

This may be indicative of previously seen phenomena where these models perform well on a subset of relevance values, but poorly on the highest relevance values. This is potentially indicated because MAPE $y \geq 0$ still reports that the CNN model trained with SERA is better. However, it is clear that the model created to predict 12-hours ahead using IW-MSE seems to have some form of advantage over the one using SERA. The impact this has on the COP is to be seen.

As can be seen, the accuracy of models generally declines as the prediction horizon increases. This is a well-documented effect in literature [MCM21]; however, this may cause issues when solving COPs, as longer horizons are unreliable.

3.6 Conclusion

In this chapter, an empirical evaluation of 54 DL configurations was conducted to determine the impact that loss metrics and resampling strategies for imbalanced regression datasets would have on the performance of DL approaches. In total, three different loss metrics, two of which are specially designed for imbalanced regression datasets, were evaluated against MSE, and two specialised resampling methods were compared to the use of no resampling method across 6 different DL models.

Evaluation of resampling methods originally showed that not using any resampling method was preferred over the other examined methods. Additional evaluation with more specialised test metrics showed that configurations using resampling methods performed well for a subset of the test dataset. However, when evaluating the entire test dataset and a subset of rare cases, it was found that configurations using no resampling strategy outperform those which do. The evaluation also showed that a CNN model trained with a SERA loss function with no resampling strategy performs the best across all three examined test metrics. This was followed by a CNN model trained with IW-MSE loss function with no resampling strategy. Additional investigation yielded results showing that, while CNN SERA None generalised well across approximately 80% of rare case in the test dataset, CNN IW-MSE outperforms it for the most extreme rare cases examined.

It was also found that of the top 10 ranked configurations for each test metric, those trained with SERA appeared the most (5.667 times per test metric on average). However, it should be noted that such configurations have the largest training cost, with the calculations required for the SERA loss function extending the total training time significantly. While configurations which utilise SERA loss function and a TRU resampling strategy perform worse, they are less costly to train, which may prove advantageous in some scenarios.

One interesting thing to note is the presence of DL configurations trained with MSE and no resampling method in the top 10 configurations for each test metric. This is attributed to the use of a stationary dataset for these models being better able to capture changes in regression values, rather than the values themselves. Regardless, the use of imbalanced regression methods, such as specialised loss functions and resampling strategies, proves beneficial when using DL models for an imbalanced regression dataset. This evaluation has filled a literature gap discussed in Chapter 2 Section 2.11 (Page 45), as is considered Contribution 1 of this dissertation as outlined on Page 5.

The two best-performing configurations, CNN SERA None and CNN IW-MSE None, are also extended to estimate 12, 18, and 24 hours into the future. These are used in Chapter 4 to predict the available Upscaled Excess for use in a BEB-CSP. The results of which are compared to a scenario with full information about the available Upscaled Excess to determine the usefulness of using predictions. Performance of the BEB-CSP using both type of predictive configurations are also compared to determine if a generalised model outperforms one that predicts extremes better. Furthermore, due to the increasing error for predictions made for longer horizons, a method is developed in Chapter 4 to alleviate this.

Chapter 4

Sustainable Scheduling of Charging Events for Battery Electric Buses

4.1 Introduction

As discussed in Chapter 2 Section 2.2 (Page 10) Battery Electric Bus (BEB) are vehicles that produce lifecycle emissions based on the electricity used to charge their batteries. Furthermore, in Chapter 2 Section 2.7.6 (Page 36) it was found that BEBs charged using solely renewable energy sources produce the lowest amount of lifecycle emissions. As such, it is important to consider the source of electricity used to charge BEBs, in particular, Ireland has a high potential for wind-generated energy as outlined in Chapter 3.

As summarised in Chapter 2 Section 2.7.2 (Page 30), the Battery Electric Bus Charging Scheduling Problem (BEB-CSP) has several considerations that complicate the generation of charging schedules, such as limited charging infrastructure, and reduced operational range. In this chapter, a Constrained Optimisation Problem (COP) is defined to solve the BEB-CSP for a fleet of BEBs to minimise the use of non-clean energy (i.e., energy that is not a part of the Upscaled Excess defined in Chapter 3).

Additionally, as discussed in Chapter 2 Section 2.11 (Page 45) there is a gap in the literature related to using predictive models to estimate the value of an unknown variable (such as Upscaled Excess) and evaluating the performance of the BEB-CSP using the predictive model against other scenarios. To this end, four scenarios are explored regarding the knowledge of available clean energy. A naïve scenario which assumes no information is available. Two predicted scenarios utilising the two best-performing models from Chapter 3, and finally an Ideal scenario which uses actual historical information about how much clean energy is available. Furthermore, some methods are discussed to deal with potential mispredictions from prediction models with longer prediction horizons. The evaluation of these scenarios and methods for a BEB-CSP tested on three Irish cities fills the gap in the literature and is considered Contribution 2 of this dissertation.

The remainder of this chapter is divided into the following sections. Section 4.2 defines the BEB-CSP and the considerations made. In addition, mitigation strategies for dealing with inaccurate predictions are discussed, and the datasets used for evaluation are detailed. Then Section 4.3 presents the formal COP model of the problem definition, including the nomenclature of the constants and variables used. Next, Section 4.4 outlines the experiment configurations and performs an extensive evaluation of results. Finally, Section 4.5 summarises the chapter and ties it to the thesis statement and Contribution 2.

4.2 Problem definition

Given a set of bus routes ($b \in S$) which contain several stops ($i \in S_b$) in a repeating sequence. Each with an originally scheduled arrival time (τ_{bi}), assign a new arrival time (t_{bi}) and a charge time (ct_{bi}), such that a BEB-CSP can be solved with a BEB assigned to all bus routes. The resulting BEB-CSP is subject to a minimisation objective function to reduce the total amount of non-clean energy consumed by the BEB fleet (nc_{bi}). Where non-clean energy is considered energy that occurs outside a Time-of-Use (TOU) window and above the amount of clean energy available in each TOU window.

TOU windows ($k \in |\Gamma|$) are periods where wind-generated energy exceeds the national grid demand, as outlined in Chapter 3. Each TOU window has a start time (p_k), end time (q_k), and available energy (ce_k). Energy available during a TOU window can only be used if a BEB charges between the associated start and end times. Furthermore, the total available wind-generated excess energy is limited, meaning a maximum amount of clean energy can be consumed per window. The available wind-generated excess energy per window is portioned out based on the total energy required to operate fleets of BEBs for all examined datasets.

Each BEB has a battery capacity at each stop (c_{bi}) which falls between a maximum and minimum battery capacity ($C_{\max}$ and $C_{\min}$ respectively) represented in Kilowatt hour (kWh). The BEB's battery loses energy as it travels between two bus stops (D_{ij}), however, if a stop is considered a charging station (X_i) the BEB may recharge its battery (x_{bi}). If a BEB recharges its battery, the amount of energy gained (e_{bi}) is the linear product of the charge time (ct_{bi}) and the charge rate ($\mathcal{R}$). The amount of energy gained by a BEB at one stop is limited, as the maximum battery capacity cannot be exceeded. An upper and lower bound is also placed on the total amount of time a BEB can charge for (β and γ respectively). This is done to avoid battery overheating, micro-charging, and overcharging. Additionally, usage of an upper bound for charging times helps to dissuade solutions where battery capacity may exceed the cut-off point for the Constant Current (CC) phase of battery charging from a single charge event, as described in Chapter 2 Section 2.7.4 (Page 32).

The new arrival time of a BEB at a given stop depends on the arrival and charge time at the previous stop and the time taken to travel between the two stops (T_{ij}). The maximum difference between the original arrival time and the new arrival time at each stop (Δt_{bi}) is limited (Δt). This allows for control over the maximum difference in arrival time, which can be beneficial to guarantee service quality will not deteriorate past a certain point.

Fast charging stations are assumed to be placed at certain intervals throughout the bus route to provide opportunity charging as described in Chapter 2 Section 2.7.1 (Page 2.7.1) and are unary in nature. That is, each charging station may only have one BEB at a time. As such, if two BEB charge at a shared charging stop (Z_{bdij}), the charge time must not overlap.

While it is assumed that overnight charging takes place at the depot, it is also assumed that the slow charging rate associated with this type of charging means buses may not start with full battery capacity.

4.2.1 Dealing with mispredictions for scheduling problems

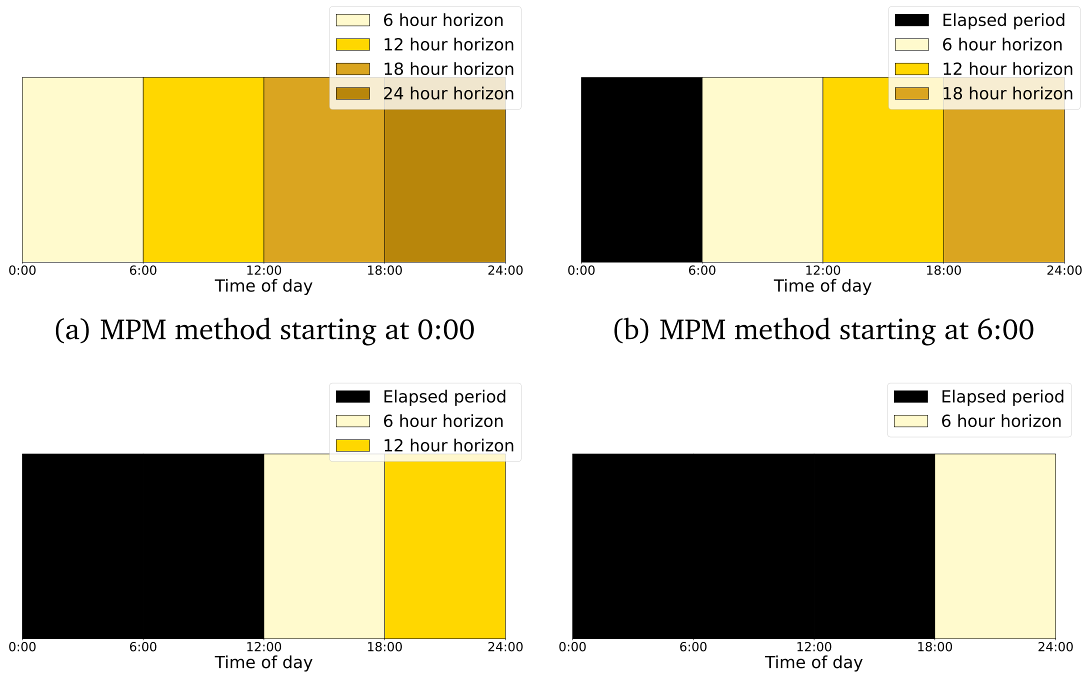
In Chapter 3 imbalanced regression methods were evaluated along with DL models to find the best configurations to estimate wind-generated energy excess. The two best-performing configurations were found to be Convolutional Neural Network (CNN) models trained with Squared Error Relevance Area (SERA) or Inverse Weighted Mean Squared Error (IW-MSE) loss functions using no resampling strategies. While the former loss function was shown to generalise well for approximately 80% of rare cases, the latter appeared to perform well at predicting extremes in Upscaled Excess. These configurations were extended to predict 12, 18, and 24-hours ahead in addition to the 6-hour prediction model. Despite the extensive evaluation to select a prediction model, mispredictions may still occur. Especially as the forecast horizon increases, which causes an increase in prediction model error as seen in Chapter 3 Section 3.5.3 (Page 81).

Even minor mispredictions may result in losses in optimality when considering the output of a prediction model as a deterministic input for a COP such as the BEB-CSP. As such, some consideration should be made to diminish the impact of such mispredictions. To this end, a reactive approach is taken, such that when a misprediction is identified, a new, more accurate set of predictions is provided to the BEB-CSP to generate a new schedule.

Two methods are examined in this chapter for the interaction between the prediction and optimisation model. The first is a Multi-Period reactive Model (MPM) method, which focuses on the aggregation of predictions from four different forecasting models to determine the availability of wind-generated energy excess throughout the day. Each forecasting model has a different prediction horizon (i.e., 6, 12, 18, and 24 hours) as denoted in Chapter 3 Section 3.5.3 (Page 81).

At the beginning of the day, the aggregated predictions are used to generate a charging schedule for the entire day with 100% degree of confidence. When each prediction horizon has passed (i.e., a "checkpoint" has been reached), the predictive models are updated with the most recent time series data. If a deviation from the previous estimations is detected, the schedule from that checkpoint onwards is recalculated. In particular, Figure 4.1a calculates a schedule for the entire day with checkpoints at 0:00, 6:00, 12:00, and 18:00. The fleet of BEBs begins operations as prescribed by the optimisation model at the first checkpoint. Figure 4.1b outlines the behaviour at the second checkpoint (6:00). At this point, current knowledge of wind-generated energy excess is incorporated and updates the predictions for subsequent prediction models. The method repeats this behaviour for all remaining checkpoints, as seen in Figures 4.1c and 4.1d.

As previously shown, the estimations of predictive models are more reliable and accurate for short-term time horizons. This is the basis of the rolling horizon approach [ZAdAR22], where a schedule is divided up into several periods,



(a) MPM method starting at 0:00

(b) MPM method starting at 6:00

(c) MPM method starting at 12:00

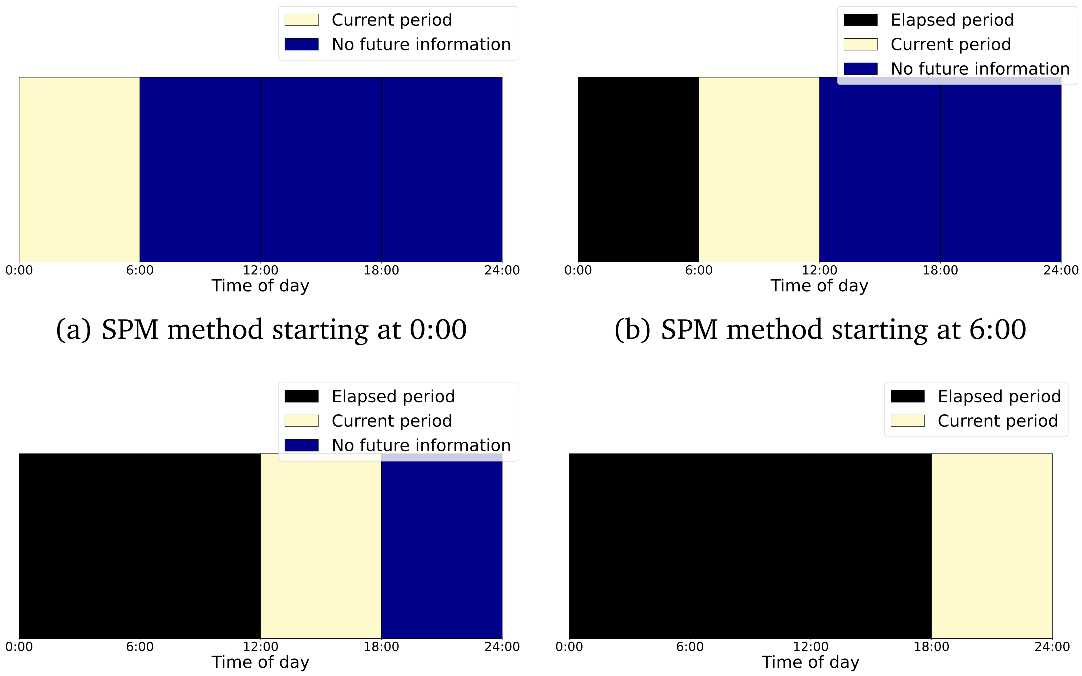
(d) MPM method starting at 18:00

Figure 4.1: Visual representation of the MPM method

each of which is provided with a different prediction of an unknown variable, which is used to solve a COP unique for that period. However, in the rolling horizon approach, it is assumed that predictions for later periods do not affect the schedule of an earlier period. For the BEB-CSP, this is not the case, as an incorrect prediction of no wind-generated energy excess during later periods may result in BEBs charging using non-clean energy during earlier periods, thus leading to sub-optimal solutions to the optimisation problem. Therefore, a Single-Period reactive Model (SPM) method is employed that makes assumptions that future forecasts will impact the early schedule generated and expects long-range predictions to be inaccurate.

As a result, SPM only relies on the prediction model for the current period. It will then schedule charges to consume the available wind-generated energy excess. However, any charges that use non-clean energy are postponed to the next period where possible, working under the assumption that more wind-generated energy excess will become available later.

Figure 4.2a shows a schedule for the entire operational day using SPM with predictions for the 0:00–6:00 period, assuming there is no knowledge of future wind-generated energy excess past 6:00. Figure 4.2b then shows the calculation



(a) SPM method starting at 0:00
(b) SPM method starting at 6:00
(c) SPM method starting at 12:00
(d) SPM method starting at 18:00

Figure 4.2: Visual representation of the SPM method

of a new schedule for the 6:00–12:00 period using a new set of predictions. This continues for the remaining periods as seen in Figures 4.2c and 4.2d. As a result, the SPM method should be more tolerant of mispredictions as it only uses the most accurate prediction model. It should be noted that for both examined models, a full charging schedule is created at every checkpoint. Therefore, any potential loss in connectivity will not impact operations.

4.2.2 Wind-generated energy excess dataset

As outlined in Chapter 2 Section 2.7.6 (Page 36) lifecycle emissions of BEBs are lowest when charged solely on renewable energy. In Chapter 3 the high capacity of wind-generated power is discussed, with the expected growth of this capacity extrapolated into an Upscaled Excess dataset, representing the potential wind-generated energy by 2026. This Upscaled Excess showcases periods where electricity generated by wind turbines exceeds the national grid's demand. Ordinarily, this clean energy may be lost if a Battery Energy Storage System (BESS) is not in place. As such, the TOU windows are used to represent times when the Upscaled Excess is above the value of 0, as it represents the use of clean energy to charge BEBs, while electricity outside of these windows can be considered non-clean energy.

For evaluating the performance of the BEB-CSP using TOU windows based on Upscaled Excess, 14 test days are used to determine the impact that different amounts of clean energy will have on solutions. Each of these days ranges from the 14th to the 27th of February 2022. These dates were selected, as they showcase a variety of situations regarding the availability of clean energy.

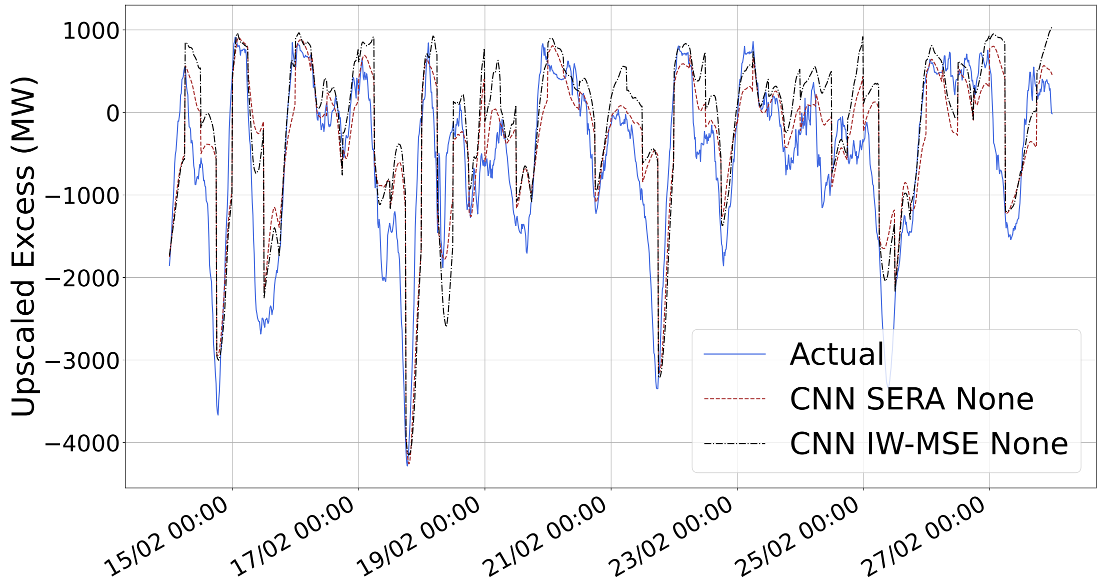


Figure 4.3: Actual Upscaled Wind-generated Energy Excess versus CNN SERA None and CNN IW-MSE None 6-hour model predictions

Figure 4.3 shows the availability of clean energy across the 14 examined days, as well as the predicted amount of clean energy from the two best configurations from Chapter 3 with a 6-hour forecast horizon. As shown, some days showcase a large amount of clean energy, such as the 26th where clean energy is available all day. However, some days have little or no Upscaled Excess above 0, such as the 17th, 18th, 21st, and 25th. This provides a variety of days to examine the impact of different amounts of clean energy availability on the generation of solutions for the BEB-CSP. This figure also clearly shows that the CNN IW-MSE None prediction model used overestimates the amount of clean energy available compared to CNN SERA None. This may have a detrimental impact on the results of the BEB-CSP.

4.2.3 City bus route datasets

The BEB-CSP is evaluated using three bus route datasets. These datasets represent three major cities in Ireland: Cork, Galway, and Limerick. Each has a unique number of buses, bus routes, stations, and operational times, allowing

Table 4.1: Cork, Galway, and Limerick dataset characteristics

Dataset	Bus routes	Buses	Stations	Operational time
Cork	11	81	578	5:30–24:00
Galway	6	33	254	5:50–23:32
Limerick	6	23	288	6:30–23:28

for a diverse evaluation regarding the impact of dataset size on solution quality. These datasets were developed by an as-of-yet unpublished report completed for *Work Package 2 Milestone 2* for the *SMARTeBuses* project [LA20][1]. The generation process of these datasets is briefly described below.

The generation of these datasets consists of two parts. First, the bus routes, stops, and arrival times are collected from the Bus Éireann website[2]. Global Positioning System (GPS) coordinates are collected, and the *haversine* formula is used to calculate the distance between each stop.

Next, a heuristic function is used to determine the number of buses assigned to each bus route. Each bus route has several *trips* that must be assigned to a bus. A bus is assigned a trip, then, based on a minimum speed parameter v, the arrival time at the endpoint of the trip is calculated. The bus is then assigned another trip, starting from the previous endpoint, if another feasible trip exists with regard to time. Once no more trips can be added to a bus, a new bus is added to the route. This process continues until all trips in the route have been assigned. Additional work was also carried out to include driver rests every 20 Kilometre (km), however, this is not used in this dissertation. For all instances used in this chapter, the speed parameter v's value is set to 20 Kilometres per hour (km/h). However, the average speed used by buses may be different from this value.

Table 4.1 shows the statistics for the resulting datasets. As can be seen, the Cork city dataset is the largest, with 81 buses operating from 5:30 to midnight, as well as double the number of stations of the other datasets. One deviation from the report [LA20] is the exclusion of the bus route 350 Ennis (Bus Station) - Galway (Bus Station) from the Limerick city dataset. This is an intercity bus route and does not match the urban nature of the bus routes featured in the rest of the datasets. The removal of this route is included in the statistics feature in Table 4.1.

[1] SMARTeBuses project website can be found at `https://smartebuses.github.io/web/`

[2] GPS location of the bus stations and timetables are available from https://www.buseireann.ie.

Table 4.2: Chapter 4 constants

Constant	Description
$\|B\|$	Set of identical buses
b	The b^{th} BEB of set B
$\|S_b\|$	Sequence of ordered stations for BEB b
s_i	The i^{th} stop of sequence S_b
X_i	Binary parameter which is 1 when there is a charging station at stop i
C_{min}	Minimum allowed capacity of a BEB battery
C_{max}	Maximum allowed capacity of a BEB battery
Δt	Max. allowed deviation time from the original schedule
D_{ij}	Energy required to travel from stops i to j
T_{ij}	Time needed to travel from stops i to j
τ_{bi}	Scheduled arrival time of BEB b at stop i
β	Maximum charge time.
γ	Minimum charge time.
$\mathcal{R}$	Charge rate in kWh.
M	Suitably large enough number to deal with time-based constraints
ϕ	A value between 0 and 1 representing how much of a reduction is applied to non-clean energy
Ω	End time of a 6-hour prediction horizon
$\|\Gamma\|$	Sequence of ordered TOU windows
Γ_k	The k^{th} TOU window of Γ
P_k	Starting time of TOU window Γ_k
Q_k	Ending time of TOU window Γ_k
ce_k	Total amount of clean energy available during TOU window Γ_k

4.3 Constraint Optimisation Problem model

This section describes the COP model for the BEB-CSP examined in this chapter. The nomenclature for the constants and variables used can be found in Tables 4.2 and 4.3 respectively.

Table 4.3: Chapter 4 variables

Variable	Domain	Description
x_{bi}	$\{0, 1\}$	Binary variable which is assigned the value of 1 when BEB b charges at stop i
t_{bi}	$[0, 24]$	Continuous variable representing actual arrival time of BEB b at stop i in an hour decimal format
c_{bi}	$[C_{\min}, C_{\max}]$	Continuous variable representing battery capacity of BEB b at stop i
ct_{bi}	$[0, \beta]$	Continuous variable representing time BEB b spent charging at stop i
e_{bi}	$[0, \mathcal{R} \cdot \beta]$	Continuous variable representing energy gained by BEB b at stop i
Z_{bdij}	$\{0, 1\}$	Binary value assigned the value 1 when stop i of BEB b and stop j of BEB d are the same charging station
z^{BD}_{bdi}	$\{0, 1\}$	Binary variable which is assigned 1 when BEB b arrives after BEB d finishes charging at stop i
z^{DB}_{dbi}	$\{0, 1\}$	Binary variable which is assigned 1 when BEB d arrives after BEB b finishes charging at stop i
nc_{bi}	$[0, \mathcal{R} \cdot \beta]$	Continuous variable representing the total amount of non-clean energy used by BEB b charging at stop i
kt_{bik}	$\{0, 1\}$	Binary variable which is assigned 1 when the time BEB b spends at stop i overlaps with Γ_k
wt_{bik}	$[0, \beta]$	Continuous variable representing the amount of time BEB b spends charging from the grid at stop i using clean energy from Γ_k
ce_{kbi}	$[0, \mathcal{R} \cdot \beta]$	Continuous variable representing the amount of clean energy used in Γ_k by BEB b at stop i
ase_{bi}	$\{0, 1\}$	Binary variable which is assigned 1 when BEB b arrives at stop i before the end of a 6-hour prediction horizon.
r_{bi}	$[0, \mathcal{R} \cdot \beta]$	Continuous variable representing the energy reduction for BEB b charging at stop i if it takes place after a 6-hour prediction horizon

4.3.1 Constraints

$$
\begin{aligned}
&\forall_{b \in B} \forall_{s_i \in S_b} : \\
&\quad C_{\text{min}} \leq c_{bi} + e_{bi} \leq C_{\text{max}} && (4.1)\\
&\quad \beta \cdot x_{bi} \geq ct_{bi} && (4.2)\\
&\quad X_i \geq x_{bi} && (4.3)\\
&\quad e_{bi} \leq \mathcal{R} \cdot ct_{bi} && (4.4)\\
&\quad ct_{bi} \geq \gamma \cdot x_{bi} && (4.5)
\end{aligned}
$$

Constraint (4.1) restricts the battery capacity to be within the bounds of the minimum and maximum values, while Constraint (4.2) sets the maximum allowed charge time to the value of β. Constraint (4.3) enforces that a charging station must be present for a BEB to charge. Constraint (4.4) then indicates the energy gained by charging BEB b at stop i is proportional to the charging rate ($\mathcal{R}$) and the charge time at stop i. Constraint (4.5) enforces a minimum charge time on all charges, while Constraint (4.6) denotes the capacity of BEB b at stop i.

$$
\begin{aligned}
&\forall_{b \in B} \forall_{s_i \in S_b \setminus s_0, j=i-1} : \\
&\quad c_{bi} \leq c_{bj} + e_{bj} - D_{ij} && (4.6)\\
&\quad t_{bi} \geq t_{bj} + ct_{bj} + T_{ij} && (4.7)\\
&\quad \Delta t \geq t_{bi} - \tau_{bi} && (4.8)\\
&\quad \Delta t \geq \tau_{bi} - t_{bi} && (4.9)
\end{aligned}
$$

Constraint (4.7) sets the arrival time of BEB b at stop i to be greater than the arrival time at the previous stop, plus the time spent charging, and the travel time between stops i and j. Constraints (4.8) and (4.9) denote the deviation time of BEB b at stop i with regards to the expected arrival time. Therefore, the arrival time of a BEB at any of its stops will be within the maximum deviation time from the original schedule.

$$\forall_{b,d \in B | b \neq d} \forall_{s_i \in S_b, \forall_{s_j} \in S_d | s_i = s_j} :$$

$$Z_{bdij} \leq x_{bi} \tag{4.10}$$

$$Z_{bdij} \leq x_{dj} \tag{4.11}$$

$$x_{bi} + x_{dj} \leq Z_{bdij} + 1 \tag{4.12}$$

$$t_{bi} \geq t_{dj} + ct_{dj} - M \cdot z^{DB}_{dbi} \tag{4.13}$$

$$t_{dj} \geq t_{bi} + ct_{bi} - M \cdot z^{BD}_{bdi} \tag{4.14}$$

$$z^{BD}_{bdi} + z^{DB}_{dbi} - (1 - Z_{bdij}) \leq 1 \tag{4.15}$$

Additionally, to ensure the unary nature of charging stations, there must be no overlapping charging times for BEBs at the same stop. Constraints (4.10)–(4.12) determine if BEBs b and d are scheduled to charge at the same charging station. Constraints (4.13) and (4.14) assert the order in which BEB arrives and finishes charging first at the shared charging station. Constraint (4.15) enforces that Constraints (4.13) and (4.14) are mutually exclusive for BEBs sharing a charging station, which is only applied if the scheduled arrival times of BEBs b and d are close enough to overlap at a given stop.

4.3.2 Time-of-Use windows

Constraints (4.16)–(4.26) capture the amount of clean energy used in Γ_k by every bus at every stop.

$$\forall_{b \in B} \forall_{s_i \in S_b} \forall_{k \in \Gamma} :$$

$$t_{bi} + ct_{bi} + M \cdot (1 - kt_{bik}) \geq p_k \tag{4.16}$$

$$t_{bi} - M \cdot (1 - kt_{bik}) \leq q_k \tag{4.17}$$

$$kt_{bik} \geq wt_{bik} \tag{4.18}$$

$$kt_{bik} \cdot \mathcal{R} \geq ce_{kbi} \tag{4.19}$$

$$x_{bi} \geq kt_{bik} \tag{4.20}$$

$$q_k - t_{bi} - \sum_{l=0}^{k-1} wt_{bil} + M \cdot (1 - kt_{bik}) \geq wt_{bik} \tag{4.21}$$

$$t_{bi} + ct_{bi} - p_k + M \cdot (1 - kt_{bik}) \geq wt_{bik} \tag{4.22}$$

$$ce_{kbi} \leq wt_{bik} \cdot \mathcal{R} \tag{4.23}$$

$$\forall_{b \in B} \forall_{s_i \in S_b} :$$

$$\sum_{k=0}^{|\Gamma|} wt_{bik} \leq ct_{bi} \tag{4.24}$$

$$nc_{bi} \geq e_{bi} - \sum_{k=0}^{|\Gamma|} ce_{kbi} \tag{4.25}$$

$$nc_{bi} \leq e_{bi} \tag{4.26}$$

Constraints (4.16) and (4.17) determine if the time BEB b spends at stop i overlaps with TOU window Γ_k. If so, the variable kt_{bik} is assigned the value of 1. Otherwise, it is assigned 0. Constraint (4.18) asserts that a BEB can only charge using the clean energy available in Γ_k if the time BEB b spends at stop i overlaps with Γ_k. Constraint (4.19) explicitly states that the amount of clean energy used for Γ_k is 0 if the time BEB b spends at stop i does not overlap with Γ_k. Constraint (4.20) enforces that BEB b cannot use clean energy at stop i if it does not charge. Constraints (4.19) and (4.20) are implied by the rest of the model. However, they are explicitly stated to assist the search process. Constraints (4.21) and (4.22) determine the maximum amount of time BEB b can spend using clean energy from Γ_k at stop i, while Constraint (4.21) determines how much time remains between the end of Γ_k and the time BEB b starts using energy from Γ_k at stop i.

For example, if k is 0, then Constraint (4.21) determines the time between the end of the Γ_k and the arrival time of the BEB, as there is no previous Γ to use energy from. Constraint (4.22) determines the time between BEB b finishing its charge at stop i and the beginning of Γ_k. The values from these constraints are used to constrain the upper bounds of wt_{bik}. In particular, Constraint (4.21) requires information about how much time BEB b has spent charging from previous instances of Γ_k. This determines how much time is left before the end of Γ_k. Constraint (4.23) associates the total amount of clean energy used for BEB b at stop i during Γ_k to be the amount of time the bus spent charging from that Γ, multiplied by the charging rate constant. Constraint (4.24) ensures that the total amount of time BEB b spends charging at stop i using clean energy does not exceed the total amount of time the bus spends charging at that stop. Constraint (4.25) enforces that the total amount of non-clean energy used is equal to the total energy used to charge BEB b at stop i, minus the sum of the clean energy used to charge the bus across all Γ. Finally, Constraint (4.26) asserts that the amount of non-clean energy used to charge BEB b at stop i must be less than the total amount of energy used by BEB b at stop i.

$$\forall_{k \in \Gamma} : \quad ce_k \geq \sum_{b=0}^{|B|} \sum_{i=0}^{|S_b|} ce_{kbi} \tag{4.27}$$

The amount of clean energy (ce_k) available within Γ_k is limited. Therefore, any energy exceeding this value is considered non-clean. Constraint (4.27) enforces this.

4.3.3 Single-period model

As outlined in Section 4.2.1 mispredictions can cause losses in optimality for COPs such as the BEB-CSP. Therefore, the SPM method is proposed to curtail the impact of mispredictions, where non-clean energy charge events are prioritized at future times after the current checkpoint. This way, charges are spread throughout the day, and the optimisation model has more opportunity to reschedule non-clean energy charges with more accurate predictions. To achieve this, the optimisation model applies a small reduction to the amount of energy used if the charge takes place after the current checkpoint. While the previous section defined the COP model for MPM, modifications are made to represent the SPM.

$$
\begin{aligned}
&\forall_{b \in B} \forall_{s_i \in S_b} : \\
&\quad t_{bi} + M \cdot ase_{bi} \geq \Omega && (4.28) \\
&\quad r_{bi} \leq \phi \cdot e_{bi} && (4.29) \\
&\quad r_{bi} \leq M \cdot (1 - ase_{bi}) && (4.30)
\end{aligned}
$$

Constraint (4.28) determines if the arrival time of BEB b at stop i (t_{bi}) occurs before the start of the next checkpoint (Ω). In this case, the binary variable ase_{bi} equals 1. Constraints (4.29) and (4.30) enforce how much of a reduction (ϕ) is applied to non-clean energy if the charge takes place afterwards and sets the reduction to zero if the charge takes place during the current checkpoint. Charges starting in the current checkpoint and finishing during the next one are not considered for the non-clean energy reduction provided by the SPM reactive method.

$$
nc_{bi} \geq e_{bi} - r_{bi} - \sum_{k=0}^{|K|} ce_{kbi} \tag{4.31}
$$

To apply the reduction to the non-clean energy used, Constraint (4.25) is replaced with Constraint (4.31). It is important to notice that Constraint (4.31) is identical to Constraint (4.25). However, the reduction for charges taking place after the current checkpoint (i.e., r_{bi}) is applied to the amount of non-clean energy used by BEB b at stop i. As a result, the COP will attempt to delay any charge using non-clean energy until after the current checkpoint. Thus scheduling charges later in the day and preventing clustering of BEB charging in the morning.

4.3.4 Objective function

$$
\min : \sum_{b=0}^{|B|} \sum_{i=0}^{|S_b|} nc_{bi} \tag{4.32}
$$

The objective function (4.32) aims to minimise the total amount of non-clean energy used in the schedule based on the clean energy information provided to the COP model via the prediction models used. In the scenario where $|\Gamma| = 0$ (i.e., there is no knowledge of the availability of clean energy), the objective function will minimise the total amount of energy used.

4.4 Experiment and Evaluation

4.4.1 Experiment configurations

The outlined BEB-CSP is evaluated using three Irish cities, i.e., Cork, Limerick, and Galway, with real data for two weeks, i.e., 14^{th} of February 2022 to 27^{th} of February 2022 as outlined in Sections 4.2.3 and 4.2.2 respectively.

For the initial experiments, an energy consumption rate of 1 kWh/km is considered. Further sensitivity analysis explores the impact of a discharge rate of 1.2 kWh/km. These values are common in contemporary literature for single-decker BEBs, as discussed in Chapter 2 Section 2.7.3 (Page 31). A charge rate ($\mathcal{R}$) of 600 kWh is used. The average speed of BEBs is considered 35 km/h when calculating travel time and energy expenditure in the dataset, which is the average expected speed of BEBs based on the datasets used. However, certain buses must travel faster to reach their destination per official timetables. In such cases, the minimum feasible speed is consistent with the maximum deviation times.

A value of 25 is assigned to Big-M, as this value is suitably large enough when dealing with the time-related Constraints (4.13), (4.14), (4.16), (4.17), (4.21), (4.22), (4.28), and (4.30). This is because time is represented as a continuous variable, in an hour decimal format.

The BEB-CSP is evaluated with multiple values for $C_{\max}$ (i.e., 120, 180, and 240 kWh) and $C_{\min}$ is set to a flat 12 kWh. The value of β scales with $C_{\max}$, allowing the BEB to charge up to 80% of their maximum capacity in a single charge. For example, the β value used for experiments with $C_{\max}$ of 120 is 0.16. Representing 9.6 minutes in an hour decimal format. γ is set to 1 minute (i.e., $0.01\dot{6}$). Furthermore, values of 5 and 10 minutes are used for Δt (i.e., $0.08\dot{3}$ and $0.1\dot{6}$ respectively). ϕ is set to 0.01 (or a 1% discount on non-clean energy used during later checkpoints).

As outlined in section 4.2 it is assumed that BEBs do not start with full battery capacity. To represent this, each BEB start with a capacity of 30 kWh regardless of the selected value of $C_{\max}$. This allows enough capacity for each bus to travel to its first fast-charging station while respecting the $C_{\min}$ value of 12 kWh. Furthermore, the use of a static starting capacity ensures fairness when comparing different battery capacities.

As discussed in Section 4.2.3, the earliest operational time of the datasets examined is 5:30. As a result, any TOU windows that occur before this cannot be used by the BEB fleet, and are not considered when evaluating results.

Four scenarios are considered regarding information about future TOU windows. The first is $|\Gamma| = 0$, which is a naïve approach that assumes no prior knowledge of $|\Gamma|$. Predicted configuration 1 (P_1), where information is provided from the best-performing predictive configuration from Chapter 3 (i.e., CNN SERA None). Predicted configuration 2 (P_2) uses estimations from the second-best configuration from Chapter 3 (i.e., CNN IW-MSE None). Finally, an Ideal scenario is examined where 100% accurate information about TOU windows is available. The motivation for evaluating the two best configurations from Chapter 3 is to determine the impact of a predictive model that generalises well versus one that can more accurately estimate extreme amounts of Upscaled Excess.

Additionally, the energy available in each TOU window is portioned out to each dataset based on how much energy is needed to operate the bus system. In total 26.371 MegaWatt (MW) is needed to operate the bus systems of all three cities, with Cork requiring 16.263 MW, Galway using 6.047 MW and Limerick needing 4.059 MW, assuming an energy consumption rate of 1 Kilowatt hours per kilometre (kWh/km). As a result, the Cork, Galway, and Limerick datasets receive 61.67%, 22.93%, and 15.39% of the available clean respectively. Additional experiments are also conducted where each dataset receives 100% of the available clean energy.

As previously stated, it is assumed charging infrastructure for each examined city is already available. The placement of charging station infrastructure is based on Opportunity Charging (OC), where charging stations are placed every X km [Laj18]. Two scenarios are explored for values of X:

- Infrastructure A (Inf-A): Charging stations are placed every 12 km on each bus route, for a total of 53 chargers in Cork, 13 in Limerick, and 19 in Galway.
- Infrastructure B (Inf-B): Charging stations are placed every 15 km on each bus route, for a total of 43 charges in Cork, 12 in Limerick, and 11 in Galway.

Table 4.4: Non-clean energy usage (in kWh) for Inf-A. and Δt = 5 minutes.

City	$C_{\max}$	Method	$\|\Gamma\| = 0$	P_1	P_2	Ideal
Limerick	120	MPM	3296.54	**3256.11**	3222.12	**3073.70**
		SPM	3453.85	3314.77	**3212.08**	3097.92
	180	MPM	3302.85	**3250.97**	3256.86	**3073.70**
		SPM	3239.15	3288.33	**3221.95**	3125.36
	240	MPM	3293.06	**3272.14**	**3216.52**	**3073.70**
		SPM	3256.60	3273.64	3218.62	3110.55
Galway	120	MPM	4234.81	**4153.55**	**4095.72**	**3910.47**
		SPM	4202.10	4186.60	4108.67	3944.81
	180	MPM	4232.88	**4151.94**	4138.49	**3910.47**
		SPM	4202.46	4222.06	**4111.67**	3944.42
	240	MPM	4231.32	**4162.13**	4171.80	**3910.46**
		SPM	4206.32	4190.86	**4099.29**	3957.20
Cork	120	MPM	11347.00	**11200.96**	**11021.31**	**10528.24**
		SPM	11496.84	11391.11	11032.60	10589.42
	180	MPM	11341.11	**11153.22**	**10920.01**	**10449.92**
		SPM	12089.81	11418.43	10986.22	10509.87
	240	MPM	11330.13	**11271.72**	10976.96	**10429.40**
		SPM	11687.55	11326.04	**10958.20**	10483.94

A pre-computed warming solution is used for the first schedule in a working day. For this purpose, $C_{\max}$ is set to 120 kWh, Δt is assigned 5 minutes, and $|\Gamma| = 0$ scenario is used. These warm solutions took 6.1, 0.7, and 0.4 seconds to find an optimal solution for Cork, Galway, and Limerick, respectively. Furthermore, when a schedule is recalculated due to a change in the TOU windows (i.e., mispredictions in the previous checkpoint), the previous solution is used as a warm solution. It should also be noted that values assigned to variables that occur before the recalculation of the schedule are not changed.

Experiments were implemented in C++ and conducted on a server using a 2.5 Gigahertz (GHz) Intel Xeon W-2175 Central Processing Unit (CPU) and 62GB of Random Access Memory (RAM). CPLEX version 12.10 is used to implement the COP as a Mixed Integer Linear Programming (MILP) model with a 12-minute timeout per experiment.

4.4.2 Evaluation

Table 4.4 shows the average results over all experiments with Δt = 5 minutes and Inf-A. for the MPM and SPM reactive methods with different values for $C_{\max}$. $|\Gamma| = 0$ shows the non-clean energy consumption assuming that no prior knowledge about Γ (i.e., the TOU windows) is available. P_1 presents the results of the optimisation model using predictions from the best-performing configuration in Chapter 3, while P_2 uses the second-best-performing configuration. Finally, Ideal presents the results with perfect information about TOU windows. Solutions generated using $|\Gamma| = 0$, P_1, and P_2 are overlayed with the historical values of the available clean energy. The average consumption of non-clean energy across the fourteen reference days is then reported.

As expected, the best solutions generated by the COP model occur when perfect knowledge of future clean energy availability and the MPM schedule creation method are used. Investigating the two prediction scenarios, it is clear that P_2 outperforms P_1 in all but two categories (Galway 240 MPM and Limerick 180 MPM). However, these are still worse than the P_2 results using SPM. Furthermore, when comparing results for P_1 to $|\Gamma| = 0$ shows that, in some scenarios, the COP using estimations can produce worse results than not knowing when clean energy is available. This is consistent with previous works showing COP models need to be considered in the predictive model training process [GIOS14]. However, the steps taken in Chapter 3, such as the specialised test metrics, were not enough to determine the best configuration with regards to the impact on the result for the COP. Instead, it is clear that the second-best configuration, which is better at estimating values in the 99^{th} percentile of the training dataset, provides better predictions for the COP. The reasoning for this is due to the distribution of energy across the examined datasets. Both Galway and Limerick receive little clean energy (22.93% and 15.39% respectively). Therefore, more accurate estimation of large amounts of clean energy has a higher impact on solution quality compared to better generalisation of low amount of energy. With this in mind, the results show that if the correct prediction model is used, even if it is not fully accurate, integrating a learning component into schedules has a notable benefit.

As can be seen, the SPM method outperforms the MPM method in 5 out of 9 scenarios (P_2 column). This can be attributed to its ability to delay non-clean energy charge events until after the current checkpoint. The MPM method schedules non-clean energy consumption during each checkpoint with the current knowledge, as a result, this method may be susceptible to mispredictions from the 12, 18, and 24-hour models, which were shown to have higher errors in Chapter 3 Section 3.5.3 (Page 81).

It should also be noted that results using SPM for the Cork dataset perform

Table 4.5: Relative performance difference for Table 4.4

City	$C_{\max}$	Method	P_{best} vs $\lvert\Gamma\rvert = 0$	Ideal vs P_{best}
Limerick	120	MPM	2.30	4.82
		SPM	7.52	3.68
	180	MPM	1.59	5.77
		SPM	0.53	3.09
	240	MPM	2.37	4.64
		SPM	1.17	3.47
Galway	120	MPM	3.39	4.73
		SPM	2.27	4.15
	180	MPM	2.28	5.83
		SPM	2.20	4.24
	240	MPM	1.66	6.43
		SPM	2.61	3.59
Cork	120	MPM	2.95	4.68
		SPM	4.20	4.18
	180	MPM	3.85	4.49
		SPM	10.04	4.53
	240	MPM	3.21	5.25
		SPM	6.65	4.53

worse compared to MPM. This suggests that larger datasets see fewer benefits from delaying non-clean charges. This is due to charging events using non-clean energy being delayed. This can cause BEBs to insufficent free battery capacity to capitalise on clean energy available later in the day compared to MPM. This topic is explored later in this section. Out of the 3276 experiments reported in this table, only 5 were unable to find optimal solutions, with an average optimality gap of 0.012%.

Table 4.5 shows the relative performance difference for Table 4.4. The column P_{best} vs $\lvert\Gamma\rvert = 0$ shows the relative performance difference between the naïve approach and the best results using a predictive model (P_1 or P_2) for that row. Finally, the Ideal vs P_{best} column shows the relative performance difference between results using the best P configuration results and accurate historical data. The relative performance difference represents the percentage improvement between the two scenarios. For example, values for the P_{best} vs $\lvert\Gamma\rvert = 0$ column are generated using the equation: $((\lvert\Gamma\rvert = 0 - P_{\text{best}})/P_{\text{best}}) \cdot 100$.

As shown, the optimisation model using predictions can reduce the amount of non-clean energy used by an average of 2.62% and 4.13% when using MPM and SPM respectively. When comparing the results of the COP model with predictions with those of Ideal, generated solutions are only 5.18% worse than the average optimal solution using MPM. Meanwhile, when using SPM, solutions generated with 100% accurate historical information are only 3.94% better than using prediction models. However, as seen in Table 4.4 SPM's results using Ideal are worse compared to MPM. When comparing the results for SPM P_{best} versus MPM in the Ideal scenario, the average relative performance difference is 4.89%. This means SPM is better when using predictions to generate solutions to the BEB-CSP and can handle mispredictions. It should be noted that P_2 is used for all P_{best} SPM results as it outperforms P_1. As such, the performance of SPM seems to also depend on the prediction model itself. This could indicate that the method performs well with a more accurate estimation of extreme values.

Table 4.4 also shows that increasing battery capacity may not reduce the non-clean consumed. This is particularly evident with results for Limerick and Galway in the Ideal scenario. In some cases, a battery capacity of 120 will use all available clean energy. When this occurs, increasing the capacity will not improve the solution quality. However, for larger datasets, increasing battery capacity does reduce the amount of non-clean energy used. This is not only due to increased consumption of electricity but also to the operational hours of the bus system. The Cork dataset includes buses that start their operational day at 5:30 while the earliest bus route in the Galway dataset starts at 5:50 and in Limerick at 6:30. On some days, there is clean energy available early in the morning and late at night. As a result, the Cork dataset can make use of this additional clean energy that other datasets cannot.

It is also observed that the quality of solutions generated with the SPM method and Ideal can decrease when battery capacity is increased. This may be due to SPM's limited knowledge of future clean energy after the current checkpoint. While SPM delays charges until after the current checkpoint, this may result in "spikes" in energy consumption immediately following a checkpoint. This can mean BEBs do not have enough free capacity to charge using clean energy later in the same checkpoint. Larger battery capacities have larger β values, thus allowing them to gain more energy in a single charge, potentially leading to larger energy consumption or "spikes". This is also the cause of SPM's poor performance compared to MPM in the ideal scenario. As MPM in the ideal scenario with perfect information of when clean energy occurs. Therefore, the schedule generated ensures BEBs have enough free battery capacity to consume the available clean energy.

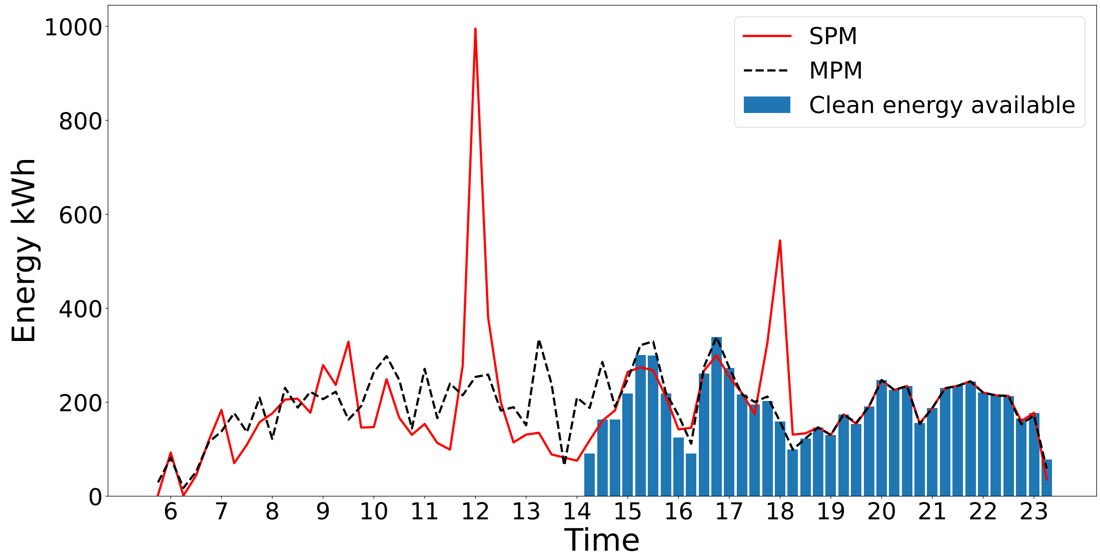


Figure 4.4: Comparison of energy used for SPM and MPM methods versus clean energy available for the 27th for Cork dataset, capacity 240, 5-minute deviation time, and the Ideal scenario

Figure 4.4 compares the amount of energy consumed by both the MPM and SPM method versus the amount of clean energy available for the Cork dataset for the 27th of February using 5-minute deviation time and 240 kWh capacity battery. This shows energy consumption "spikes" at 12:00 and 18:00 for the SPM method as previously mentioned. The COP model is shown to generate schedules that make full use of the clean energy available. Both methods consume similar amounts of energy, with SPM using 9.3 kWh less than MPM. However, SPM cannot make full use of the available clean energy compared to MPM due to the aforementioned reasons. As a result, SPM consumes 145.59 more kWh of non-clean energy compared to MPM.

Table 4.6 shows the impact of increasing deviation time to 10 minutes. As expected, the MPM model with Ideal shows an improvement in solution quality compared to Table 4.4. Similar patterns, such as the relationship between solution quality and battery capacity, are also observed. However, allowing a larger deviation time slightly deteriorates solution quality when using predictions for most experiments. This phenomenon is attributed to the prediction models' error, the increased deviation time may allow the optimisation model to schedule chargers during a given Γ_k with less clean energy than expected. While increasing the value for Δt increases solution quality for the Ideal scenario and MPM, it is only increased by 0.685% on average. However, increasing the value of this variable also increases the search space of the problem. Resulting in the compu-

Table 4.6: Non-clean energy usage for Inf-A. and Δt = 10 minutes.

City	C_{min}	Method	$\lvert\Gamma\rvert = 0$	P_1	P_2	Ideal
Limerick	120	MPM	3303.64	**3298.24**	3259.29	**3067.54**
		SPM	3360.26	3322.24	**3234.35**	3114.45
	180	MPM	3304.22	**3348.72**	3316.61	**3067.45**
		SPM	3315.33	3315.74	**3239.60**	3112.70
	240	MPM	3307.13	3371.09	3391.97	**3067.54**
		SPM	3334.60	**3323.35**	**3240.87**	3126.68
Galway	120	MPM	4233.23	**4190.28**	4168.52	**3892.87**
		SPM	4442.53	4265.00	**4142.57**	3946.41
	180	MPM	4255.35	**4264.25**	4263.16	**3892.62**
		SPM	4411.61	4283.12	**4161.87**	3945.49
	240	MPM	4258.33	4301.84	4275.77	**3892.62**
		SPM	4379.42	**4291.64**	**4149.69**	3982.89
Cork	120	MPM	11403.70	**11255.16**	**11018.52**	**10475.50**
		SPM	12160.71	11487.34	11035.62	10549.84
	180	MPM	11317.37	**11148.91**	10940.45	**10380.32**
		SPM	12274.28	11356.69	**10919.68**	10465.65
	240	MPM	11322.08	**11084.73**	10969.34	**10345.94**
		SPM	11825.65	11327.75	**10915.36**	10464.70

tation time for experiments using Δt = 10 minutes increasing by approximately 310% on average compared to Δt = 5 minutes.

Increasing the value for Δt also increase the number of experiments which could not find optimal solutions. In particular, of the 3276 experiments featured in Table 4.6, 41 are unable to find optimal solutions. However, much like the experiments using Δt = 5 minutes, the average optimality gap was small with each non-optimal solution being on average 0.034% away from the optimal solution.

Table 4.7 shows the relative performance difference for the results in Table 4.6. Here, the performance of MPM using P_{best} sees little improvement compared to the naïve scheduler, with schedules only improving by 1.13% on average. In some examined cases, solution quality using MPM and predictive models worsens compared to the naïve scenario, with the average solution for the Limerick dataset using a battery capacity of 240 performing 1.89% worse. It should be noted that this particular example is the only time in Table 4.7 where P_1 outperforms P_2. This can be attributed to the underestimation of available clean energy. For example, if clean energy occurs later in the day, and the prediction models with larger prediction horizons incorrectly estimate there is no clean energy available, then charges for BEBs may be scheduled earlier in the day. This

Table 4.7: Relative performance difference for Table 4.6

City	C_{max}	Method	P_{best} vs $\lvert\Gamma\rvert = 0$	Ideal vs P_{best}
Limerick	120	MPM	1.36	6.25
		SPM	3.89	3.84
	180	MPM	-0.37	8.11
		SPM	2.33	4.07
	240	MPM	-1.89	9.89
		SPM	2.89	3.65
Galway	120	MPM	1.55	7.08
		SPM	7.24	4.97
	180	MPM	-0.18	9.51
		SPM	6.00	5.48
	240	MPM	-0.40	9.84
		SPM	5.53	4.18
Cork	120	MPM	3.49	5.18
		SPM	10.19	4.60
	180	MPM	3.44	5.39
		SPM	12.40	4.33
	240	MPM	3.21	6.02
		SPM	8.33	4.30

would result in BEBs being unable to charge more using clean energy when the predictions update with more accurate information. This is discussed in more detail later in this section.

The Ideal scenario solutions in Table 4.7 are also shown to be on average 7.47% better than a schedule made using predictions when using MPM. When examining SPM, it can be seen that results using the P_{best} are on average 6.53% better than the naïve scenario, while the Ideal scenario only performs 4.38% better on average compared to P_{best}. When comparing the results for P_{best} SPM with those of Ideal MPM, there is an average relative performance difference of 5.85%. This once again shows SPM's suitability for the creation of BEB-CSPs when using prediction models compared to MPM.

Table 4.8 shows the average execution time in seconds and the average number of constraints for each city, method, Δt value across all dates, capacities, and scenarios. Here it can be seen that the biggest influence on the execution time and the number of constraints is the dataset used. For example, the Cork City dataset features over a million additional constraints compared to the Limerick City dataset.

As previously stated, increasing the value for Δt also increased the computation

Table 4.8: Average execution time and number constraints for all experiments featured in Tables 4.4 and 4.6.

City	Method	Δt (m)	Execution time (s)	# constraints
Limerick	MPM	5	1.72	442820.53
		10	2.11	495414.09
	SPM	5	1.26	400209.26
		10	2.72	452309.06
Galway	MPM	5	2.88	615787.74
		10	3.81	687053.53
	SPM	5	2.41	548995.09
		10	25.13	619221.85
Cork	MPM	5	11.34	1567144.93
		10	36.95	1763040.37
	SPM	5	15.20	1402917.32
		10	37.21	1596020.33

time, which is shown here. The average number of constraints for experiments featuring a Δt value of 10 minutes is also increased. This directly relates to Constraints (4.10)–(4.15). These constraints are only applied if the arrival times of BEBs b and d are close enough to overlap. As the maximum allowed deviation time increases, the range of potentially overlapping arrival times increases, leading to an increase in the number of constraints and execution time.

Of note is the impact of MPM and SPM on execution time and the number of constraints. SPM consistently features a lower number of constraints compared to the MPM. SPM features additional Constraints (4.32)–(4.30). However, as this model only takes into consideration clean energy information for the next six hours, the maximum length of $|\Gamma|$ is 24. This limits the number of times Constraints (4.16)–(4.27) feature in the optimization model. However, this does not reduce execution time, with experiments using SPM and a Δt value of 10-minutes consistently taking longer to execute compared to MPM with the same Δt value. This is further enforced by the fact that all 41 experiments which did not find an optimal solution when using a Δt value of 10-minutes used SPM.

Figure 4.5 shows box plots for the Galway dataset using the P_2 scenario and a 5-minute deviation time, comparing the methods and capacities used versus the solution quality. Here, points on each box plot denote a day from the 14 test days. Note on some days SPM method outperforms the MPM method. Upon closer inspection, these days are prone to mispredictions regarding the availability of clean energy later in the day. For example, on the 27$^{\text{th}}$ at the

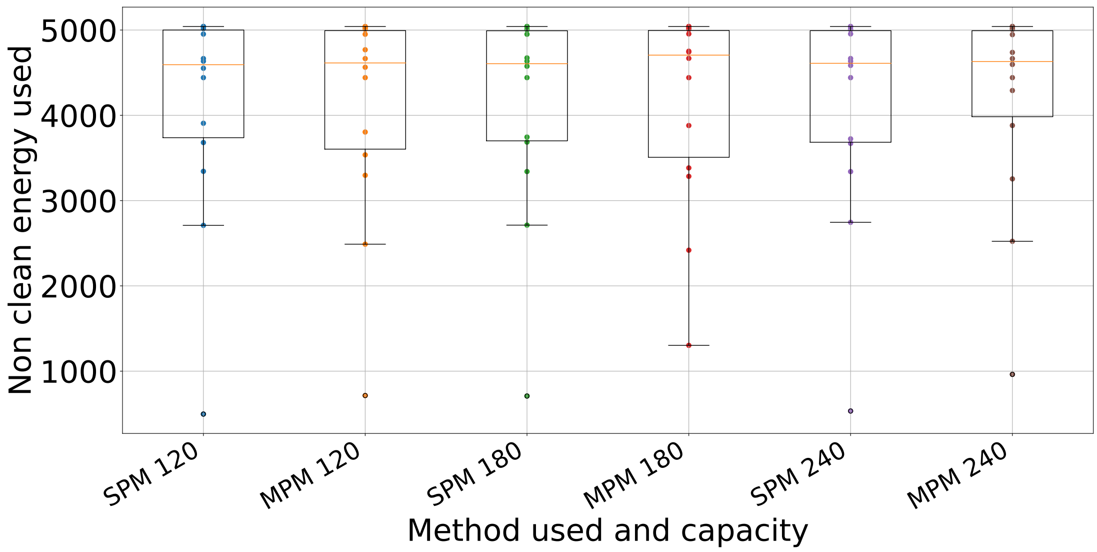


Figure 4.5: Comparison of methods and capacities used versus objective function score for the Galway dataset, 5-minute deviation time, using P_2

6:00 checkpoint, the 12 and 18-hour prediction models incorrectly estimate that there is no clean energy available after 6:00, where in reality, there is a significant amount of energy available after 14:00. At the noon checkpoint, new predictions indicates that clean energy is available between 16:30 and 18:00. Then at 18:00, the prediction model updates to show more clean energy available until the end of the day. Furthermore, the statistical outlier represents the 26th of February, where a large amount of clean energy is available throughout the day, thus leading to much lower consumption of non-clean energy compared to the other examined days.

Figure 4.6 compares the amount of energy consumption for MPM and SPM methods versus clean energy available for Galway on the 27th with 5-minute deviation time and 240 capacity for the P_2 scenario. Here, both methods consume similar amounts of energy in total and also use a large amount of energy before 13:00. Because the optimiser is not informed of any clean energy available after 18:00, the COP model will attempt to reduce the total amount of energy consumed. In these scenarios, the distribution of energy is not considered, and charge events are scheduled at the earliest possible convenience. For smaller datasets that require less energy overall, this results in most charge events occurring in the first half of the operational day. At the 12:00 checkpoint, the predictions are updated and showcase some available clean energy between 16:30 and 18:00. As this is the only known clean energy available at the time, MPM schedules a large number of charges for this time, resulting in a

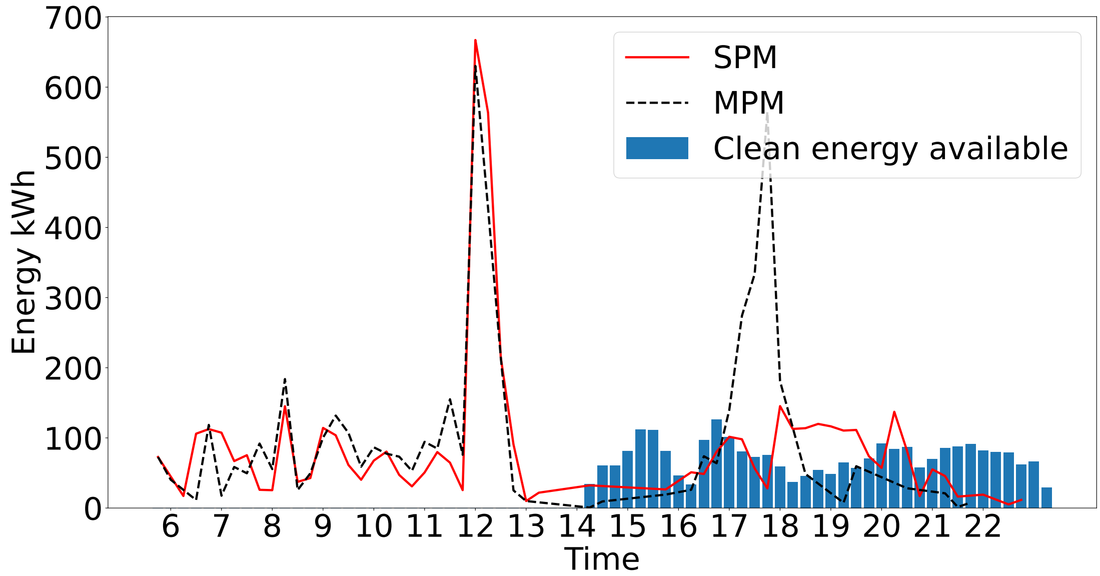


Figure 4.6: Comparison of energy used for SPM and MPM methods versus clean energy available for the 27th for Galway dataset, capacity 240, 5-minute deviation time, using P_2

large "spike". However, when predictions are updated again at 18:00, the MPM method cannot fully utilise the available clean energy, as it has already obtained most of the energy it needs to complete its daily operations. The SPM method appears to avoid this by delaying charges and only relying on predictions from the 6-hour model.

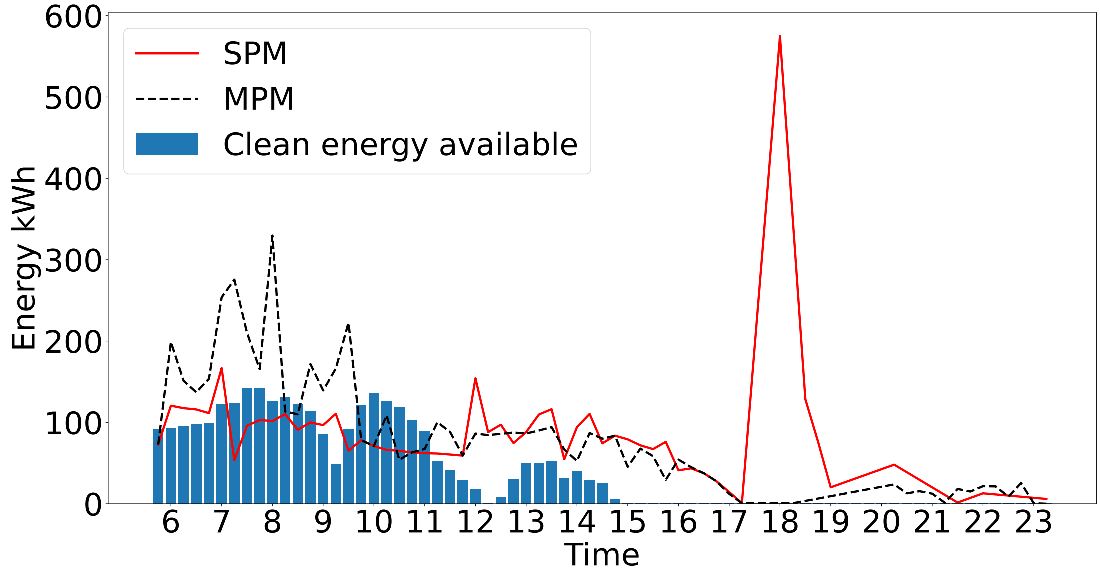


Figure 4.7: Comparison of energy used for SPM and MPM methods versus clean energy available for the 20th for Galway dataset, capacity 240, 5-minute deviation time, using P_2

Figure 4.7 shows the energy consumption for the MPM and SPM methods versus clean energy available for Galway on the 20th with 5-minute deviation time and 240 capacity using P_2. Shown here is the inverse of Figure 4.6 where all clean energy is available between 6:00 and 14:45. As can be seen, the MPM method consumes most of the clean energy available, while the SPM does not. While P_2 estimates clean energy is available between 6:00 and 12:00, the SPM method will still delay some charges until later in the day if it is unable to consume any more clean energy from the current checkpoint. As a result, the previously outlined phenomenon, where CPLEX schedules charges early in the day, works to the benefit of the MPM method, as more clean energy is available than was originally predicted. From this, an observation can be made that the most important factor for SPM's performance is if clean energy occurs at the start of the day (such as the 20th shown in Figure 4.7) or at the end of the day (as shown in Figure 4.6). However, due to the performance of SPM in Tables 4.4 and 4.5 it can be inferred that such days with the majority of clean energy available at the start of the day are less common.

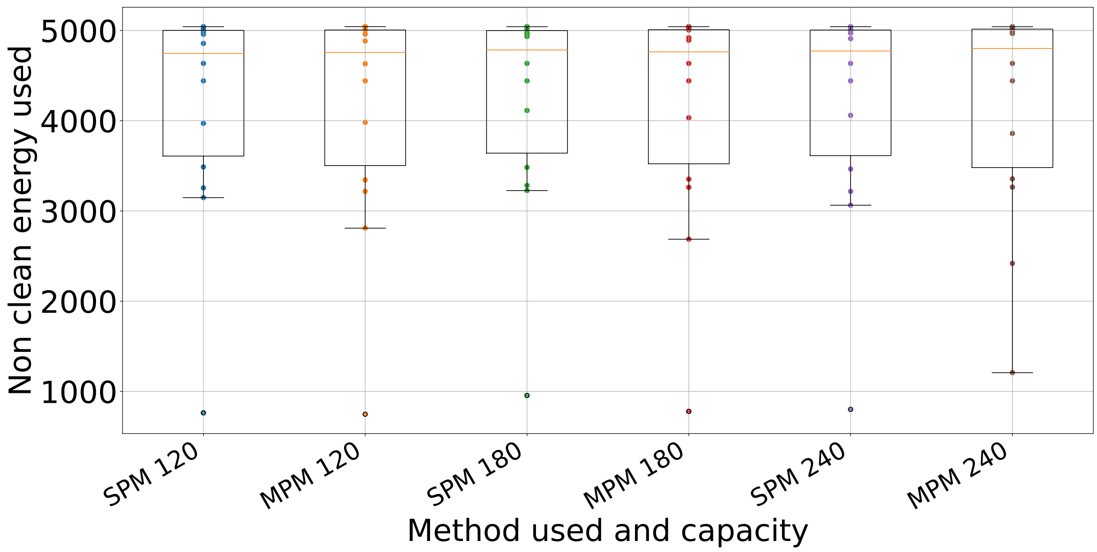


Figure 4.8: Comparison of methods and capacities used versus objective function score for the Galway dataset, 5-minute deviation time, and P_1.

Attention now shifts to the performance of scenario P_1. Figure 4.8 shows a box plot for the same dataset as Figure 4.5 using the P_1 scenario instead of P_2. As can be seen, P_1 clearly performs worse overall, as seen in the reduced mean line for all examined box plots compared to Figure 4.5. However, it is also clear that some days appear to perform better when using P_1. For example, when using MPM with a $C_{\max}$ value of 180, the 26th of February, which was previously included in the lower adjacency of the box plot, is now considered a statistical outlier. Additionally, the lower adjacency of MPM with $C_{\max}$ = 240 has increased, representing improved performance on some days, along with the decreased performance of 26th comparative to P_2. To further analyse this, the number of times P_1 outperformed P_2 for each of the 14 test days was compiled for all experiments featured in Table 4.4.

In total, the solutions generated when using the P_1 scenarios outperformed P_2 in 100% of the experiments conducted on the 27th of February. Additionally, P_1 performed well on the 24th and 14th, outperforming P_2 55% and 44% of the time respectively. Interestingly, on both the 14th and 24th there is almost no clean energy available. P_1 appears to consume less energy than P_2 for these days, leading to its improved performance. The reason for the difference in energy consumption lies in the accuracy of the predictions made. P_2 incorrectly estimates that there is clean energy available from 6:00 to 12:00. While P_1 predicts that there is no clean energy available after 7:45.

In actuality, there is no clean energy available after 6:00, as P_1 is closer to this, the schedules produced can minimise total energy consumed. However, P_2 provides incorrect information about clean energy availability, thus consuming more energy than needed, as it is offset by the "fake" clean energy used. This can be considered a "false positive" of the prediction model, where a positive amount of energy is estimated when, in actuality, there is none available. In Chapter 3 it was speculated that the influence of "false positives" was negligible compared to the improved performance of accurately estimating positive values. This is proven true by these findings. While individual days where there is no clean energy available may suffer, on average, a predictive model that can accurately estimate extreme amounts of Upscaled Excess provides the best results for the COP.

Figure 4.9 shows the same experiment as Figure 4.6 using P_1 instead of P_2. Here it can be seen that the schedule generated by P_1 does not have the "spikes" at 12:00 and 18:00 that are present in the results for P_2. This causes more clean energy to be consumed by the bus charging schedule, thus generating better solutions. It should be noted that, for SPM, both solutions consume similar amounts of energy in total (5047.30 for P_1 and 5041.57 for P_2). However, for MPM P_1 consumes 113.54 less kWh than P_2. The only discernible difference between the two scenarios for this experiment is that P_1 estimates that there is no clean energy available after the 12:00 checkpoint. At the same time, P_2 predicts that energy is available between 16:30 and 18:00. It is interesting to note

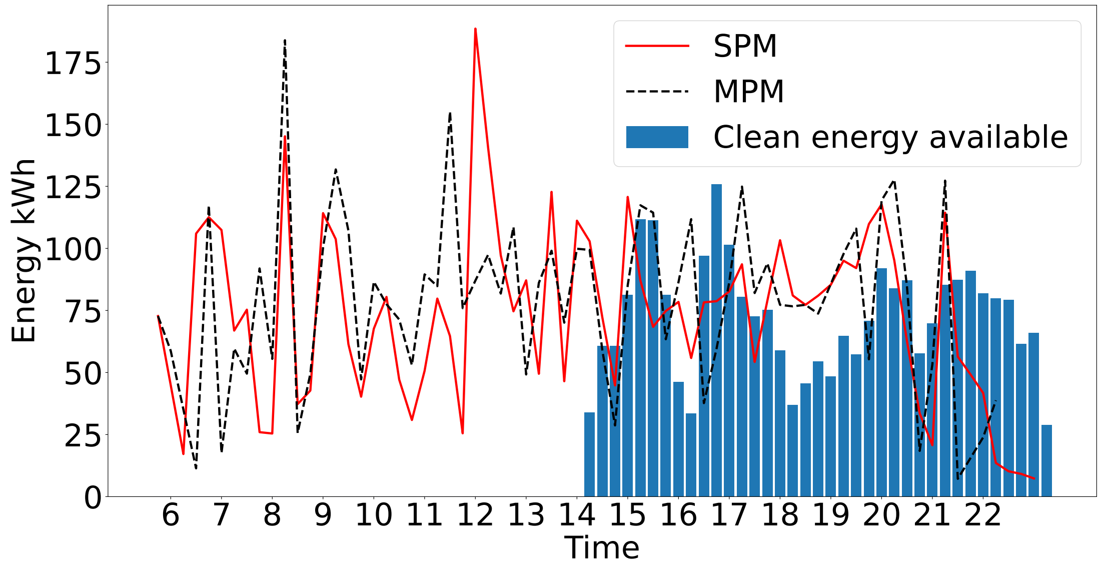


Figure 4.9: Comparison of energy used for SPM and MPM methods versus clean energy available for the 27th for Galway dataset, capacity 240, 5-minute deviation time, and P_1

that the energy consumption patterns between Figures 4.6 and 4.9 are identical before noon. This reinforces the previous observations as to the tendencies for CPLEX scheduling charging times if no clean energy is available.

The concentration of charge events around 12:00 for P_2 is likely done to ensure BEB has enough empty capacity to charge when clean energy is available later in the day. This can be seen in the low amount of energy consumed between 13:00 and 16:00 in Figure 4.6. This is juxtaposed by P_1, where energy consumption is more spread out. At 18:00, the predictions for both models are updated to show clean energy for the rest of the day. However, P_2 MPM cannot make full use of this clean energy due to the charging "spike" which occurred before 18:00 (which coincides with the previous estimations of clean energy from the 12:00 checkpoint). This, along with the more spread-out charging pattern of P_1, is responsible for the improved performance. It is also noteworthy that SPM can mitigate this to a degree. For example, when using MPM P_1 consumes 1026.66 less non-clean energy than P_2, while if both scenarios use SPM then the difference reduces to 498.81. This also highlights issues related to predictions. Even though the estimations provided by P_2 were more accurate than P_1, the latter was able to outperform it. This motivates a potential alternative to MPM and SPM where consumption of energy is evenly dispersed throughout the day if no clean energy is available. This way concentration of charging events such as the ones seen in Figure 4.6 can be avoided.

Table 4.9: Percentage of all clean energy consumed for Table 4.4

City	$C_{\max}$	Method	$\|\Gamma\| = 0$	P_1	P_2	Ideal
Limerick	120	MPM	36.82	**39.22**	42.42	**49.17**
		SPM	27.32	35.13	**42.44**	48.10
	180	MPM	36.50	37.24	41.86	**49.00**
		SPM	40.55	**39.69**	**42.21**	46.55
	240	MPM	36.92	35.94	41.85	**49.20**
		SPM	41.12	**39.85**	**42.19**	47.76
Galway	120	MPM	41.01	**49.98**	**54.95**	**61.40**
		SPM	42.29	48.58	53.84	60.61
	180	MPM	40.99	46.96	53.74	**61.35**
		SPM	42.92	**48.11**	**54.00**	60.63
	240	MPM	40.57	46.71	52.06	**61.32**
		SPM	42.32	**48.23**	**54.10**	60.11
Cork	120	MPM	48.48	**56.42**	62.77	**73.88**
		SPM	48.58	55.45	**63.56**	73.43
	180	MPM	49.05	**56.21**	**65.52**	**77.93**
		SPM	36.64	49.83	61.80	75.02
	240	MPM	49.23	**56.59**	**62.94**	**77.17**
		SPM	45.64	55.37	62.39	75.37

Table 4.9 shows the average percentage of the available clean energy consumed by the BEB fleet for the BEB-CSP results shown in Table 4.4. Here, P_2 appears to consistently consume more clean energy on average compared to P_1. Furthermore, it can be seen that the difference between the amount of clean energy consumed by $|\Gamma| = 0$, P_1, and P_2 scenarios increases as the size of the dataset increases. For example, the Limerick dataset in the P_2 scenario using MPM consumes 5.3% more clean energy on average than $|\Gamma| = 0$. This increases to 14.82% when using the Cork dataset. However, the gap between the percentage of clean energy consumed for the Ideal and P_2 scenarios is larger. When looking at experiments using MPM, the difference between Ideal and P_2 ranges from 6.45% clean energy consumed to 14.23%, with the Cork dataset featuring particularly larger differences over 11%. The existing pattern of SPM outperforming MPM on the smaller datasets of Limerick and Galway in both P_1 and P_2 scenarios is also prevalent here. This aligns with the previous observation relating to CPLEX's scheduling of charging events when there are inaccuracies in the predictions.

Since maximising the amount of clean energy consumed is not the objective function of the COP model, the percentage of clean energy consumed may be lower compared to other configurations, but may feature worse performance in Table 4.4. For example, the Cork dataset using capacity 240, MPM, and the Ideal scenario consumes 77.17% of the available clean energy, whereas the same configuration using 180 consumes 77.93%. As outlined in Table 4.4 battery capacity 240 outperforms 180 by 20.5 kWh. This is because the COP model is allowed to generate schedules that consume more energy than is needed, as long as the non-clean energy consumed is minimised. This results in symmetrical solutions where more clean energy is used.

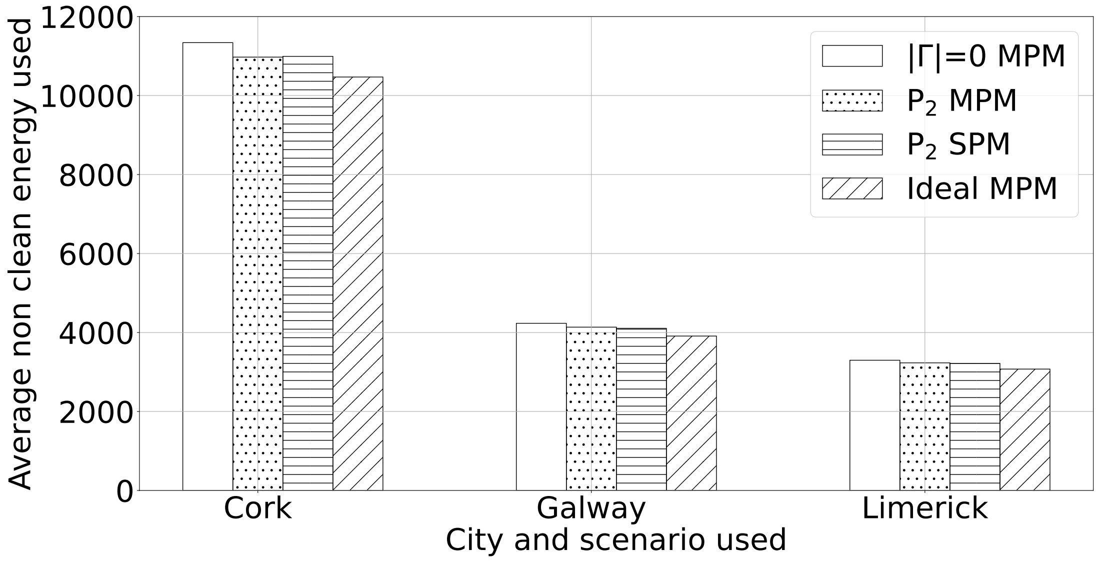


Figure 4.10: Performance of MPM and SPM P_2 versus MPM $|\Gamma|=0$ and MPM Ideal from Table 4.4

Figure 4.10 shows a bar chart with the average performance of both MPM and SPM P_2 methods versus MPM $|\Gamma|=0$ and MPM Ideal from Table 4.4 for each city. P_1 is excluded from this figure due to its worse performance on average compared to P_2. Here, it can be seen that the average difference between the three scenarios for MPM (i.e. $|\Gamma|=0$, P_2, and Ideal) heavily depends on the dataset used. For example, smaller datasets like Limerick show little difference while the larger Cork dataset shows more significant differences. Conversely, when comparing the results for P_2 MPM and SPM, it is clear that SPM performs better for the smaller datasets of Galway and Limerick, with little differences between the two methods observed for the Cork dataset. This further enforces previous findings as to the performance of the examined methods.

As previously outlined, additional experiments were conducted to evaluate the impact of certain parameters for the BEB-CSP. Table 4.10 reports the non-clean energy used in experiments where each city is given 100% of the available clean energy instead of being proportionally split. Here it can be seen that the quality

Table 4.10: Non-clean energy used for experiments with 100% energy distribution, and Δt = 5 minutes for Inf-A.

City	$C_{\max}$	Method	$\|\Gamma\| = 0$	P_1	P_2	Ideal
Limerick	120	MPM	2903.91	**2774.42**	**2909.94**	**2497.54**
		SPM	3183.45	2817.32	2918.94	2505.16
	180	MPM	2903.49	2769.01	**2823.81**	**2485.61**
		SPM	3006.26	**2741.72**	2963.71	2508.23
	240	MPM	2883.40	**2726.61**	**2787.82**	**2485.61**
		SPM	2907.21	2757.15	2841.96	2505.82
Galway	120	MPM	3923.94	3621.70	**3883.34**	**3181.09**
		SPM	3947.83	**3579.20**	4094.11	3211.92
	180	MPM	3896.84	**3483.41**	**3815.69**	**3120.39**
		SPM	3926.09	3511.99	3879.59	3150.98
	240	MPM	3910.85	3510.34	**3722.86**	**3105.25**
		SPM	3910.29	**3450.09**	3754.51	3133.96
Cork	120	MPM	11018.59	**10790.47**	10757.49	**9867.14**
		SPM	11232.35	10913.61	**10722.99**	9938.26
	180	MPM	10999.56	**10761.60**	10687.31	**9721.23**
		SPM	11864.68	10950.97	**10583.49**	9774.49
	240	MPM	10998.09	**10706.96**	10705.29	**9688.84**
		SPM	11442.79	10873.16	**10398.91**	9748.76

of the average schedule generated by the COP is improved compared to those reported in Table 4.5. However, it should also be noted that the performance of P_2 worsens regarding the performance of both the naïve and Ideal scenarios. Of note, P_2 using MPM generates schedules that are 2.47% better than $|\Gamma| = 0$ on average, while being 15.10% worse than the optimal solution in the Ideal scenario. This relative performance difference is noticeably worse than those reported in Table 4.5. However, it should also be noted that when only considering the Cork dataset, this average difference reduces to 9.81%. This would imply smaller datasets are more susceptible to mispredictions, which aligns with prior findings related to the SPM method. However, it would seem that SPM does not aid in the reduction of these mispredictions, with solutions generated using the P_2 scenario and SPM only 2.42% better than the naïve scenario, and 19.74% worse than the Ideal scenario for the Limerick and Galway datasets. This suggests that there is a trade-off between the performance of SPM and the magnitude of mispredictions made by the predictive model. Furthermore, some schedules generated using SPM with the P_2 scenario showcase worse results than the naïve scenario. This is representative of the events depicted in Figure 4.7 where a large amount of clean energy available in the morning is not used by SPM.

Table 4.11: Percentage of all clean energy consumed for experiments with 100% energy distribution, and Δt = 5 minutes for Inf-A.

City	$C_{\max}$	Method	$\lvert\Gamma\rvert = 0$	P_1	P_2	Ideal
Limerick	120	MPM	20.60	**20.95**	22.82	26.44
		SPM	13.92	18.61	**23.52**	**26.51**
	180	MPM	19.55	**22.58**	**23.78**	**26.67**
		SPM	18.50	21.56	23.54	26.39
	240	MPM	19.72	**23.11**	22.57	**26.49**
		SPM	19.05	20.40	**23.00**	26.26
Galway	120	MPM	21.94	26.43	27.65	36.10
		SPM	21.95	**33.07**	**28.20**	**36.65**
	180	MPM	23.64	28.72	29.50	37.09
		SPM	22.40	**30.97**	**29.73**	**37.57**
	240	MPM	22.93	28.20	28.47	**37.45**
		SPM	23.12	**31.32**	**28.55**	37.32
Cork	120	MPM	38.57	**47.29**	49.26	59.67
		SPM	37.72	44.60	**50.10**	**60.07**
	180	MPM	38.89	**47.38**	49.70	**63.13**
		SPM	30.00	39.10	**50.06**	61.57
	240	MPM	38.92	**47.51**	49.81	**62.60**
		SPM	34.52	45.51	**52.83**	61.59

Building on the impact of mispredictions, this table also shows that P_1 consistently outperforms schedules generated with the P_2 scenario for both the Limerick and Galway datasets. This may be due to several factors, such as the improved accuracy of future horizon models used by P_1. Furthermore, P_2 is more accurate at estimating extreme values of Upscaled Excess (i.e., values present in the 99$^{\text{th}}$ percentile of the training dataset). This proved useful for the BEB-CSP when power was distributed across multiple cities, as large amounts of clean energy would be more impactful than small amounts of energy. However, as P_1 is better at generalising for Upscaled Excess values above 0, these smaller values have more impact on the overall solution quality when power is not distributed. This is especially noticeable on the smaller datasets as they receive significantly lower amounts of clean energy compared to Cork in prior experiments (22.93% and 15.39% of available power for Galway and Limerick, respectively).

Table 4.11 shows the percentage of clean energy consumed for the experiments reported in Table 4.10. Here, the results show the percentage of clean energy consumed is lower in all reported configurations compared to those shown in Table 4.9. From this, it is apparent that, while the total amount of non-clean

Table 4.12: Non-clean energy used for experiments with 100% energy distribution, 1.2 kWh/km consumption rate, and Δt = 5 minutes for Inf-A.

City	$C_{\max}$	Method	$\lvert\Gamma\rvert = 0$	P_1	P_2	Ideal
Limerick	120	MPM	3642.54	**3396.76**	**3592.39**	**3136.80**
		SPM	3863.32	3488.26	3686.00	3147.97
	180	MPM	3609.79	**3388.74**	3550.22	**3095.33**
		SPM	3699.24	3448.74	**3529.89**	3113.97
	240	MPM	3637.78	**3399.48**	**3503.71**	**3095.33**
		SPM	3572.13	3431.93	3527.43	3123.94
Galway	120	MPM	4828.01	**4491.12**	**4774.87**	**4038.92**
		SPM	4860.12	4520.35	4952.76	4064.09
	180	MPM	4813.11	4359.65	**4729.95**	**3936.79**
		SPM	4798.74	**4356.49**	4804.99	3968.36
	240	MPM	4802.12	4357.46	4598.49	**3922.31**
		SPM	4762.40	**4295.18**	**4551.15**	3948.28
Cork	120	MPM	13547.94	**13374.19**	**13180.10**	**12426.64**
		SPM	13767.97	13630.90	13219.21	12494.68
	180	MPM	13547.77	**13298.32**	13168.56	**12332.66**
		SPM	14570.48	13621.38	**13126.22**	12391.19
	240	MPM	13551.98	**13243.35**	13100.36	**12305.34**
		SPM	14446.29	13501.71	**13009.57**	12390.68

energy used may be reduced, the BEB fleet with an energy consumption rate of 1 kWh/km reaches a limit on how much clean energy it can consume. This may be due to previously observed restrictions, such as battery capacity, maximum deviation time, and availability of charging stations, which all limit how much a BEB can charge. Additionally, SPM sometimes consumes more clean energy than MPM for the Ideal scenario. This is due to the objective function of the BEB-CSP minimising non-clean energy consumption and not the maximisation of clean energy consumption. While SPM may perform better in the consumption of clean energy, it uses more non-clean energy to do so, resulting in a worse solution compared to the Ideal scenario using MPM. Furthermore, the inability to consume additional clean energy enforces the previous observation regarding the improved performance of P_1 over P_2. As the BEB-CSP cannot consume all of the available energy, the accuracy of larger clean energy estimations becomes less relevant. This is because the energy needs of the bus fleet can be supplied by smaller occurrences of clean energy throughout the day.

Table 4.12 shows non-clean energy consumption for experiments using 100% clean energy distribution and an increased BEB energy consumption rate of 1.2 kWh/km. This increase simulates larger passenger loads and the usage of air conditioning or heating on a bus during various times of the year. Here, the difference between results for P_2 and the Ideal scenario decreases relative to Table 4.10, reducing from 15.10% to 12.69% on average for MPM. However, the naïve and P_2 scenarios are only subjected to minor changes.

Similar patterns regarding the performance of P_1 and P_2 also appear, with the former outperforming the latter on average for experiments using the Galway and Limerick datasets. For the Limerick dataset the difference between the two prediction models widens, with the average relative performance difference of 3.97% in Table 4.10 increasing to 4.06% in Table 4.12. Interestingly, the performance of P_1 worsens relative to P_2 when examining the Galway dataset. Where P_1 previously performed 9.41% better on average compared to P_2, the inclusion of a higher energy consumption rate decreases this to 6.63%. This supports the previous findings regarding the correlation between performance of P_1 and P_2, and the consumption of energy needed for the bus system. As the energy demand of the bus system increases a higher reliance is placed on accurate prediction of large amounts of clean energy, rather than small amounts of energy that might occur during the day. This can further be seen in the Cork dataset, as the performance of P_2 compared to P_1 increases by 0.55% from Table 4.10. It should also be noted that the optimal solution for the Ideal scenario is worse compared to experiment only using 100% energy distribution. This is due to the total energy consumption of the BEB fleet increasing by 20%.

Table 4.13 shows the percentage of clean energy consumed for experiments shown in Table 4.12. When results are compared to Table 4.11 it can be seen that the amount of clean energy consumed increases with energy consumption rate, with an average of 2.62% for the Ideal scenario using MPM across all cities. However, when examining results from the Cork City dataset, this increase is 3.89%. This would imply that, while having more free battery capacity to consume clean energy has an impact, the availability of charging stations plays a large role in how much clean energy a BEB fleet can consume.

As previously outlined, an additional charging station architecture, Inf-B., is also explored to determine the impact of a reduced amount of charging stations on each dataset. Table 4.14 shows the average non-clean energy used for experiments using Inf-B. and $\Delta t = 5$. Of the 3276 experiments reported in this table, only 3 were unable to find an optimal solution with an average optimality gap of 0.01%. As expected, decreasing the number of charging stations results in a decrease in solution quality. Intuitively, more BEBs will need to use

Table 4.13: Percentage of all clean energy consumed for experiments with 100% energy distribution, 1.2 kWh/km discharge rate, and Δt = 5 minutes for Inf-A.

City	$C_{\max}$	Method	$\|\Gamma\| = 0$	P_1	P_2	Ideal
Limerick	120	MPM	21.51	**25.33**	24.15	28.05
		SPM	16.00	21.16	**24.52**	**28.55**
	180	MPM	22.09	**24.67**	24.67	**28.61**
		SPM	19.64	22.86	**24.78**	28.11
	240	MPM	20.66	**24.59**	23.90	28.15
		SPM	21.84	22.99	**25.03**	**28.22**
Galway	120	MPM	24.80	28.51	**30.26**	38.17
		SPM	24.86	**31.95**	29.72	**38.73**
	180	MPM	24.94	32.60	30.70	39.52
		SPM	25.69	**33.18**	**31.40**	**40.14**
	240	MPM	25.22	28.84	31.07	39.57
		SPM	26.43	**35.08**	**32.73**	**40.12**
Cork	120	MPM	43.29	**50.83**	**55.16**	**63.80**
		SPM	43.80	46.78	54.09	63.32
	180	MPM	42.32	**51.22**	**54.16**	**66.37**
		SPM	31.64	41.05	53.57	64.68
	240	MPM	41.80	21.86	54.56	**66.91**
		SPM	32.96	**47.32**	**56.16**	64.82

Table 4.14: Non-clean energy usage for Inf-B. and Δt = 5 minutes.

City	$C_{\max}$	Method	$\|\Gamma\| = 0$	P_1	P_2	Ideal
Limerick	120	MPM	3198.54	**3212.08**	**3191.02**	**3074.53**
		SPM	3326.83	3293.18	3195.19	3103.65
	180	MPM	3227.25	**3228.13**	3242.09	**3074.53**
		SPM	3348.79	3300.08	**3211.08**	3098.38
	240	MPM	3208.49	**3253.01**	3256.39	**3074.54**
		SPM	3313.85	3295.20	**3195.82**	3110.10
Galway	120	MPM	4212.29	**4209.01**	**4126.34**	**3950.95**
		SPM	4504.42	4301.81	4151.33	3989.19
	180	MPM	4240.46	4326.23	4241.83	**3950.96**
		SPM	4534.92	**4310.27**	**4160.70**	3992.85
	240	MPM	4245.47	**4258.36**	4245.62	**3950.97**
		SPM	4565.24	4317.26	**4160.18**	4001.46
Cork	120	MPM	11206.95	**11135.06**	11059.49	**10548.16**
		SPM	11981.90	11405.17	**11022.32**	10590.90
	180	MPM	11333.46	**11166.52**	11134.27	**10500.62**
		SPM	12174.38	11406.43	**11008.52**	10560.03
	240	MPM	11334.62	**11340.22**	11067.66	**10493.91**
		SPM	11991.74	11463.81	**10954.23**	10559.79

Table 4.15: Relative performance difference for Table 4.14.

City	$C_{\max}$	Method	P_{best} vs $\|\Gamma\| = 0$	Ideal vs P_{best}
Limerick	120	MPM	0.23	3.78
		SPM	4.12	2.94
	180	MPM	-0.27	4.99
		SPM	4.28	3.63
	240	MPM	-1.36	5.80
		SPM	3.69	2.75
Galway	120	MPM	2.08	4.43
		SPM	8.50	4.06
	180	MPM	0.00	7.36
		SPM	8.99	4.20
	240	MPM	0.00	7.45
		SPM	9.73	3.96
Cork	120	MPM	1.33	4.84
		SPM	8.70	4.07
	180	MPM	1.78	6.03
		SPM	10.59	4.24
	240	MPM	2.41	5.46
		SPM	9.47	4.24

the charging station which may be shared by multiple routes at the same time. BEBs also may not be able to reach a charging station in time to charge using clean energy from Γ_k, meaning they may not have enough time to make full use of the clean energy available. Furthermore, this table shows that solution quality does not directly scale with the number of charging stations. For instance, empirical results show the Cork dataset with 53 charging stations (i.e., Inf-A.), Ideal predictions, MPM method, and $C_{\max} = 240$ kWh only performs 0.006% better than Inf-B. with 43 charging stations.

Table 4.15 shows the relative performance difference for the experiments featured in Table 4.14. Results for P_{best} are poor when compared to the naïve scenario when using MPM, with only an average performance difference of 0.69%. However, the reoccurring trend of the improved performance of scenarios using predictive models with SPM is also present here. With solutions using P_{best} and SPM having a 7.56% higher solution quality compared to those using no clean energy information. Furthermore, using MPM, the P_{best} scenario is 5.57% worse than the Ideal scenario on average, when using SPM, this changes to 3.78%.

4.5 Conclusion

In this chapter, a COP model was created to optimise the BEB-CSP to minimise the total amount of non-clean energy consumed by a BEB fleet. To this end, the two best-performing prediction configurations from Chapter 3 were used to provide estimations as to how much clean energy is available. Results for the BEB-CSP using these two predictive scenarios are compared to two different benchmarks. First was a naïve scenario with no knowledge of available clean energy, which acts as a baseline for comparison. Second was an Ideal scenario, with 100% perfectly accurate information about the availability of clean energy. Two methods for dealing with mispredictions for the prediction models were also examined, MPM and SPM. Where MPM utilises predictions up to 24-hours in the future to create a schedule, SPM only uses predictions from a 6-hour model to ensure the most accurate information is used.

Extensive experimental results with real data from the Irish national grid and three major cities' bus systems in Ireland show that the COP model solved with MILP generates schedules that consume up to 77.93% of the available clean energy when using an Ideal scenario. It was also found that schedules generated with a predictive model specialised at estimating large amounts of clean energy (P_2) reported an improvement of up to 3.85% against the naïve scenario when using MPM and an average discrepancy of 5.23% with regards to the Ideal scenario. When using SPM, these values change to 10.04% and 3.94% respectively. However, results for SPM in the Ideal scenario are worse compared to those generated using MPM. When comparing SPM using the P_2 scenario to the Ideal scenario using MPM the average discrepancy is 4.89%. This means that when using estimations from prediction models, SPM provides the best results.

Additional sensitivity analysis with additional availability of clean energy reports P_2 improves solutions by up to 11.04% against the naïve scenario and an average discrepancy of 13.15% with regards to the Ideal scenario when using SPM. However, the average relative performance difference between the predicted and Ideal scenarios falls to 12.69% using MPM, which insinuates both SPM and MPM have their strengths and weaknesses. Furthermore, a more generalised predictive model (P_1) outperforms P_2 when the overall amount of clean energy available is increased. This was due to the energy consumption of the BEB fleet not needing as high a percentage of the available clean energy to meet its daily energy demand.

In addition, the scheduler using P_2 uses up to 65.52% of the available clean energy for the baseline experiments and up to 56.16% for experiments with additional availability of clean energy and energy consumption. These results prove that the use of predictive models can improve the lifecycle emissions for BEBs in the BEB-CSP. This extensive evaluation of the use of predictive models to estimate an unknown variable for the BEB-CSP fills a gap in literature outlined in Chapter 2 Section 2.11 (Page 45) and considered Contribution 2 of this dissertation as outlined on Page 5.

In Chapter 5 the COP model evaluated in this chapter is expanded upon. In particular, the impact on service quality is examined in more detail by modelling public transport users' Value-of-Time (VOT). Furthermore, Carbon Dioxide Equivalent (CO_{2eq}) from all sources on the national grid is minimised and fuel costs are incorporated to showcase potential operational cost savings. Additional aspects for the BEB-CSP discussed in Chapter 2 Section 2.7 (Page 27) are also added, such as non-linear charging of BEB batteries, non-static energy consumption rates, and consideration of a mixed-fleet of BEBs and Internal Combustion Engine Bus (ICEB).

Chapter 5

Mixed-Fleet Battery Electric Bus Charging Schedule Problem for reducing fuel, emissions, and negative service quality costs

5.1 Introduction

In Chapter 2 Section 2.11 (Page 45) it was outlined that little work exists in the area of minimising lifecycle emissions and negative service impact for a Battery Electric Bus Charging Scheduling Problem (BEB-CSP). Furthermore, such works do not consider the impact of different weighting of Carbon Dioxide Equivalent (CO_{2eq}), or the effect of public transport passengers' Value-of-Time (VOT) when modifiers are applied. It was also noted that no work in the literature models a BEB-CSP to consider a mixed-fleet of Battery Electric Bus (BEB) and Internal Combustion Engine Bus (ICEB), non-linear charging, partial charging, charging station limitations, and weight-based energy consumption. Additionally, BEB-CSPs which minimise lifecycle emissions and negative service impact do not evaluate large datasets.

To address these gaps in the literature, Contribution 3 detailed in this Chapter extends the Constrained Optimisation Problem (COP) model developed in Chapter 4 to solve a BEB-CSP with non-linear charging, partial recharging, charger station limitations, and weight-based energy consumption for a mixed-fleet of BEBs and ICEBs. Mixed Integer Linear Programming (MILP) is used to implement this model, and the objective of this problem is to minimise the total daily operational costs. Where fuel costs of both BEB and ICEBs are considered, along with both tailpipe and lifecycle emissions of the fleet modelled as the Social Cost of Carbon (SCC) and public transport users VOT.

Datasets from the three largest cities in Ireland are used to empirically evaluate the performance of this model. In particular, the largest dataset, Dublin, contains over 1000 buses, which exceeds the number of buses examined in literature that focuses on emissions and service impact. Decomposition and simplification approaches are also explored to improve the scalability and determine the impact of particular aspects of the model. Furthermore, two weights for converting CO_{2eq} into SCC are investigated to determine the impact of more aggressive carbon-reducing policies.

This chapter is divided into the following sections. Section 5.2 defines the problem definition. Additionally, the datasets for electricity prices and CO_{2eq} of the Irish national grid are discussed. Data is then aggregated to represent an average weekday. Then, Section 5.3 formalises the extension of the COP model defined in Chapter 4 to solve the problem outlined for Contribution 3. This includes new constraints to model a piecewise linear approximation for non-linear charge time. Next, Section 5.4 outlines the experiment configuration and values used. The evaluation of the COP using three Irish cities, including specific details about negative service impact, emissions reduction, and the effects of higher weighting on CO_{2eq}. Finally, Section 5.5 closes the chapter and ties the work done to the thesis statement and Contribution 3.

5.2 Problem definition

Given a set of bus routes ($b \in S$) which contain several stops ($i \in S_b$) in a sequence, which may repeat. Each stop has an originally scheduled arrival time (τ_{bi}), an assigned new arrival time (t_{bi}) and a charge time (ct_{bi}) to designate if a bus route is serviced by an ICEB or BEB (eb_b). The assignment of bus types is based upon the objective function to minimise total daily operational costs, which is determined by fuel price (f_b), SCC (ssc_b), and change in user VOT (Δvot_{bi}) for both ICEB and BEBs.

Each BEB has a battery capacity at each stop (c_{bi}) which falls between a maximum and minimum battery capacity ($C_{\max}$ and $C_{\min}$ respectively) represented in Megawatt hour (MWh). The BEB's battery loses energy as it travels between two bus stops (D_{ij}), which increases with higher passenger loads (L_{bi}). If a stop is considered a charging station (X_i), the BEB may recharge its battery (x_{bi}).

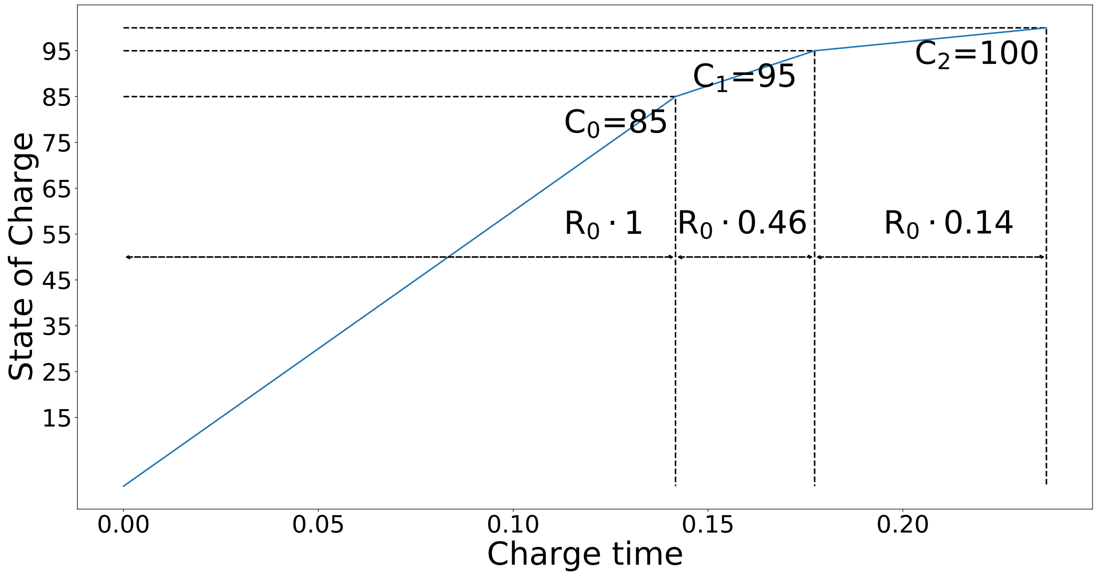


Figure 5.1: Piecewise linear approximation of non-linear charge time

If a BEB recharges its battery, the amount of energy gained (e_{bi}) is based on a piecewise linear function of non-linear charging with three linear sections ($r \in |\mathcal{R}|$) depicted in Figure 5.1. The first section represents the Constant Current (CC) phase of the battery charging process, where the battery is charged linearly at the fully available charge rate ($\mathcal{R}_0$) until the BEB's State of Charge (SOC) reaches a cut-off threshold (C_0). At this point, the Constant Voltage (CV) phase of battery recharging begins, and a lower charging rate is used (R_1) until the final cut-off threshold value is reached (C_1). Then the final section of the linear approximation begins where the lowest charge rate (R_2) is used. This process continues until the SOC of the BEB is 100, or the charge time (ct_{bi}) has elapsed. Additional details about non-linear charging are discussed in Chapter 2 Section 2.7.4 (Page 32). Furthermore, a minimum charge amount is enforced (γ) to ensure micro charges do not occur.

The new arrival time of a BEB at a given stop depends on the arrival and charge time at the previous stop and the time taken to travel between the two stops (T_{ij}). The maximum difference between the original arrival time and the new arrival time is limited (Δt). To further reduce the impact on service quality, passengers' VOT is considered as a value of currency, as well as modifiers which represent users' opinion of late arrivals (ζ_a^+), late departures (ζ_d^-), early arrivals (ζ_a^-), and early departures (ζ_d^+). The change in time spent on public transport, amount of passengers on board, and modifiers are used to calculate the change in users' VOT (Δvot_{bi}).

Both the CO_{2eq} and the price of electricity on the national grid fluctuate throughout the day. To optimise a BEB-CSP to reduce carbon emissions and fuel costs, a modification is made to the Time-of-Use (TOU) windows from Chapter 4. An operational day is split into TOU windows ($k \in |\Gamma|$), each with an associated start time (P_k), end time (Q_k), electricity price (F_k), and SCC (V_k). Where SCC is the CO_{2eq} multiplied by an weight to convert it into a financial value, as detailed in Chapter 2 Section 2.7.7 (Page 38). The total SCC ($scce_{bi}$) and fuel cost (fe_{bi}) for a BEB at each stop are determined by the amount of electricity consumed at that stop for each TOU window across all sections of the piecewise linear function (e_{bikr}). For ICEBs, the SCC (DV_b) and fuel cost (DF_b) of diesel are considered to be static throughout the day based on the distance travelled by that bus.

The considerations for the unary nature of charging stations are the same as stated in Chapter 4 Section 4.2 (Page 85). While it is also assumed BEBs do not begin the operational day with a full battery capacity, a stochastic approach is taken to the starting value rather than a deterministic one. In Chapter 4 Section 4.4.2 (Page 101), it was found that reducing the number of charging stations only *slightly* diminishes solution quality. This chapter assumes charging stations are placed at terminal bus stops instead of every X Kilometre (km) [Laj18]. While this diminishes solution quality, it would lower the capital investment for charging infrastructure, and also ensure adequate space to idle while charging.

5.2.1 Electricity price and carbon intensity dataset

To optimise the fuel cost and SCC for the BEB-CSP with a mixed-fleet of ICEB and BEBs, knowledge of electricity prices and carbon emissions is required. In Chapter 4 the power of using predictive models was proven concerning the performance of the BEB-CSP when using historical data as a part of Contribution 2. As the power of predictive models has already been examined, this chapter and the contribution made within focus on using historically accurate information to determine the full extent of reducing daily operational costs.

To this end, data was gathered from February 2023 to January 2024 for a full 12 months of data. Electricity price information was gathered from the Single Energy Market Operator (SEMO), which reports the wholesale price of electricity [1]. The temporal resolution of this time series dataset is one hour. To bring this dataset in line with other used, values were interpolated for every 15-minute segment between two data points.

Data related to CO_{2eq} of the national grid is retrieved from Smart Grid Dashboard[2], which defines the average grams of Carbon Dioxide per Kilowatt hour (gCO_2/kWh) in 15-minute intervals. As outlined in Chapter 4 Section 2.9 (Page 43), SCC is the cost of burning one Tonne of Carbon Dioxide Equivalent (tCO_{2eq}). There exist several different values for SCC. The two values explored for this contribution are the Paris Agreement extended scenario 2025 [BN24], and the Irish Government's appraisal guidelines for the Shadow Price of Carbon (SPC) for 2050 [Dep23b][3]. Therefore, two datasets were used for SCC, the first is the CO_{2eq} for every 15-minute interval multiplied by the SCC value set out by the Paris Agreement extended scenario 2025 (€88.675 per tCO_{2eq}). While the second uses the CO_{2eq} multiplied by the Irish government's value for Shadow Price of Carbon (SPC) (€252.235 per tCO_{2eq}). Here, both values have been adjusted for inflation.

The resulting datasets are aggregated by day, then aggregated by weekend and weekday. The final result is a dataset containing the average electricity price and SCC for weekdays and weekends. The dataset representing an average weekday was selected as it contains the majority of observations. Additionally, weekends traditionally feature lower system demand and electricity prices, thus artificially lowering overall operational costs. Furthermore, as will be later discussed, the VOT used for public transport users is based on a work commuting value, thus motivating the use of weekdays over weekends.

Figure 5.2 shows the cost of one MWh of electricity throughout an average weekday, as well as the carbon intensity reported as SCC using both the 2025 and 2050 scenarios. As can be seen, the electricity price peaks at 77.49% higher than the pit prices. This represents an approximate difference of 68.6264 euros per MWh consumed, implying that significant savings can be made for fuel prices if these peaks are avoided. Fluctuations are also observed for both SCC scenarios. However, it is not as prevalent. While there is some overlap with

[1]Electricity information as gathered from https://reports.sem-o.com/documents/

[2]National grid dataset available from https://www.smartgriddashboard.com/#roi/co2

[3]The value for the SPC was retrieved in May 2024, the document has since been replaced with a new value. [Dep23b] provides a link to the archived version of the report, while [Dep24c] provides a link to the new version.

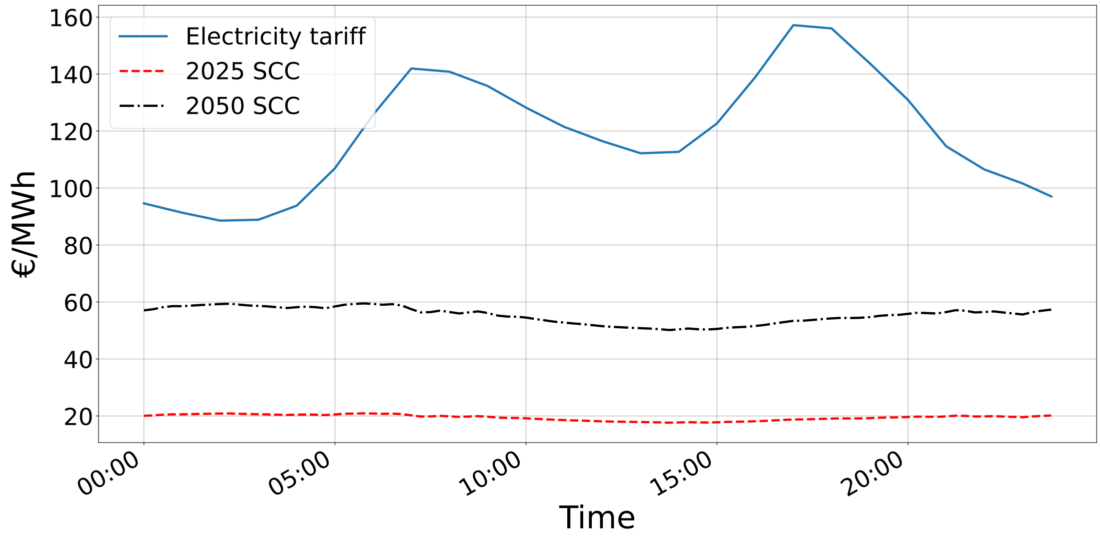


Figure 5.2: Cost of electricity and SCC throughout an average weekday in €/MWh

electricity prices, differences are present. For example, the peak carbon intensity occurs at 5:45, whereas the peak electricity price occurs at 17:00.

5.2.2 Dublin bus dataset

In Chapter 4 Section 4.2.3 (Page 90), the process of generating the Cork, Galway, and Limerick datasets was described, summarising the unpublished report they were developed in [LA20]. To evaluate the performance of the new BEB-CSP described in this chapter, both the Cork and Limerick city datasets are used. One change made is the inclusion of the bus route 350 Ennis (Bus Station) - Galway (Bus Station) from the Limerick dataset, which was originally omitted due to being an intercity bus route. The use of a mixed-fleet of BEBs and ICEBs motivates this route's inclusion in this chapter, as well as the placement of charging stations solely at terminal stops. In addition to these datasets, a third, much larger dataset is also examined, Dublin City.

Instead of using timetables from the bus operator's website, an online dataset was used to generate the Dublin Bus dataset[4]. This dataset contains over 44 million Global Positioning System (GPS) coordinates for Dublin Bus vehicles for January 2013. The same authors who previously worked on the Cork, Galway, and Limerick datasets [LA20] adapted this data into a dataset. This process took the actual GPS coordinates of the Dublin Bus dataset and converted it into

[4]Dublin Bus dataset can be found at https://data.gov.ie/dataset/dublin-bus-gps-sample-data-from-dublin-city-council-insight-project

the same format as the previously used bus datasets. This results in a dataset with 1169 buses and 4279 unique bus stops.

One point of note is that the Dublin Bus dataset does not contain strict timetables like the previously discussed datasets for Cork, Galway, and Limerick. Due to this, it is difficult to identify which stops are terminal stops for a bus. To resolve this, it is assumed that any stop where the bus is expected to idle for 30 minutes is a terminal stop. Such that if the scheduled arrival time of a bus at its next stop minus the scheduled arrival time of the current stop is 30 minutes more than the time it takes to travel the distance between stops.

5.3 Mathematical optimisation model

The model developed in Chapter 4 is extended to solve the previously outlined problem. While the majority of constraints have been modified, some are duplicated for this model. These include constraints (4.1) which enforces minimum and maximum battery capacities, (4.7) which sets the arrival time of BEBs, and the non-overlapping charge time constraints (4.10)–(4.15). For the sake of brevity, these constraints are omitted from this section.

Tables 5.1 and 5.2 show the updated constants and variables used in the model. Similar to the constraints, some constants and variables are unchanged from Chapter 4. These have been omitted from these tables. These include the constants $|B|$, b, $|S_b|$, s_i, X_i, T_{ij}, τ_{bi}, M and the variables x_{bi}, Z_{bdij}, z^{BD}_{bdi}, and z^{DB}_{dbi}.

5.3.1 Constraints

$$\forall_{b \in B} \forall_{s_i \in S_b} :$$

$$ct_{bi} \leq x_{bi} \tag{5.1}$$

$$\gamma \cdot x_{bi} \leq e_{bi} \leq x_{bi} \tag{5.2}$$

$$x_{bi} \leq X_i \cdot eb_b \tag{5.3}$$

Constraint (5.1) sets an upper bound for ct_{bi}, which is limited to one hour if a charge event occurs. Constraint (5.2) enforces a lower limit on the e_{bi} to γ if a charge event occurs. This protects the longevity of the batteries by avoiding overheating and regular tiny charges. While these constraints are present in

Table 5.1: Chapter 5 constants

Constant	Description
C_{min}	Minimum allowed capacity of a bus battery in MWh
C_{max}	Maximum allowed capacity of a bus battery in MWh
$D_{b_{ij}}$	Energy required by bus b to travel from stop i to j in MWh
γ	Represents the minimum allowed energy gain in MWh
$\lvert\mathcal{R}\rvert$	Set of values denoting the charging rate used by each section of the piecewise linear approximation for non-linear sharing in MWh.
$\lvert\Gamma\rvert$	Sequence of ordered TOU windows
Γ_k	k^{th} TOU window
P_k	Start time of TOU window Γ_k
Q_k	End time of TOU window Γ_k
V_k	SCC for using 1 MWh of electricity during TOU window Γ_k
F_k	Electricity cost for using 1 MWh of electricity during TOU window Γ_k
ζ_a^+	Modifier applied to the difference in VOT if a bus arrives at a stop late
ζ_a^-	Modifier applied to the difference in VOT if a bus arrives at a stop early
ζ_d^-	Modifier applied to the difference in VOT if a bus departs at a stop early
ζ_d^+	Modifier applied to the difference in VOT if a bus departs at a stop late
$\lvert C\rvert$	Set of values in MWh denoting the bounds for the section of the piecewise linear approximation for non-linear charging
DV_b	SCC of bus b if it an ICEB.
DF_b	Fuel cost of bus b if it an ICEB.
L_{bi}	Passenger load of bus b when departing stop i.
VOT	VOT constant for one passenger spending one hour on public transport
VOT_{bi}	Original VOT for bus b departing stop i.
Δt_{max}	Maximum allowed deviation time between the arrival time at any stop and the original schedule
F'	Average electricity cost for using 1 MWh
V'	Average SCC for using 1 MWh

Table 5.2: Chapter 5 variables

Variable	Domain	Description
t_{bi}	$[\tau_{bi}-\Delta t_{\max}, \tau_{bi}+\Delta t_{\max}]$	Arrival time of bus b at stop i
c_{bi}	$[C_{\min}, C_{\max}]$	Battery capacity of bus b at stop i
ct_{bi}	$[0, 1]$	Time bus b spent charging at stop i
e_{bi}	$[0, C_{\max}-C_{\min}]$	Energy gained by bus b at stop i
ΔVOT_{bi}	$[0, VOT_{bi} + \Delta t_{\max}\cdot VOT \cdot \zeta_d^+ \cdot L_{bi}]$	The financial cost of bus b deviating from the original schedule at stop i.
scc_b	$\{scce_{bi}, DV_b\}$	The total SCC of bus b.
f_b	$\{ef_b, DF_b\}$	The total fuel cost of bus b.
eb_b	$\{0, 1\}$	Binary variable assigned 1 if bus b is a BEB
ev_b	$\sum_{i=0}^{\lvert S_b \rvert} scce_{bi}$	The SCC of bus b if it is a BEB.
ef_b	$\sum_{i=0}^{\lvert S_b \rvert} fe_{bi}$	The fuel cost of bus b if it is a BEB.
e_{bikr}	$[0, (Q_k - P_k) \cdot \mathcal{R}_r]$	The amount of energy used by bus b at stop i during TOU window k in piecewise linear approximation section r
ct_{bikr}	$[0, Q_k - P_k]$	The amount of charge time spent by bus b at stop i during TOU window k in piecewise linear approximation section r
o_{bikr}	$\{0, 1\}$	Binary variable denoting an overlap between bus b time at stop i and both TOU window k and piecewise linear approximation section r
fe_{bi}	$[0, (C_{\max}-C_{\min}) \cdot max(F)]$	The electricity cost of bus b at stop i if it is a BEB
$scce_{bi}$	$[0, (C_{\max}-C_{\min}) \cdot max(V)]$	The SCC of the electricity bus b consumes at stop i if it is a BEB
λ_{abi}^+	$[0, \Delta t_{\max}\cdot VOT \cdot \zeta_a^+]$	Amount of VOT associated with late arrival for bus b at stop i
λ_{abi}^-	$[0, \Delta t_{\max}\cdot VOT \cdot \zeta_a^-]$	Amount of VOT associated with early arrival for bus b at stop i
λ_{dbi}^+	$[0, \Delta t_{\max}\cdot VOT \cdot \zeta_d^+]$	Amount of VOT associated with early departure for bus b at stop i
λ_{dbi}^-	$[0, \Delta t_{\max}\cdot VOT \cdot \zeta_d^-]$	Amount of VOT associated with late departure for bus b at stop i

Chapter 4, changes have been made, such as the removal of β, thus increasing the upper bound for ct_{bi}, and changing γ from a time-based parameter to energy based one. These changes are due to the consideration of non-linear charging time. Constraint (5.3) enforces that bus b can charge (x_{bi}) only if i is a charging station and if b is a BEB.

$$\forall_{b\in B}\forall_{s_j\in S_b\setminus s_0, i=j-1}:$$

$$c_{bj} \leq c_{bi} + e_{bi} - D_{b_{ij}} \cdot eb_b \tag{5.4}$$

$$\lambda^{+}_{abj} \geq (t_{bj} - \tau_{bj}) \cdot (vot \cdot \zeta^{+}_{a}) \tag{5.5}$$

$$\lambda^{+}_{dbj} \geq (\tau_{bi} - t_{bi} - ct_{bi}) \cdot (vot \cdot \zeta^{+}_{d}) \tag{5.6}$$

$$\lambda^{-}_{abj} \leq (\tau_{bj} - t_{bj}) \cdot (vot \cdot \zeta^{-}_{a}) \tag{5.7}$$

$$\lambda^{-}_{dbj} \leq (t_{bi} + ct_{bi} - \tau_{bi}) \cdot (vot \cdot \zeta^{-}_{d}) \tag{5.8}$$

$$\Delta vot_{bi} \geq (vot_{bi} + \lambda^{+}_{abi} + \lambda^{+}_{dbi} - \lambda^{-}_{abi} - \lambda^{-}_{dbi}) \cdot L_{bi} \tag{5.9}$$

$$\Delta vot_{bi} \leq vot_{bi} + ((\Delta t_{\max} \cdot vot) \cdot \zeta^{-}_{d}) \cdot L_{bi} \tag{5.10}$$

Constraint (5.4) enforces the energy expenditure of BEBs. Constraints (5.5)–(5.8) defines the deviations in VOT for passengers representing late arrival, early departure, early arrival, and late departure respectively. Constraint (5.9) applies the deviations to the total VOT deviation for bus b at stop i considering the number of passengers on the bus. Here, late arrival and early departure are considered to increase VOT, while early arrival and late departure decrease it. This is because the former represents additional time spent on public transport, while the latter results in the opposite. Constraint (5.10) defines the upper limit to the total VOT deviation for bus b at stop i, which depends on the maximum allowed deviation time per stop Δt_{bi}.

$$\forall_{b \in B} \forall_{s_i \in S_b} \forall_{k \in \Gamma} \forall_{r \in \mathcal{R}}:$$

$$e_{bikr} \leq C_r - c_{bi} - \sum_{h=0}^{k} \sum_{l=0}^{r-1} e_{bihl} - \sum_{h=0}^{k-1} \sum_{l=0}^{r} e_{bihl} + C_r \cdot (1 - o_{bikr}) \tag{5.11}$$

$$C_{r-1} \cdot o_{bikr} \leq c_{bi} + \sum_{h=0}^{k} \sum_{l=0}^{r-1} e_{bihl} + \sum_{h=0}^{k-1} \sum_{l=0}^{r} e_{bihl} \tag{5.12}$$

$$e_{bikr} \leq \mathcal{R}_r \cdot ct_{bikr} \tag{5.13}$$

$$t_{bi} + \sum_{h=0}^{k} \sum_{l=0}^{r} ct_{bihl} - M \cdot (1 - o_{bikr}) \leq Q_k \tag{5.14}$$

$$P_k \cdot o_{bikr} \leq t_{bi} + \sum_{h=0}^{k} \sum_{l=0}^{r-1} ct_{bihl} + \sum_{h=0}^{k-1} \sum_{l=0}^{r} ct_{bihl} \tag{5.15}$$

$$ct_{bikr} \leq (Q_k - P_k) \cdot o_{bikr} \tag{5.16}$$

$$o_{bikr} + o_{bik-1r+1} \leq 1 \tag{5.17}$$

The TOU windows from Chapter 4 Section 4.3.2 (Page 96) are adapted to be used with the piecewise linear approximation for non-linear charging. Constraints (5.11)–(5.17) defines the TOU constraints for electricity price, SCC, and the piecewise linear charging function. Constraint (5.11) defines the maximum energy bus b can gain at stop i during section r, considering any charging from elapsed TOU windows and previous piecewise linear function sections. Here, C_r is used in place of Big-M as it is a smaller value that proves suitable for the constraint. Constraint (5.12) ensures the viability of the piecewise linear function section r for bus b at stop i. Constraint (5.13) associates the energy gained during TOU window k at charge rate r with a linear sum of $\mathcal{R}_r$ and ct_{bikr}. Constraint (5.14) enforces the maximum allowed charge time during section r before the end of TOU window k. Likewise, Constraint (5.15) denotes that charge time can only occur in k if charging happens after the start of the window. Constraint (5.16) ensures that charging only occurs if it overlaps with the time bus b spends at stop i, TOU window k, and charge section r. Finally, Constraint (5.17) denotes that there can be no overlap with TOU window k and charge section r if a higher charge rate section was previously used to charge bus b at stop i.

$$\forall_{b\in B}\forall_{s_i\in S_b}$$

$$ct_{bi} \geq \sum_{k=0}^{|\Gamma|}\sum_{r=0}^{|\mathcal{R}|} ct_{bikr} \tag{5.18}$$

$$e_{bi} \geq \sum_{k=0}^{|\Gamma|}\sum_{r=0}^{|\mathcal{R}|} e_{bikr} \tag{5.19}$$

$$scce_{bi} \geq \sum_{k=0}^{|\Gamma|}\sum_{r=0}^{|\mathcal{R}|} e_{bikr}\cdot V_k \tag{5.20}$$

$$fe_{bi} \geq \sum_{k=0}^{|\Gamma|}\sum_{r=0}^{|\mathcal{R}|} e_{bikr}\cdot F_k \tag{5.21}$$

$$\sum_{i=0}^{|S_b|} scce_{bi} \leq ev_b \tag{5.22}$$

$$\sum_{i=0}^{|S_b|} fe_{bi} \leq ef_b \tag{5.23}$$

Constraints (5.18) and (5.19) associate the charge time and energy gained via the TOU constraints with the total charge time and energy gained by bus b at stop i. Constraints (5.20) and (5.21) set the values for SCC and fuel cost based on the time electricity was consumed. While Constraints (5.22) and (5.23) calculate the total amount of SCC and fuel cost for bus b if it is a BEB.

$$\forall_{b\in B}$$

$$ev_b + DV_b\cdot(1-eb_b) \leq scc_b \tag{5.24}$$

$$ef_b + DF_b\cdot(1-eb_b) \leq f_b \tag{5.25}$$

Constraints (5.24) and (5.25) then decide if the BEB values for SCC and fuel cost are used, or the SCC and fuel cost associated with ICEB, depending on the type of bus assigned to bus b's route.

5.3.2 Objective function

$$\min: \sum_{b=0}^{|B|}\sum_{i=0}^{|S_b|}\Delta vot_{bi} + \sum_{b=0}^{|B|} scc_b + \sum_{b=0}^{|B|} f_b \tag{5.26}$$

The objective function (5.26) aims to simultaneously minimise the total daily operational costs of the fleet subject to the SCC, fuel cost, and difference in passenger VOT, based on the information provided for carbon intensity, fuel price, and passenger loads.

5.3.3 Decomposition and simplifications

$$\forall_{b \in B} \forall_{s_i \in S_b} :$$

$$e_{bi} \leq ct_{bi} \cdot \mathcal{R}_0 \tag{5.27}$$

$$scce_{bi} \geq e_{bi} \cdot V' \tag{5.28}$$

$$fe_{bi} \geq e_{bi} \cdot F' \tag{5.29}$$

To deal with large datasets, decomposition and alternative approaches are explored for non-linear charging and TOU windows. For the decomposition approach, TOU windows are not initially considered. Constraints (5.11)–(5.21) are instead replaced with constraints (5.27)–(5.29) which enforce a linear charge rate and use a daily average value for SCC and electricity cost represented by V' and F' respectively. Charging is limited to the first section of the piecewise linear charging function, that is, for the first part of the decomposition approach, C_{max} is set to C_0. The solution is translated into a warming solution for the full problem with TOU constraints and non-linear charging by setting key variables such as eb_b, t_{bi}, c_{bi}, ct_{bi}, and e_{bi}.

A simplification is also explored where TOU constraints are considered, but charging is limited to the first section of the piecewise linear function. To this end, Constraints (5.11) and (5.12) are removed, $|\mathcal{R}|$ is 1 (i.e., only using $\mathcal{R}_0$), and C_{max} is set to C_0. Theoretically, this approach will have a worse global minima than the previously outlined methods but should simplify the problem. A decomposition approach for this simplification is also examined, where the first step follows the same definition as the decomposition for the full problem. However, instead of translating the resulting solution into the full problem with TOU and piecewise linear charging, the solution is used as a warming solution for the simplification outlined above.

5.4 Experiments and Evaluation

5.4.1 Experiment configuration

The COP model outlined in Section 5.3 is evaluated using datasets from three Irish cities. The Limerick and Cork datasets previously outlined in Chapter 4 Section 4.2.3 (Page 90) are evaluated, along with a dataset for Dublin, which is detailed in Section 5.2.2. As previously outlined, this chapter assumes charging stations are placed at terminal stops instead of every X km. This results in 13, 28, and 410 charging stations for Limerick, Cork, and Dublin, respectively. Data for electricity price and SCC is outlined in section 5.2.1.

The starting capacities for BEBs are assigned a random value between 40%–60% C_{max}, which itself is assigned a value of 0.454 MWh. A $\mathcal{R}_0$ value of 0.15 MWh is used. Both C_{max} and $\mathcal{R}_0$ values are consistent with the specifications of the BEB's employed by Dublin Bus. C_{min} is assigned the value 0.0454 MWh (i.e., 10% C_{max}), an average speed of 20 Kilometres per hour (km/h) is used, and γ is assigned 0.01 MWh. Note that while Chapter 4 and common literature use Kilowatt hour (kWh) to represent energy consumed by BEBs, this Chapter uses MWh instead. This is due to the inclusion of the Dublin dataset, which was found to cause rounding errors during some experiments, due to Big-M trickle-flow as discussed in Chapter 2 Section 2.5.2 (Page 22). To solve this, MWh was used instead of kWh to bring scaling of time and energy consumed closer together. Furthermore, the reduced average speed of 20 km/h compared to 35 km/h used in the previous chapter is representative of the more urban development of Dublin city.

During peak operation hours, it is assumed the passenger load of each bus is 88, which is the maximum passenger capacity for BEBs employed by Dublin Bus, while off-peak hours feature a passenger load of 23. This results in an average passenger load of 33.83 over a 24-hour period, which is consistent with the average passenger load for Dublin Bus in 2018 [ORG+22]. When not operating during peak hours, the energy expenditure of BEBs is assumed to be 0.0016 Megawatt hours per Kilometre (MWh/km). Outside these times, energy consumption increases to 0.0032 MWh/km. This is due to the increased total weight from the additional passenger load, which was shown in Chapter 2 Section 2.7.3 (Page 31) to be a major contributing factor to the operational range of BEBs. It is assumed that peak operating hours are between 7–10 and 17–19.

To further deal with Big-M trickle-flow, a value of 3 was assigned to Big-M instead of 25 used in Chapter 4. This was possible due to the tightening of bounds related to arrival times of BEBs, as seen in the domain for variable t_{bi} in Tables 4.3 and 5.2. The maximum allowed deviation time per stop (Δt_{max}) is set to $0.066\dot{6}$ (4 minutes) to emulate tighter restrictions on timetable deviation compared to Chapter 4. VOT is set €10.87 per hour [Dep24c][5], and the values for ζ_a^+, ζ_a^-, ζ_d^+, and ζ_d^- are set to 1.97, 0.92, 0.33, 0.43 respectively [WCdJ16]. The cost of diesel fuel is €0.3393274 per Kilometre (km), and emissions for ICEBs are set to 1360 Grams of Carbon Dioxide Equivalent per Kilometre (gCO_{2eq}/km) [Dep19][6]. As a result, the SCC for ICEBs is 0.120598 and 0.3430396 €/km for the 2025 and 2050 scenarios, respectively. For the decomposition approaches, F' is set to 119.6795 €/MWh, while V' is 21.0858 €/MWh for the 2025 scenario and 59.9783 for the 2050 scenario.

The COP model is implemented in C++ and uses CPLEX version 12.10. All experiments are executed five times with a different random seed. Experiments were conducted on a server with 62GB of Random Access Memory (RAM), and a 2.5 Gigahertz (GHz) Intel Xenon W-2175 Central Processing Unit (CPU). A timeout of 1200 seconds is used for the Limerick and Cork datasets, while a timeout of 3600 is used for Dublin. When executing the decomposition approaches, each step is limited to half the total timeout.

5.4.2 Evaluation

Table 5.3 shows the mean solution time in seconds, the mean solution quality, the score Standard Deviation (STD), and the daily percentage saving compared to a fully ICEB fleet across all five random seeds for all Limerick and Cork experiments. Best-performing results are highlighted in bold, and the carbon scenarios and approaches are abbreviated into config. column (i.e., carbon scenario 2025 full problem is 2025 F, carbon scenario 2050 simplification decomposition is 2050 SD, etc.). All experiments for Limerick and Cork find the optimal solution within the runtime. However, slight differences in optimal solutions are reported. This is due to CPLEX's relative Mixed Integer Programming (MIP) gap tolerance, where solutions with a gap $\leq$ 1x10^{-4} are considered optimal. Optimal solutions report relative gap values between approximately 1x10^{-5} and 9x10^{-5}, which causes slight variations in the final score of the solution. However, this variation is tiny in proportion to the scores reported. The standard deviation

[5] Commuting non-work price adjusted for inflation
[6] It is assumed all buses use a euro VI engine, which is the highest euro emissions standard currently used

Table 5.3: The mean search time, mean solution quality, the score STD, and percentage saving compared to a fully ICEB fleet for the Limerick and Cork dataset for both carbon scenarios, and all approaches

City	Config.	Time (s)	Score (€)	STD(€)	∆ICEB(%)
Limerick	2025 F	6.38	150980.6811	**571.8977**	5.3997
	2025 D	4.78	150943.1586	586.3493	5.4232
	2025 S	4.73	150943.2969	584.0584	5.4231
	2025 SD	**4.09**	**150941.9809**	585.8825	**5.4239**
	2050 F	6.03	151729.0736	**616.5176**	5.6566
	2050 D	4.87	151689.3326	630.2135	5.6813
	2050 S	4.86	151690.3671	627.5853	5.6806
	2050 SD	**4.05**	**151688.7073**	630.3015	**5.6817**
Cork	2025 F	37.65	452864.3374	1082.7145	5.0624
	2025 D	42.41	452861.7713	1073.3037	5.0629
	2025 S	**15.55**	**452848.7509**	**1071.2811**	**5.0657**
	2025 SD	23.42	452857.0186	1072.8210	5.0639
	2050 F	40.26	454970.1445	1278.3058	5.2931
	2050 D	40.95	454880.3992	1148.4161	5.3118
	2050 S	**15.40**	**454874.2333**	1151.0025	**5.3131**
	2050 SD	22.10	454887.9297	**1146.4702**	5.3102

for solution scores reports limited variance, further reinforcing this.

Table 5.3 shows that the simplification and simplification decomposition approaches often outperform other approaches in search time and resulting solution score. For example, the Cork dataset carbon scenario 2050 simplification approach is 161.43% faster and generates a solution of 0.02% improved quality compared to the full problem approach. The improved score performance is noteworthy, as the global optima should be worse due to charging being restricted to the CC phase. This is due to the instance size and the selected constant's values. Optimal solutions using these datasets do not require the capacity of a BEB to exceed C_0. As a result, the global optima for the full problem and the simplification approach are the same. This is due to several factors, such as the battery capacity used, the number of charging stations available, and the maximum allowed deviation time per stop, all of which limit the total amount of time a BEB can charge.

When comparing the generated mixed-fleet schedules to a fleet of fully ICEB significant savings can be observed for the Cork dataset, with up to €25523.90 saved per day due to decreases in fuel price and SCC, as well as VOT reductions. VOT may be subjective to the customer base and passenger load information. Excluding VOT, total savings of up to €4909.10 and €6946.42 per day are observed for carbon scenarios 2025 and 2050, respectively. This highlights the importance of VOT and considering passenger wait times when generating schedules, as it has a high level of impact. The differences between carbon scenarios 2025 and 2050 are also highlighted by Table 5.3 with the solution scores several thousand lower in the former compared to the latter. However, the use of carbon scenario 2050 accomplishes an overall reduction in emissions. For the 2025 carbon scenario, emission reductions of up to 8.31 tCO_{2eq} are observed for the Cork dataset per day, when considering the 2050 scenario emissions reduce further by 0.0025 tCO_{2eq}.

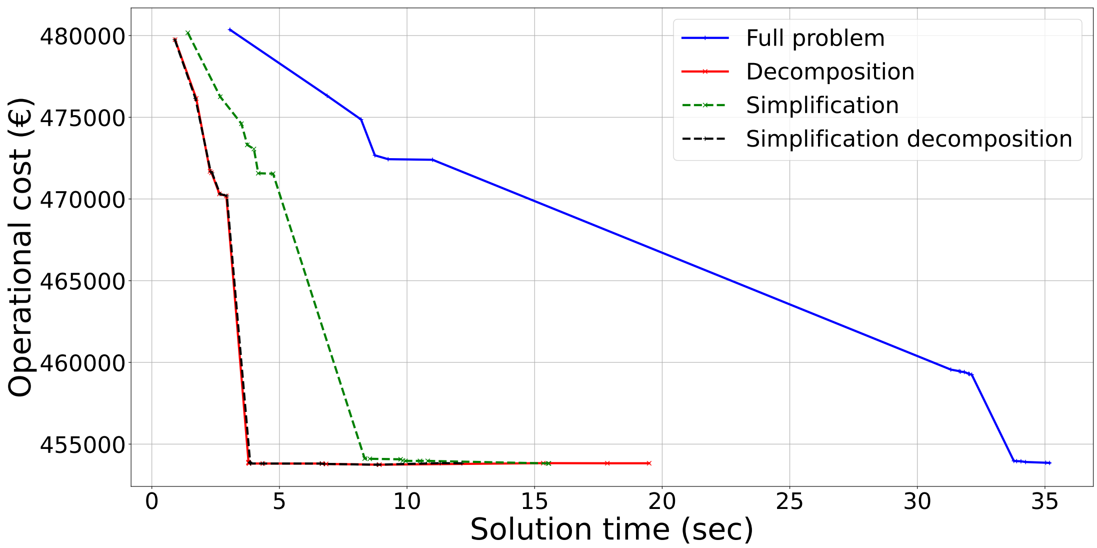


Figure 5.3: Solution quality (€) versus solution time (seconds) for Cork dataset for the four examined approaches with carbon scenario 2025

Figures 5.3 and 5.4 show the solution quality in daily operational cost versus solve time in seconds for all approaches for the Cork and Limerick datasets respectively using the 2025 SCC scenarios for one of the five executions which were randomly selected. Figures 5.3 and 5.4 show both decomposition approaches quickly solve their problem, finding optimal solutions before their respective non-decomposition approaches. It is observed that the performance of both decomposition approaches progresses along the same trajectory during the initial stages of the search process. This is because the first step of the decomposition approach is the same, with variance only occurring during the second stage.

Figures 5.5 and 5.6 show the solution quality in operational costs versus solve time for the Cork and Limerick datasets for the 2050 carbon scenario. Here, both decomposition approaches find optimal solutions in approximately 20 sec-

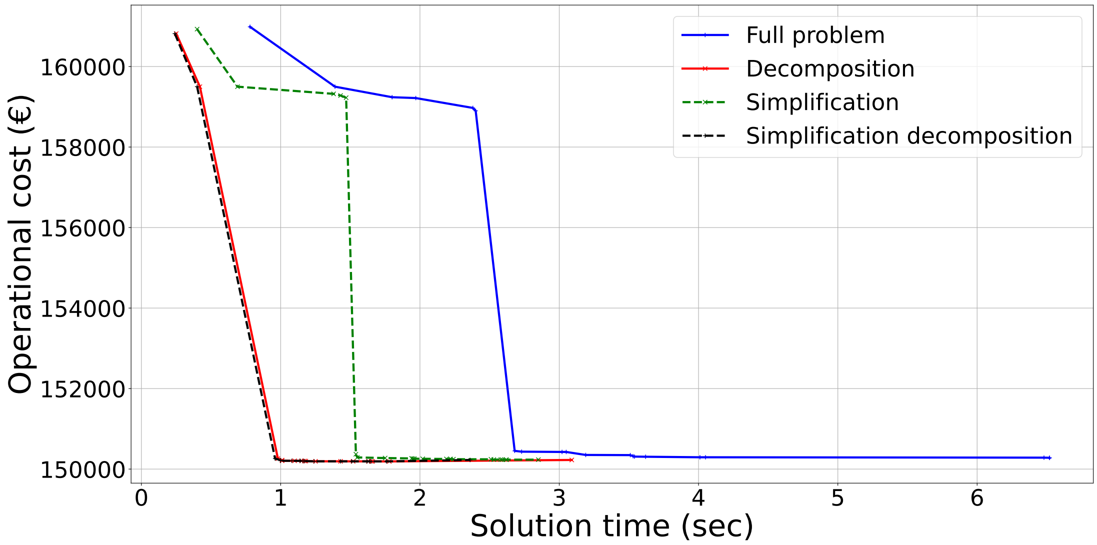


Figure 5.4: Solution quality (€) versus solution time (seconds) for Limerick dataset for the four examined approaches with carbon scenario 2025

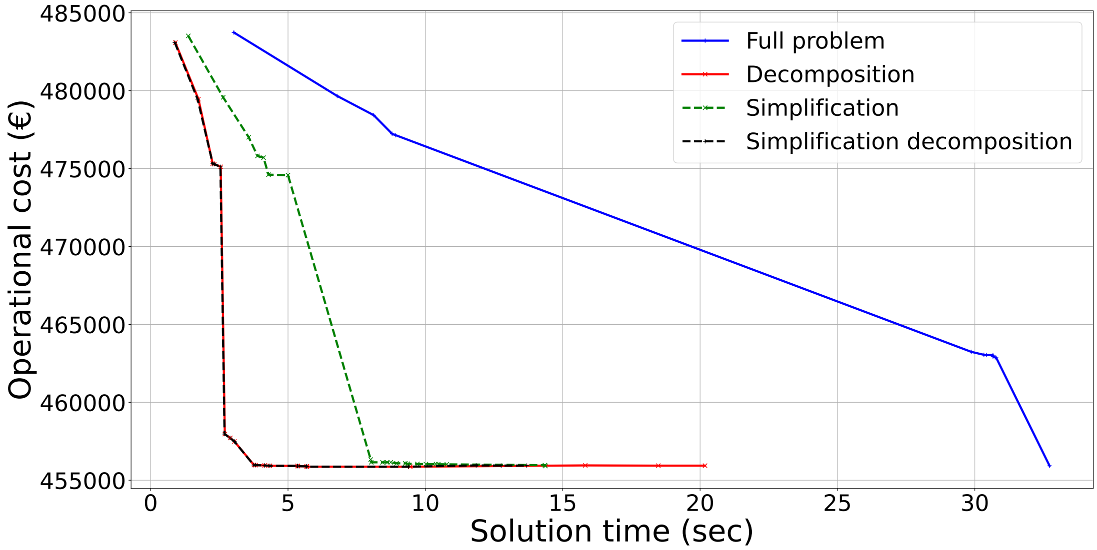


Figure 5.5: Solution quality (€) versus solution time (seconds) for Cork dataset for the four examined approaches with carbon scenario 2050

onds for the Cork dataset and 4 seconds for the Limerick dataset. Of note, the performance of the full problem approach in the Cork dataset is poor compared to the other approaches examined. With the approach takes 62.14% longer compared to the decomposition approach to find the optimal solution. This can also be compared to the approaches' performance on the Limerick dataset, where it is only 60.85% slower compared to the decomposition approach. This indicates that the full problem approach scales poorly with the dataset size. However, when examining the carbon scenario 2025 for Cork, the full problem

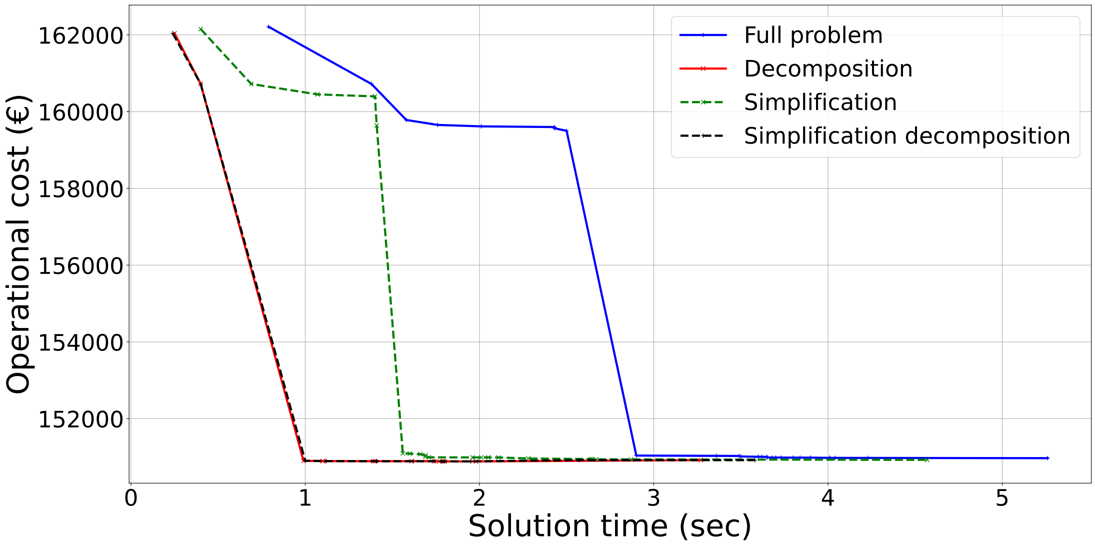


Figure 5.6: Solution quality (€) versus solution time (seconds) for Limerick dataset for the four examined approaches with carbon scenario 2050

Table 5.4: The mean search time before TOU constraints are introduced for decomposition approaches, and the mean improvement between the VOT, Fuel, SCC, and score between the first step of the decomposition approaches, and the final solution for Limerick and Cork datasets

City	Config.	Pre TOU time	ΔVOT	ΔFuel	ΔSCC	ΔScore
Limerick	2025 D	2.16	0	0	0	0
	2025 SD	2.14	0.3934	0.7384	0.0459	1.1777
	2050 D	2.08	0	0	0	0
	2050 SD	2.08	1.4587	-0.5147	-0.3188	0.6253
Cork	2025 D	9.01	0.8355	8.451	0.2629	9.5496
	2025 SD	8.97	4.0909	9.9521	0.2592	14.3023
	2050 D	8.10	4.7832	13.1753	1.3737	19.3322
	2050 SD	8.13	0.3584	10.5995	0.8439	11.8017

is only 80.50% worse than the decomposition approach, while it is 111% worse for the Limerick dataset. This may be indicative of some speed up in exploration of the search space due to a higher weighting of one of the components of the objective function.

It is also observed that the score of the solutions found by the decomposition approach worsens at a particular point during the search process. This signifies the translation of the first part of the decomposition approaches to the second, where the average electricity price and SCC are exchanged for their realised values, and TOU constraints are introduced. The scores for both decomposition approaches before this point are approximate.

Table 5.4 reports the mean search time for the first part of the decomposition approaches, and the mean improvement between the two decomposition steps for VOT, fuel, SCC, and solution scores. As previously outlined, the difference between the first and second steps of both decomposition approaches is the exclusion of TOU constraints from the former and inclusion in the latter. As such, the first part of both decomposition approaches is the same, which is reflected in the low difference in reported solution times (between 0 and 0.04 seconds). However, the difference between the VOT, fuel, SCC, and solution score values should differ for the second step. As previously mentioned, with the particular values used, the global optima of the final solution are the same for both approaches. This is reflected in the low variance between metrics reported for the full problem and simplification decompositions. With a maximum variation of 4.2×10^{-3} of the optimal solution (Cork carbon scenario 2050 using decomposition approach). This falls within the previously mentioned CPLEX's relative MIP gap tolerance.

The difference between the first step and the final solution is small, with some instances reporting no change. This is due to the size of the datasets examined. For example, the improvements for the Limerick dataset are barely different. This is due to the shorter operation times of buses in the city, and the small number of charging stations, which result in less flexibility for when BEBs can charge. For the Cork dataset, more improvements in score can be observed when introducing TOU constraints, due to the longer operational hours, and the higher number of charging stations and buses. This would imply that larger datasets benefit more from TOU constraints.

Figure 5.7 shows the total electricity consumed for both parts of the decomposition approach versus the electricity price and SCC for the Cork dataset with carbon scenario 2050 for the randomly selected seed shown in Figure 5.5. Here, the electricity consumption is aggregated into 15-minute increments. Blue bars show the electricity consumption before TOU constraints are introduced, while orange represents consumption after TOU constraints are added. The introduction of TOU constraints makes a visible difference in when electricity is consumed, as some of the consumption is moved from electricity price and SCC peak around 17:00 to the pit at 13:30. The resulting change in when electricity is consumed in this figure is reflected in Table 5.4.

Figure 5.8 shows the average deviation in minutes from the original schedule per BEB versus the passenger load for the Limerick dataset. Carbon scenario

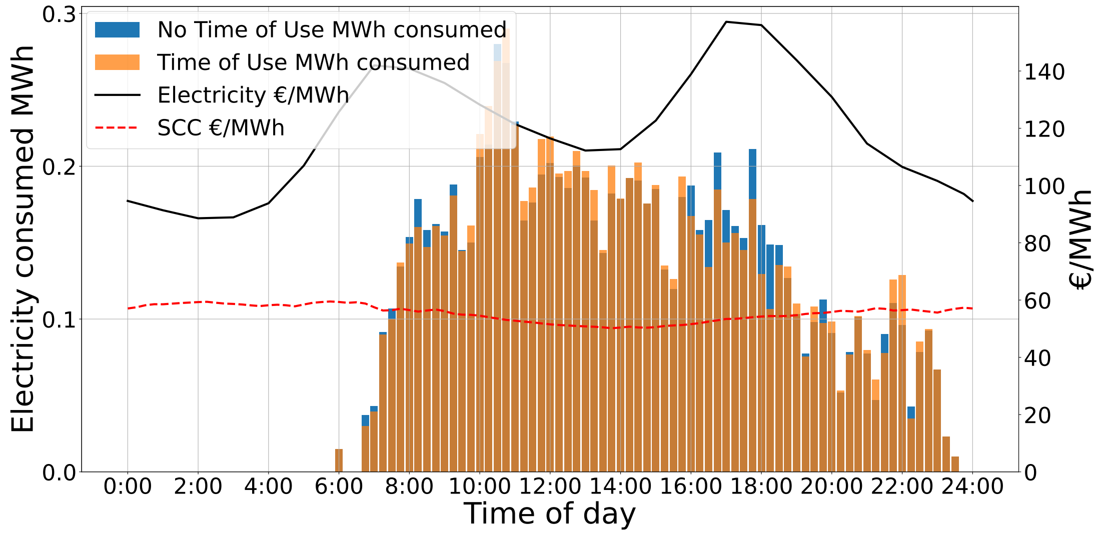


Figure 5.7: The total electricity consumed for parts one and two of the decomposition approach throughout the day versus the electricity price and SCC for the Cork dataset with carbon scenario 2050 for one randomly selected seed

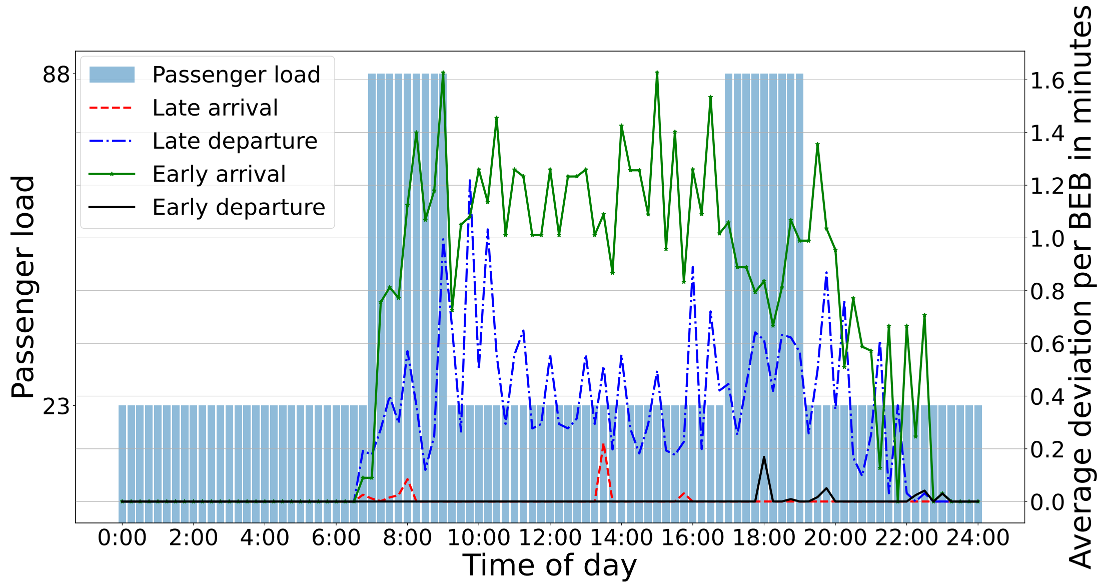


Figure 5.8: The average late arrival, late departure, early arrival, and early departure per BEB throughout the day versus passenger load for the Limerick dataset with carbon scenario 2050 and the decomposition approach for one randomly selected seed.

2050 is used as well as the decomposition approach, which is broken into 15-minute segments for the randomly selected seed examined in Figure 5.6. Here, each coloured line represents late arrivals, late departures, early arrivals, and early departures throughout the day. In particular, both negative timetable deviations, late arrival and early departure, are kept to a minimum. With the highest average negative delay peaking at 13.32 seconds, which occurs during a period with a lower passenger load. The highest average early departure does occur during peak passenger load. However, it results in passengers only spending an additional 10.14 seconds on public transport.

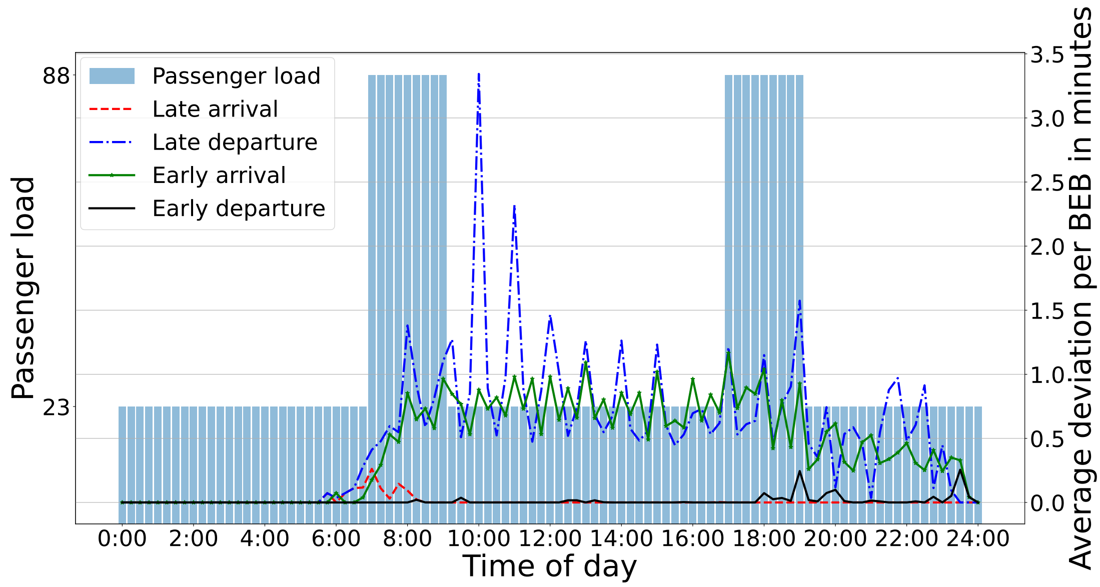


Figure 5.9: The average late arrival, late departure, early arrival, and early departure per BEB throughout the day versus passenger load for the Cork dataset with carbon scenario 2050 and the decomposition approach for one randomly selected seed.

Figure 5.9 shows the average deviation in minutes per BEB versus passenger load for the Cork dataset using carbon scenario 2050 and decomposition approach broken into 15-minute segments. Where late arrivals, late departures, early arrivals, and early departures throughout the day are represented as series. Here, similar trends where the highest average late arrival and early departure are kept to a minimum. The highest mean late arrival is 15.6 seconds, while early departure peaks at 15.3.

For both Figures 5.8 and 5.9, late departure and early arrival are considered beneficial. As they can reduce the total amount of time passengers spend on public transport, thus lowering VOT. These positive deviations occur throughout the day. However, for the Cork dataset, these exceed more than a 3.5-minute deviation per BEB in a given 15-minute segment, limiting the total divergence from the original schedule. Thus, the timetable deviation per BEB is minimal considering the aforementioned financial and emissions savings per day when adopting the generated mixed-fleet schedule.

Table 5.5 shows the mean solution score, the relative MIP gap, and percent-

Table 5.5: The mean score, gap, and the percentage saving between the solution found and a fully ICEB fleet for the Dublin dataset for both carbon scenarios, and all approaches

Config.	Score (€)	Gap(%)	Δ ICEB(%)
2025 F	3776397.0628	2.88x10^{-4}	4.3957
2025 D	**3776191.9620**	2.26x10^{-4}	**4.4009**
2025 S	3776244.1452	2.78x10^{-4}	4.3996
2025 SD	3776194.2948	**1.72x10^{-4}**	4.4009
2050 F	3786121.4762	4.17x10^{-4}	4.7752
2050 D	**3785539.2134**	2.21x10^{-4}	**4.7899**
2050 S	3785546.3044	2.65x10^{-4}	4.7897
2050 SD	3785541.0956	**1.7x10^{-4}**	4.7898

age savings compared to a fully ICEB fleet across all five random seeds for the Dublin dataset. No experiment using the Dublin dataset found the optimal solution within the timeout. However, the relative gap is very small. With the worst-performing experiment configuration obtaining an average relative gap of 4.17x10^{-2}.

The trend of both decomposition approaches outperforming their non-decomposition counterparts continues, with the score, relative gap, and percent savings consistently favouring the former over the latter. Unlike experiments for the Cork and Limerick datasets, the full problem decomposition approach obtains a better global minima compared to simplification decomposition, despite having a worse mean relative gap. As previously outlined, the global minima for the full problem approaches should be lower compared to the simplification approaches due to the inclusion of non-linear charging. While experiments with the Cork and Limerick datasets did not need to fully exploit non-linear charging to reach the global minima, this is not the case for the Dublin dataset. Of the 10 experiments conducted with the Dublin dataset and the decomposition approach, 2 used battery capacities above C_0. However, it should be noted that the difference between the mean scores for both decomposition approaches is only €2.33 and €1.88 per day for carbon scenarios 2025 and 2050, respectively. This would imply that there is very little difference between non-linear sharing and limiting capacity to 85% of the maximum rated capacity when using realistic datasets and terminal-based charging. Alternatively, it can also motivate the investigation for lower battery capacities compared to the ones currently in operation for Dublin bus.

When evaluating the percentage savings compared to a fully ICEB fleet, savings for the Dublin dataset are lower than those of the Cork and Limerick experiments. With the highest percentage savings for the Dublin dataset 0.5219% lower than the same configuration for Cork (Decomposition with carbon scenario 2050). However, the total euro saved per day is much higher, with up to €173836.87 and €190444.82 saved per day for carbon scenarios 2025 and 2050, respectively. As previously outlined, changes to VOT can greatly impact overall savings values. Excluding VOT, €24661.88 can be saved for carbon scenario 2025 and €33979.49 for carbon scenario 2050 per day when transitioning from a fully ICEB fleet to a mixed one. This also includes a carbon emissions reduction of 101.52 tCO_{2eq} per day for carbon scenario 2025, with a further reduction of 59.5 Kilograms of Carbon Dioxide Equivalent (kg/CO_{2eq}) when considering the increased weighting of SCC in carbon scenario 2050.

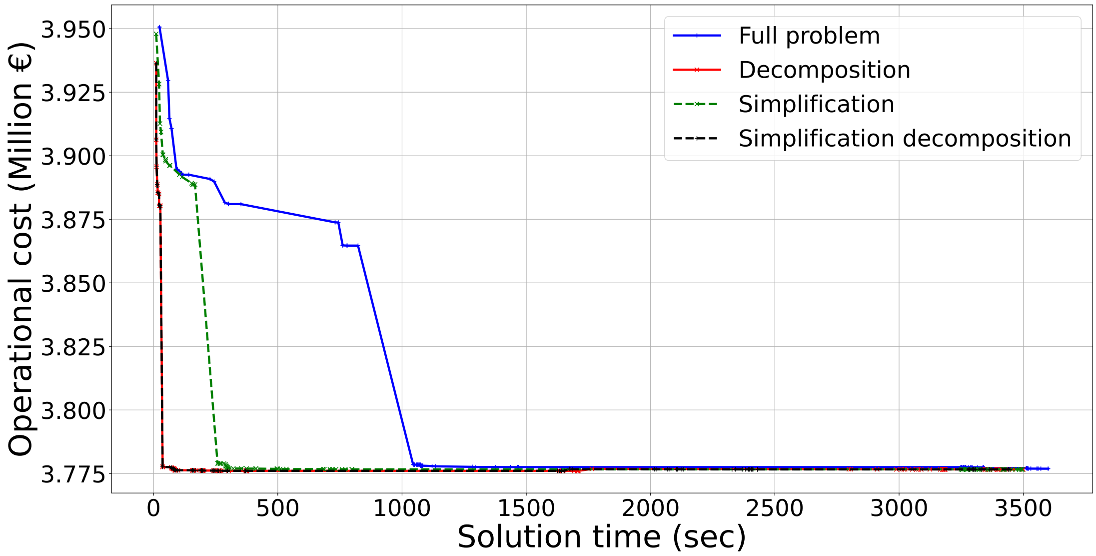


Figure 5.10: Solution quality (Million €) versus solution time (seconds) for Dublin dataset for the four examined approaches with carbon scenarios 2025.

Figures 5.10 and 5.11 show the solution quality in million € versus solve time in seconds for the Dublin dataset for the 2025 and 2050 carbon scenarios for one randomly selected seed. These figures show similar patterns observed in Figures 5.3–5.6. Where both decomposition approaches make quick progress towards their respective global minimas and can obtain a solution with a relative MIP gap of less than 2% within two minutes for both carbon scenarios. As previously stated, no experiment using the Dublin dataset was able to find the optimal solution within the timeout, this includes the first part of the decomposition approaches, as all experiments are observed to plateau when reaching a relative MIP gap of approximately 0.05% and make very little progress after-

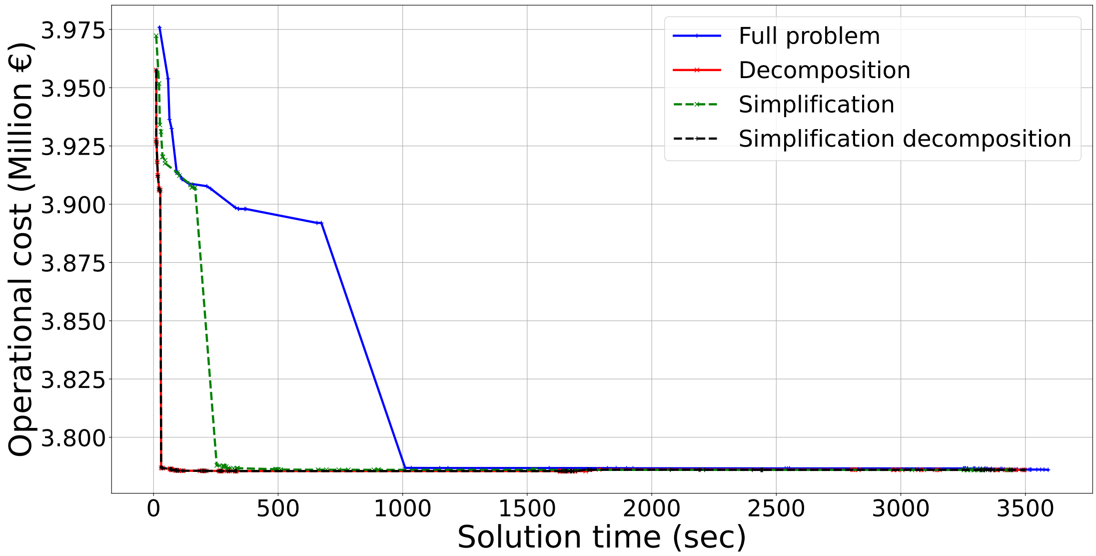


Figure 5.11: Solution quality (Million €) versus solution time (seconds) for Dublin dataset for the four examined approaches with carbon scenarios 2050.

Table 5.6: The mean change in VOT, Fuel, SCC and Score when introducing TOU constraints in part two of the decomposition approaches for the Dublin dataset for both carbon scenarios

Config.	ΔVOT(€)	ΔFuel(€)	ΔSCC(€)	ΔScore(€)
2025 D	15.8522	187.2486	8.6491	211.7500
2025 SD	15.5240	183.3760	7.0935	205.9934
2050 D	13.8591	183.6443	24.9140	222.4175
2050 SD	17.7716	181.8571	21.2320	220.86

wards. However, this highlights the need for decomposition approaches when dealing with larger datasets.

Table 5.6 reports the difference in VOT, Fuel, SCC, and score between the best solution found for parts one and two of the decomposition approaches. The difference between the best solution found without TOU and with TOU constraints is greater for the full problem decomposition. Both full problem decomposition and simplification decomposition are subject to the same constraints in part one. This difference reflects the better global minima that part two of the full problem decomposition has, as the score difference of the two decomposition approaches is consistent with the score difference seen in Table 5.5.

Introducing TOU constraints showcases additional savings of up to €211.75 and €222.41 per day for carbon scenarios 2025 and 2050, respectively. However, it should be noted that these additional values are small compared to the €24661.88 and €33979.49 saved per day for each scenario. Constituting

0.8586% and 0.6545% of the total savings for the 2025 and 2050 scenarios, respectively. Further reductions of carbon emissions are observed when introducing TOU constraints, with an additional reduction of 95 kg/CO_{2eq} for carbon scenario 2025, and 98.8 kg/CO_{2eq} for 2050.

While additional savings are small compared to the final solution, they are more significant compared to the maximum additional saving of 4.2x10^{-3} observed in Table 5.4 (Cork dataset using carbon scenario 2050 and the decomposition approach). This enforces the claim that larger datasets benefit more from TOU constraints, as longer operating hours and a higher number of charging stations and buses allow more flexibility when energy is consumed.

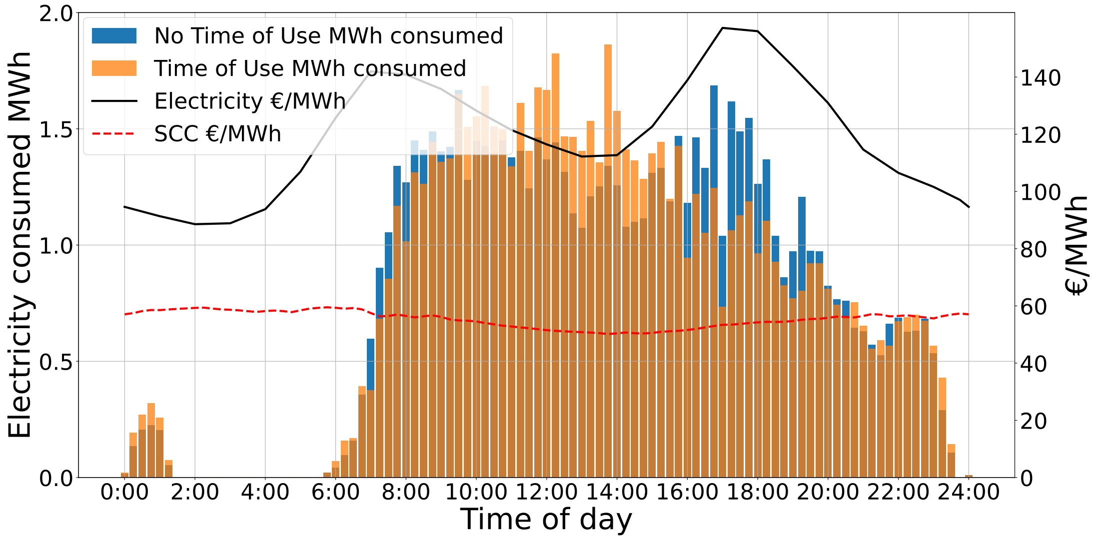


Figure 5.12: The total electricity consumed for parts one and two of the decomposition approach throughout the day versus electricity price and SCC for Dublin dataset with carbon scenario 2050 for one randomly selected seed

Figure 5.12 shows the energy consumed for parts one and two of the full problem decomposition approach with carbon scenario 2050 for the Dublin dataset for one randomly selected seed with electricity prices and SCC. Similar to Figure 5.7, introducing TOU constraints results in a notable change in when electricity is consumed during the operational day. Like the results for the Cork dataset, the electricity consumption during the electricity price peak around 17:00 and 7:00 is moved to the pit around 13:30. However, for the Dublin dataset, the movement of electricity consumption is more pronounced, with 0.55 MWh less electricity consumed at 17:15 and 0.52 MWh more electricity consumed at 13:45. This movement of electricity consumption only represent a small portion of savings as seen in Table 5.6 compared to the overall reduction in fuel cost and SCC when transitioning to BEBs. However, these findings do prove that the consideration of TOU windows results in reductions for both fuel prices and carbon emissions.

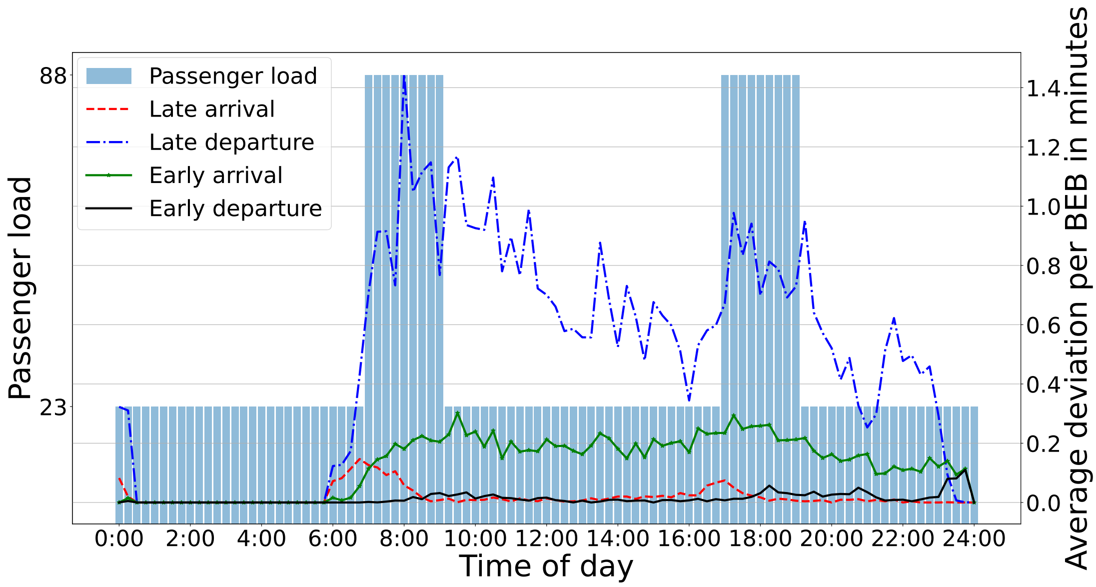


Figure 5.13: The average late arrival, late departure, early arrival, and early departure per BEB throughout the day versus passenger load for the Dublin dataset with carbon scenario 2050 and the decomposition approach for one randomly selected seed.

Figure 5.13 shows the average late arrival, late departure, early arrival, and early departure per BEB versus passenger load throughout the day for the Dublin dataset. With carbon scenario 2050 and the decomposition approach for one randomly selected seed aggregated into 15-minute segments. Here, similar patterns are observed from Figures 5.8 and 5.9, where delays caused to each BEB are relatively small for the savings obtained. For example, the highest average late arrival per BEB in a 15-minute segment is 8.802 seconds. In comparison, the highest early departure value is 6.63 seconds, with both peaks occurring outside the highest passenger load times. In terms of positive schedule deviations such as early arrivals and late departures, they are kept relatively low. For example, the maximum deviation for late departure is 1.1453 minutes (68.71 seconds). This shows that the overall timetable deviation per BEB is low compared to the significant savings that can be obtained. Of note, delays incurred for the Dublin dataset are lower compared to those for the Cork dataset, this can be attributed to a higher level of adoption of BEBs for the former compared to the latter.

Table 5.7 shows a summary of experiments conducted for this chapter in terms of the best-performing approach per dataset. This includes execution time in

Table 5.7: The best-performing approach for each dataset using carbon scenario 2050 with mean execution time, the mean percentage of the fleet that is BEB, the mean percentage of fuel savings, the mean percentage of carbon emission reductions, and the highest mean negative delayed per BEB

City	Approach	Time(s)	BEB%	Fuel%	Carbon%	Max. delay(s)
Limerick	SD	4.05	60	26.76	39.26	13.32
Cork	S	15.40	69.63	26.35	40.20	15.73
Dublin	D	3600	88.60	50.56	64.02	8.8

seconds, percentage saving in fuel cost, percentage of carbon emission reductions, and the highest average late arrival per BEB for each dataset using carbon scenario 2050. From this, it can be seen that the larger the dataset, the higher the penetration of BEBs in the bus fleet. Consequently, the total percentage saving in fuel cost and carbon emissions also increases, with the Dublin dataset able to reduce fuel costs by over 50% and reduce carbon emissions by almost 2/3. This is contrasted by the execution time. Experiments with the Dublin dataset were unable to find optimal solutions within the timeout. However, a solution very close to the global minima is found very quickly when using the decomposition approaches.

Overall, the decomposition approaches prove to perform better compared to their full problem counterpart. While simplification and simplification decomposition outperform the full problem decomposition for the Cork and Limerick datasets, the differences are minimal. Negative service impact is also shown to be kept to a minimum, with the highest average late arrival per 15-minute segment peaking at 8.8 seconds per BEB for the largest dataset. Smaller datasets observe higher late arrival values. However, this is attributed to a smaller penetration of BEBs in the fleet and a lower number of charging stations.

5.4.3 Objective function analysis

To determine the effectiveness of the three aspect objective function, additional experiments are conducted to evaluate the performance of COP when only optimising fuel price or SCC, alongside minimising VOT. To this end, ten additional experiments took place, with a one-hour timeout, using the full problem solution approach and the Dublin dataset. Five of these experiments are given the objective of minimising fuel price and VOT, while the remaining executions minimise SCC and VOT. The five experiments for each objective feature the same random seeds used in previous experiments, and carbon scenario 2050

Table 5.8: Mean score, VOT, SCC, and fuel costs for experiments conducted with different objective functions for the Dublin dataset with full problem solution approach using carbon scenario 2050

Objective function	Score (€)	VOT(€)	SCC(€)	Fuel(€)
Min. SCC+VOT	3785479.01	3751721.04	14028.11	19729.85
Min. Fuel+VOT	3785421.53	3751758.86	14042.11	19620.56
Min. Fuel+SCC+VOT	3786121.47	3751978.18	14500.05	19643.22

is used for experiments to minimise SCC. After the experiments, the fuel price of solutions generated using the SCC and VOT objective function is calculated, as well as the SCC for solutions found with the fuel price and VOT objective function. Results are then averaged across the five seeds for each category.

Table 5.8 shows the mean score, VOT, SCC, and fuel costs for the Dublin dataset, full problem solution approach, and carbon scenario 2050 for the different objective functions explored. Here, it can be seen that the objective functions that minimise SCC or fuel price separately achieve better scores than the objective function that optimises both costs and VOT. None of the additional experiments using different objective functions found the optimal solution within the one-hour timeout. However, the mean and median gap to the optimal solution for both objective functions were smaller compared to the objective function which minimises SCC, fuel, and VOT costs. The SCC+VOT objective reports a mean gap of $3.23\text{x}10^{-4}$, fuel+VOT reports $3.41\text{x}10^{-4}$, and SCC+fuel+VOT achieves a mean gap of $4.17\text{x}10^{-4}$. This highlights the difficulty of optimising all three objectives at once. Examining the two objective functions, it is clear that both achieve similar scores, with both scoring solutions within 0.001% of each other. As expected, the objective function which minimises SCC and VOT achieves a better SCC score, while the one that minimises fuel and VOT obtains a better fuel score. One thing of note is the improved score for VOT when optimised with SCC compared to when it is optimised with fuel costs. This is likely due to the lower weighting of SCC compared to the electricity price per MWh, as previously seen in Figure 5.12. This opens some interesting avenues of future work when integrating passenger VOT along with other objectives. Also of note is an increase in penetration of BEBs for both objective functions compared to fuel+SCC+VOT. With fuel+VOT featuring 1 more BEB on average than fuel+SCC+VOT, and SCC+VOT increasing the number of BEBs by an additional 0.4 on average.

Figure 5.14 shows the energy consumption for the Dublin dataset, full problem

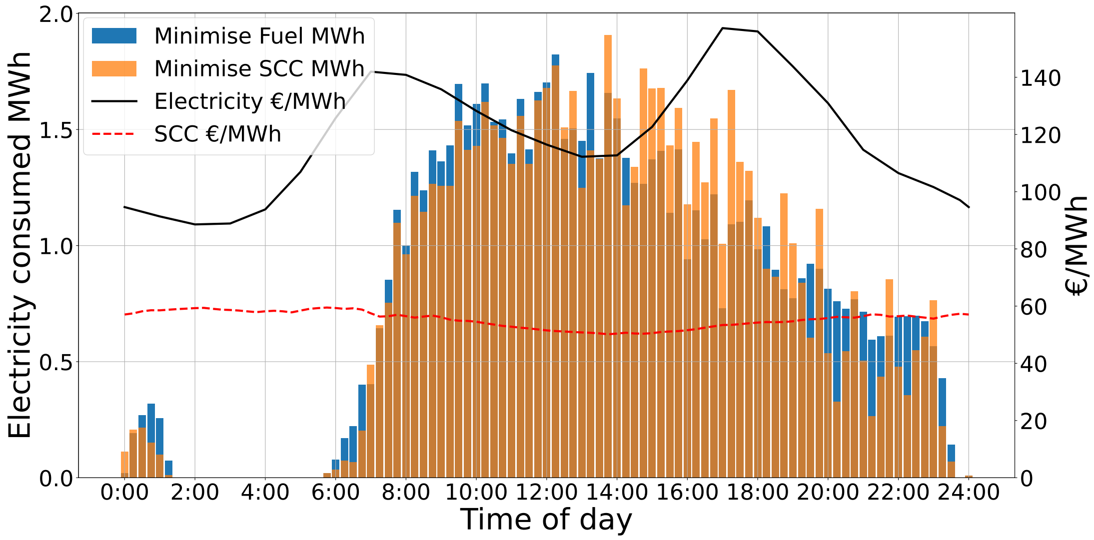


Figure 5.14: The total electricity consumed for both the fuel and SCC focused objective function throughout the way versus electricity price and SCC for the Dublin dataset with carbon scenario 2050 for one randomly selected seed with the full problem solution approach.

solution approach, and carbon scenario 2050 for the two objective functions for the same randomly selected seed used in Figure 5.12, where the blue bars represent energy consumption for the objective function that minimises fuel and VOT, while the orange bars represent the energy consumption for the objective function that minimises SCC and VOT and the darker colour represents an overlap of the two. Here, it can be seen that, while both objectives achieve similar scores overall, the distribution of charge events and enrgy consumption vary between the two objectives, with the SCC focused objective function consumping a large amount of energy around the 17:00 electricity price peak, and the fuel based objective function consuming more neergy around the carbon emission peak around 22:30. As seen in Figure 5.12, a more balanced distribution of energy is possible. This highlights the need to consider both costs in the same objective function, as improvements to the solution approach (such as using decomposition approaches) can improve the overall performance of the COP. This is further enforced by the fact that the mean score for the decompositional solution approach for the Dublin dataset in the carbon scenario 2050 is only 117 euro worse than the mean score reported by the fuel+VOT objective function, despite the fact that the former features a more complex objective function.

5.5 Conclusion

In this chapter, the optimisation model created in Chapter 4 was extended to consider a mixed-fleet of BEBs and ICEBs. A piecewise linear approximation for non-linear battery charging was also developed, and TOU windows are extended to measure the price and carbon emissions of the electricity used. This is combined with considerations for passengers' VOT to create an objective function to minimise the daily operational costs of the bus fleet. The resulting model is evaluated on the Cork and Limerick dataset outlined in Chapter 4 Section 4.2.3 Page 90), as well as a dataset for Dublin, which features 1088 more buses and 3701 more bus stops than the next largest dataset. Decomposition approaches are also examined as well as simplification approaches to improve the performance of larger datasets.

Evaluation with electricity price and carbon emissions shows that the COP model generates schedules that can reduce carbon emissions by up to 64% for the largest dataset, which constitutes a reduction of 101.52 tCO_{2eq} per day. Fuel costs can also be reduced by up to 50% per day, which represents significant financial savings which can offset the initial investment in adopting an BEB fleet. All of this can be done while keeping negative customer impact to a minimum, with the highest average late arrival only 8.8 seconds per BEB for the Dublin dataset with an average BEB adoption rate of 88.60%.

While the smaller datasets show lower adoption rates of BEBs, and consequently lower fuel cost and carbon emissions reductions, high-quality schedules are generated very quickly. The best-performing approaches find optimal solutions in 4.05 seconds for the Limerick dataset and 15.40 seconds for the Cork dataset. While the optimal solution for the Dublin dataset is never obtained, the decomposition approaches employed show that solutions close to the global minima can be obtained quickly. Evaluation also shows that considering non-linear charge times does not make a major impact on solution quality, with smaller datasets such as Limerick and Cork seeing no benefit. While the larger Dublin dataset sees an improvement in solution quality, the global minima are only approximately €3 different. The use of decomposition approaches also shows that the consideration of TOU windows for electricity price and carbon emissions can be reduced by an additional $\approx$180 euros and 95–98.8 kg/CO_{2eq} per day, respectively, for large datasets such as Dublin.

Additional experiments evaluated the impact of the different aspects of the objective function. While it was found that optimising for SCC and fuel costs separately generates similar solutions, the distribution of charge events and energy consumption is very different. It was also found that optimising for SCC leads to improved scores for VOT. This highlights that minimising emissions has a more positive impact on service quality compared to minimising fuel costs.

This chapter focused on the creation of schedules for the BEB-CSP to minimise CO_{2eq} while considering passengers' VOT and its modifiers. Additionally, multiple weightings for carbon emissions were examined and the performance of large datasets such as Dublin was evaluated. This fills the literature gap outlined in Chapter 2 Section 2.11 (Page 45). Furthermore, the presence of such reductions to CO_{2eq} while maintaining low negative service impacts such as late arrivals and early departures for the BEB-CSP is considered Contribution 3 of this thesis as detailed on Page 6.

Chapter 6

Conclusions & Future Work

This dissertation proves the need to take into consideration Carbon Dioxide Equivalent (CO_{2eq}) emissions of electricity used to charge Battery Electric Bus (BEB) for the Battery Electric Bus Charging Scheduling Problem (BEB-CSP). To this end, an empirical evaluation of the method for estimating available clean energy was performed. An extensive examination of two predictive models' impact on solving a COP for the BEB-CSP was carried out. Potential negative service impact on extremely large datasets has also been analysed. In this chapter, each contribution is tied to the original thesis statement in Section 6.1. In Section 6.3 potential future works on the topic of reducing CO_{2eq} emissions and the BEB-CSP are discussed.

6.1 Conclusion

To conclude this dissertation, the thesis statement from Chapter 1 is recalled:

> *The lifecycle emissions of battery electric buses can be reduced by using predictive models to forecast low-emission energy availability and by optimising charging schedules, all while maintaining high service quality.*

This claim is supported by the three contributions made within this thesis.

6.1.1 Predictive models

To utilise a predictive model to forecast low-emission energy availability, analysis must be performed to generate a dataset and select the best model for the task. In Chapter 3 this analysis is performed by gathering data from the Irish national grid. A dataset representing the future projected excess of wind-generated energy in Ireland was created. However, estimating this low-emission energy posed an imbalanced regression problem. To resolve this, an empirical evaluation of imbalanced regression methods, such as resampling strategies and loss metrics, is performed. Furthermore, Deep Learning (DL) approaches such as Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), and Convolutional Neural Network (CNN) models were also examined.

It was proven that CNNs trained with an Squared Error Relevance Area (SERA) loss function generalised well when predicting the availability of low-emission energy. However, it was also found that CNNs trained with a Inverse Weighted Mean Squared Error (IW-MSE) provided the second-best results, and can more accurately estimate extreme amounts of clean energy. These two configurations were extended to provide forecasts 12, 18, and 24-hours into the future, as well as the 6-hour model used during the evaluation. This was done to determine which type of estimation is most beneficial when solving a BEB-CSP.

6.1.2 Battery Electric Bus Charging Schedule Problem and low-emission energy

In Chapter 4 the two best predictive configurations from Chapter 3 were used to provide estimates to solve a COP model for the BEB-CSP. The formulated model represented a pure BEB-CSP for a fleet of BEBs with partial charging. Limits were placed on charging stations with an objective function of minimising energy consumption outside the Time-of-Use (TOU) windows, denoting the availability of low-emission energy (dubbed non-clean energy). The performance of the BEB-CSP using estimations from the two predictive configurations is analysed and compared to a naïve and ideal scenario. Where P_1 and P_2 represent the best and second-best configurations from Chapter 3, respectively. This extensive evaluation took place using 14 test days, three different datasets, and three different battery capacities. In addition, two methods for dealing with the predictive model were used, the first uses all forecasts. The second only utilises TOU information for the next 6-hours, and manipulates the schedule to be more adjustable for future prediction changes.

Baseline analysis showed that the use of P_2 can reduce the amount of non-clean energy by up to 10.04% in some scenarios, with solutions only improving by 4.89% on average when using 100% accurate information about low-emission energy. Further sensitivity analysis took place, evaluating the impact of alternative charging infrastructure, maximum allowable schedule deviations, increased energy availability, and increased energy consumption. These additional experiments showed P_1 outperformed P_2 in cases where the energy demand of the BEB fleet is low and more clean energy is available.

This showed that, on average, use of predictive models was significantly better compared to the naïve scenario. Additionally, if there is a large energy demand for the BEB fleet, predictive models that can accurately estimate extremes in clean energy available generate better solutions.

6.1.3 Lifecycle emissions and service quality

In Chapter 5 the COP from Chapter 4 was expanded to include public transport users Value-of-Time (VOT) representing service quality. This is combined with the overall CO_{2eq} emissions and fuel costs for a mixed fleet of BEBs and Internal Combustion Engine Bus (ICEB) to minimise overall daily operational costs. Furthermore, VOT modifiers were considered, such as the impact of late arrivals versus early departures. Two datasets from Chapter 4 was evaluated along with a dataset comprised of over 1000 buses, along with historical information for electricity price and CO_{2eq} emissions of the Irish national grid. The impact of non-linear charging functions was examined, along with different weightings for CO_{2eq} emissions, with decomposition approaches explored for dealing with large datasets.

The evaluation showed replacing parts of an ICEB fleet with BEBs can lead to significant savings. For the medium-sized dataset examined with 81 buses and a 69.63% adoption rate of BEBs fuel costs were reduced by 26.35%, and CO_{2eq} emissions are 40.2% lower. This constituted a saving of 8.31 Tonne of Carbon Dioxide Equivalent (tCO_{2eq}) per day. This was accomplished while limiting negative service quality impact to an average of 15.6 seconds of late arrivals, and 15.3 seconds of early departures per 15-minute segment of a BEB. Analysis of the large dataset showed much more significant savings. With a BEB adoption rate of 88.6%, fuel costs can be halved, and CO_{2eq} emissions are reduced by 64.02%. This represented a decrease of 101.52 tCO_{2eq} per day while only increasing average late arrivals of BEBs by 8.8 seconds.

Analysis of the impact of non-linear charging showed no major effects on smaller datasets, due to the battery capacity selected, while larger datasets saw a marginal impact of additional savings of €1.88–€2.33 per day. Consideration of TOU windows for electricity prices and grid CO_{2eq} emissions leads to an additional savings of €211.75 and 95 Kilograms of Carbon Dioxide Equivalent (kg/CO_{2eq}) per day for the largest dataset. When examining a higher weighting for CO_{2eq}, the amount of reductions increased to 98.8 kg/CO_{2eq} per day. This showcases that the BEB-CSP can significantly reduce CO_{2eq} emissions while inflicting minimal negative service quality changes.

6.2 Applications and limitations

While data and datasets from Ireland have been primarily used to support this dissertation's statement, the methodologies and frameworks developed can be applied to other countries and domains. Countries such as Scotland also have a high penetration of wind energy in their national grid [Sco21]. The methodologies employed in Chapter 3 can be used to estimate potential wind energy excess from wind-generated power in such countries. The COP outlined in Chapter 4 is agnostic of the type of clean energy used to charge BEB. Therefore, any type of excess clean energy, such as solar and hydro, could be used instead of wind power. Furthermore, both COP models created in this dissertation can be applied to other cities that have a fleet of BEBs and on-route charging. For instance, the COP in Chapter 5 can be used in different cities, as long as the relevant information, such as electricity price, carbon emissions, and passenger VOT is provided. In addition, while Chapter 5 uses naïve values for passenger load, the formulation of the COP allows for more detailed passenger load information to be used, if it is available. With this in mind, the COPs developed in this dissertation can be used by bus operators to determine the best deployment of BEBs for both minimising operational costs (such as fuel) and reducing the overall emissions of the fleet to meet any emission targets put in place by policy makers. Furthermore, the combination of DL models and COP developed in Chapters 3 and 4 can be used to reduce the lost low-emission energy from oversupply of renewable sources. This can have an effect of lowering the overall emissions of the energy sector of a country, and aid in meeting carbon emission targets.

There are, however, some limitations to this work. For instance, it is assumed that charging stations are placed along bus routes or at terminal stops, and there is a basis for this in both literature and the real world (e.g. the pantograph used to charge BEBs in Berlin). However, some countries, including Ireland, plan to start with a depot charging infrastructure where BEBs can only charge at the depot. To ensure the COP models developed in this dissertation work with this type of infrastructure, minor modifications would have to be made to represent the time and energy needed to travel back to the depot when charging is required. Additionally, the first contribution of this work focuses on the estimation of excess wind energy above national demand. While there is precedent for this, such as avoiding wind curtailment and the current estimated oversupply of renewable energy, future development of energy-intensive services (such as data centres) may reduce the likelihood of such excesses occurring. As such, the usefulness of evaluating imbalanced regression methods for DL approaches to estimate excess wind energy may be limited.

6.3 Future work

This dissertation performed an extensive analysis of predictive models for low-emission energy and COPs for the BEB-CSP to reduce CO_{2eq} emissions. However, there are a number of ways works featured within this thesis can be extended to further reduce CO_{2eq} emissions and model the BEB-CSP in more accurate ways.

Looking to the immediate future, additional work can adapt metaheuristic approaches like those discussed in Chapter 2 Section 2.5.3 (Page 25) for the BEB-CSPs outlined in this thesis. While work has been done using approaches such as Genetic Algorithm (GA) to solve the BEB-CSP, none have been used for the model featured in Chapter 5. Furthermore, this thesis represented time as a continuous domain variable, while other works represented time as a discrete variable. A comparison between the two representations of time can be undertaken for the BEB-CSP. This can include the TOU windows featured in this dissertation, to determine if a reduction in search time is achieved on large datasets and the degree of negative impact on solution quality. Additionally, the Nondeterministic Polynomial time (NP)-hardness of the BEB-CSP used in this thesis can be proven. While Chapter 3 performed an evaluation of imbalanced regression methods, the examined methods could be expanded, and additional datasets evaluated. Furthermore, in Chapter 4 it was noted that an alternative method of dealing with mispredictions can be examined. With charging events evenly spread throughout the day. As discussed in Chapter 3 Section 3.5 (Page 64), there is room for a comprehensive analysis for using both linear models, such as DLinear, and transformers for imbalanced regression time series problems with non-linear datasets. Comparisons can be drawn to the performance of the DL models evaluated in this thesis. Furthermore, the selection of imbalanced regression methods used could be expanded to include popular resampling strategies such as SMOTER.

In the longer term, the BEB-CSP used in this dissertation can be extended to represent the stochastic nature of some variables used. For example, travel time, energy consumption, and availability of low-emission energy are all highly volatile values that can be subject to change but were treated as deterministic. While the availability of low-emission energy was estimated via prediction models, those estimations were treated with 100% confidence. This can be addressed with contemporary methodologies such as quantile regression, which gives a probabilistic estimation of predictions made by the model [BSCG24]. This allows the estimation of the prediction interval of the model, representing the deviation from the true value. This can help create a more robust model and has already been implemented for predicting travel time. Additionally, while the use of $C_{\min}$ provides some buffer against uncertainty for energy consumption [CWSB13], more sophisticated approaches can be taken. Research into this area of stochastic scheduling for the BEB-CSP has gained significant attention in past years [ZWQ24, ZCL24, HNBM23]. While some works do consider the availability of renewable energy [ZL21], they do not analyse the impact of stochastic scheduling on the lifecycle emissions of the BEB-CSP.

An area of research, which could be the subject of an additional thesis, is the Vehicle-to-Grid (V2G) problem. This problem entails charging Battery Electric Vehicle (BEV) such as BEBs to store energy from the grid at times of low emissions or cheap electric tariffs and return it during points of high load or emissions [RNI+22, MTHA22]. This area has received a large amount of research in recent years. However, some works have brought up issues regarding battery degradation. After repeated charging and discharging cycles, batteries begin to degrade, reducing the amount of capacity that can be held. Eventually reaching their End-of-Life (EoL) as discussed in Chapter 2 Section 2.7.4 (Page 32). Additional charging and discharging cycles in the V2G problem may exacerbate this degradation, leading to more frequent battery replacement. The degradation of batteries, along with the lifecycle emissions associated with the creation of replacement batteries, is not considered in this dissertation, as these are potentially very large areas of research. For example, the most common type of battery used in BEV, and by extension BEBs, are Lithium-ion batteries. These batteries are very energy-dense and can take longer to reach their EoL compared to other battery types. However, recycling such batteries can prove problematic [CSF+25]. Furthermore, there are serious safety concerns with the use of such batteries [CKZ+21]. Alternatives such as Sodium-ion batteries prove to be more carbon efficient to manufacture than Lithium-ion batteries [LCC+23]. Additionally, these types of batteries are partially biodegradable due to their sodium component, aiding the recycling process. However, these batteries also face issues, being less energy dense and reaching their EoL faster compared to Lithium-ion batteries. The complexity of modelling different types of battery production CO_{2eq} emissions, along with issues faced when recycling batteries, could be integrated into the BEB-CSP, especially if a V2G component is considered. A new problem could be created to determine the best type of battery for a fleet of BEBs, where non-homogeneous battery types could be employed. The trade-off of these different battery types in terms of scheduling can be examined to determine the full lifecycle emissions of a fleet of BEBs. Furthermore, the benefits for using V2G as an integrated Battery Energy Storage System (BESS) can be compared to dedicated stationary facilities.

This is a non-exhaustive look at the areas of research the topic of reducing lifecycle emissions for the BEB-CSP can take. In addition, the models and methods developed can be extended to other BEVs, such as BEMUs. Extension of this research into these areas poses great interest, as the Irish government looks to upgrade the Irish rail network with electrified routes [Dep24b]. Despite this, some routes may not be suitable for electrification, either due to the remote and infrequent nature of track usage. Such scenarios pose an opportunity for the use of BEMUs to reduce reliance on diesel trains. This use of BEMUs can then engender a charging schedule problem, where the contributions within this thesis can be applied to reduce lifecycle emissions. Additionally, Ireland's second-largest city, Cork, is currently in the planning phase for introducing electric trams [Tra25]. This poses another opportunity, as overhead electrified lines that are traditionally seen for tram systems may not be possible in densely populated city centres where public transport lines may be shared with large buses. Thus compelling the adoption of battery-electric trams seen in other countries such as Spain, Australia, and Luxembourg.

Acronyms

ANN Artificial Neural Network. 12, 37, 41

ARIMA AutoRegressive Integrated Moving Average. 11–13, 19

BEB Battery Electric Bus. v, viii, 2–8, 10, 21, 27–32, 35–41, 43–47, 84–89, 92–100, 103, 104, 106–108, 113, 114, 118, 119, 121–127, 129, 130, 132, 133, 135, 137–139, 143–145, 149–152, 154, 156–162

BEB-CSP Battery Electric Bus Charging Scheduling Problem. 4–9, 30, 31, 33, 35–41, 44–47, 83–85, 87, 88, 90, 92, 97, 99, 104, 107, 114, 115, 117, 118, 122–124, 127, 129, 155–157, 159–162

BEMUs Battery Electric Multiple Units. 162

BESS Battery Energy Storage System. 3, 89, 162

BEV Battery Electric Vehicle. 1, 3, 27, 28, 161, 162

CC Constant Current. 32, 86, 126, 139

CNN Convolutional Neural Network. iv, vii, 5, 7, 16, 20, 48, 65, 66, 68–70, 72–83, 86, 90, 100, 157

CO_{2eq} Carbon Dioxide Equivalent. 1, 2, 6, 10, 37–39, 41, 43, 45, 46, 48, 50, 123–125, 127, 128, 155, 156, 158–160, 162

CO_2 Carbon Dioxide. 1

COP Constrained Optimisation Problem. iv, 7, 9, 21, 25, 26, 41–43, 45, 47, 53, 55, 73, 82, 84, 85, 87, 88, 92, 97, 98, 101–103, 105, 109, 112, 114–116, 122–125, 137, 138, 151, 153, 154, 156–160

CPU Central Processing Unit. 65, 101, 138

CSLP Charging Station Location Problem. 30

CSP Constraint Satisfaction Problem. 9, 21

CUDA Compute Unified Device Architecture. 64

cuDNN CUDA Deep Neural Network. 64

CV Constant Voltage. 32, 36, 126

DICE Dynamic Integrated model of Climate and the Economy. 43

DL Deep Learning. 3–5, 7, 9, 11–13, 16, 17, 19, 20, 28, 45, 48, 51, 54, 55, 64–66, 68, 71, 73, 74, 80, 82, 83, 86, 157, 159–161

DOD Depth of Discharge. 32

DT Decision Trees. 18

E-VRP Electric Vehicle Routing Problem. 27, 28

E-VSP Electric Vehicle Scheduling Problem. 7, 27, 30, 35

EoL End-of-Life. 32, 161, 162

EV Electric Vehicle. 1, 2, 7, 9, 10, 28, 43

FRVCP-BEV Fixed Route Vehicle Charging Problem for Battery Electric Vehicles. 27

G-VRP Green Vehicle Routing Problem. 27

GA Genetic Algorithm. 25, 36, 37, 160

gCO_{2eq}/km Grams of Carbon Dioxide Equivalent per Kilometre. 43, 138

gCO_{2eq}/kWh Grams of Carbon Dioxide Equivalent per Kilowatt hour. 48, 49

gCO_2/kWh grams of Carbon Dioxide per Kilowatt hour. 128

GDP Gross Domestic Profit. 1, 43, 44

GHG Greenhouse Gases. 1

GHz Gigahertz. 65, 101, 138

GN Gaussian Noise. 18, 58, 67–74, 76, 78

GPS Global Positioning System. 38, 91, 129

GPU Graphics Processing Unit. 65

GRU Gated Recurrent Unit. iv, 7, 14–16, 19, 20, 48, 65, 68–70, 72–76, 157

HEV-TSP Hybrid Electric Vehicle Travelling Salesman Problem. 28

HPO Hyperparameter Optimisation. vii, 65, 68, 74

ICE Internal Combustion Engine. 27, 43

ICEB Internal Combustion Engine Bus. viii, 2, 3, 6, 8, 27, 30–32, 39, 123–125, 127, 129, 131, 135, 138, 139, 146, 147, 154, 158

ILP Integer Linear Programming. 22, 38, 39

IP Integer Programming. 22, 23

IW-MSE Inverse Weighted Mean Squared Error. iv, vii, 5, 18, 61, 62, 64, 66–83, 86, 90, 100, 157

kg/CO_{2eq} Kilograms of Carbon Dioxide Equivalent. 147, 149, 154, 159

km Kilometre. 91, 100, 127, 137, 138

km/h Kilometres per hour. 91, 99, 137

kWh Kilowatt hour. viii, 31, 86, 92, 99, 101, 105, 112, 115, 121, 137

kWh/km Kilowatt hours per kilometre. viii, 31, 32, 99, 100, 118, 120

LP Linear Programming. 22–25, 37

LS Local Search. 25

LSTM Long Short-Term Memory. iv, 7, 11, 14–16, 19, 20, 48, 64, 65, 68–70, 72–81, 157

MAE Mean Absolute Error. 12, 42

MAPE Mean Absolute Percentage Error. vii, 81, 82

MD-VSP Multiple Depot Vehicle Scheduling Problem. 26

MILP Mixed Integer Linear Programming. 22, 37, 38, 40, 46, 101, 122, 124

MIP Mixed Integer Programming. 22–24, 41, 138, 143, 145, 147

MLP Multilayer Perceptrons. 18

MP Mathematical Programming. 7, 9, 22, 24–26, 36

MPM Multi-Period reactive Model. iv, v, 87, 88, 97, 101–122

MSE Mean Squared Error. vii, 12, 42, 61, 62, 64, 67, 69–83

$MtCO_{2eq}$ Million Tonnes of Carbon Dioxide Equivalent. 1, 2

MVT-VSP Multiple Vehicle Type Vehicle Scheduling Problem. 26

MW MegaWatt. 2, 3, 48, 49, 52, 100

MWh Megawatt hour. v, 125, 128, 129, 131, 137, 138, 149

MWh/km Megawatt hours per Kilometre. 137

NP Nondeterministic Polynomial time. 23, 26, 27, 42, 160

OC Opportunity Charging. 28–31, 36–38, 100

PCHIP Piecewise Cubic Hermite Interpolating Polynomials. iv, 17, 55, 57

PSO Particle Swarm Optimisation. 38, 40

RAM Random Access Memory. 65, 101, 138

ReLu Rectified Linear Unit. 16

RF Random Forest. 18

RMSE Root Mean Squared Error. 12

RNN Recurrent Nerual Network. iv, 13–16, 20, 69

SARIMA Seasonal AutoRegressive Integrated Moving Average. 11–13, 20

SCC Social Cost of Carbon. v, vi, viii, 6, 43, 46, 125, 127–129, 131, 132, 134–140, 142–144, 147–149, 151–153, 155

SD-VSP Single Depot Vehicle Scheduling Problem. 26

SDK Software Development Kit. 64

SEMO Single Energy Market Operator. 127

SER Squared Error Relevance. iv, 61, 62, 67, 75, 77–81

SERA Squared Error Relevance Area. iv, vii, 5, 18, 61, 62, 64, 66–83, 86, 90, 100, 157

SMOTE Synthetic Minority Over-Sampling Technique. 18

SMOTER Synthetic Minority Over-Sampling Technique for Regression. 18

SOC State of Charge. 28, 32, 126

SPC Shadow Price of Carbon. 43, 128

SPM Single-Period reactive Model. iv, v, 88, 89, 97, 98, 101–118, 120–122

SSE Sum of Squared Error. 61, 62, 66, 67, 71, 75, 79

STD Standard Deviation. vii, viii, 49, 50, 52, 59, 138, 139

tCO_{2eq} Tonne of Carbon Dioxide Equivalent. 6, 38, 40, 43, 49, 128, 140, 147, 154, 158

TOU Time-of-Use. viii, 5, 36, 38–43, 46, 85, 89, 90, 92, 96, 100–102, 127, 131, 132, 134–136, 142, 143, 148–150, 154, 157, 159, 160

TRU Temporal Relevance Undersampling. 58, 67, 69–74, 76, 78, 83

TSP Travelling Salesman Problem. 23, 27, 28

TWA-VRP Time Window Assignment Vehicle Routing Problem. 27

V2G Vehicle-to-Grid. 161, 162

VOT Value-of-Time. viii, 6, 8, 39, 40, 44, 46, 123–128, 131–133, 136, 138–140, 142, 143, 145, 147, 148, 151–155, 158, 159

VRP Vehicle Routing Problem. 26, 27

VRP-TW Vehicle Routing Problem with Time Windows. 27

VSP Vehicle Scheduling Problem. 7, 9, 25–27, 31, 33

VSP-TW Vehicle Scheduling Problem with Time Windows. 26

References

[AAAS20] Furkan Ahmad, Mohammad Saad Alam, Ibrahim Saad Alsaidan, and Samir M. Shariff. Battery swapping station for electric vehicles: Opportunities and challenges. *IET Smart Grid*, 3:280–286, 2020.

[AAK21] Matías Alvo, Gustavo Angulo, and Mathias A. Klapp. An exact solution approach for an electric bus dispatch problem. *Transportation Research Part E: Logistics and Transportation Review*, 156, 2021.

[AC20] Alejandro Arbelaez and Laura Climent. Transition to ebuses with minimal timetable disruptions. *Proceedings of the 13th International Symposium on Combinatorial Search, SoCS 2020*, None:119–120, 2020.

[ACC24] Juscimara G. Avelino, George D.C. Cavalcanti, and Rafael M.O. Cruz. Resampling strategies for imbalanced regression: a survey and empirical analysis. *Artificial Intelligence Review*, 57, 2024.

[AGS+07] R. Alejo, V. García, J. M. Sotoca, R. A. Mollineda, and J. S. Sánchez. Improving the performance of the rbf neural networks trained with imbalanced samples. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 4507 LNCS:162–169, 2007.

[ALKS+25] Alan Anis Lahoud, Ahmad Saeed Khan, Erik Schaffernicht, Marco Trincavelli, and Johannes Andreas Stork. Predict-and-optimize techniques for data-driven optimization problems: A review. *Neural Processing Letters*, 57, 2025.

[ALS+20] Syed Muhammad Arif, Tek Tjing Lie, Boon Chong Seet, Syed Muhammad Ahsan, and Hassan Abbas Khan. Plug-in electric bus depot charging with pv and ess and their impact on lv feeder. *Energies*, 13, 2020.

[AM17] Jonathan D. Adler and Pitu B. Mirchandania. The vehicle scheduling problem for fleets with alternative-fuel vehicles. *Transportation Science*, 51:441–456, 2017.

[AWAL20] Yaseen Alwesabi, Yong Wang, Raul Avalos, and Zhaocai Liu. Electric bus scheduling under single depot dynamic wireless charging infrastructure planning. *Energy*, 213, 2020.

[BHG21] Yiming Bie, Mingjie Hao, and Mengzhu Guo. Optimal electric bus scheduling based on the combination of all-stop and short-turning strategies. *Sustainability (Switzerland)*, 13:1–21, 2021.

[BJR13] George E.P. Box, Gwilym M. Jenkins, and Gregory C. Reinsel. *Time series analysis: Forecasting and control: Fourth edition*. John Wiley & Sons, 2013.

[BJWQ21] Yiming Bie, Jinhua Ji, Xiangyu Wang, and Xiaobo Qu. Optimization of electric bus scheduling considering stochastic volatilities in trip travel time and energy consumption. *Computer-Aided Civil and Infrastructure Engineering*, 36:1530–1548, 2021.

[BK09] Stefan Bunte and Natalia Kliewer. An overview on vehicle scheduling models. *Public Transport*, 1:299–317, 2009.

[BN24] Lint Barrage and William Nordhaus. Policies, projections, and the social cost of carbon: Results from the dice-2023 model. *Proceedings of the National Academy of Sciences of the United States of America*, 121, 2024.

[BRVN16] Kris Braekers, Katrien Ramaekers, and Inneke Van Nieuwenhuyse. The vehicle routing problem: State of the art classification and review. *Computers and Industrial Engineering*, 99:300–313, 2016.

[BSCG24] Dhivya Bharathi, Juan Manuel González Sopeña, Siobhan Clarke, and Bidisha Ghosh. Travel time prediction utilizing hybrid deep learning models. *Transportation Research Record*, 2678:56 – 65, 2024.

[BT19] Jatin Bedi and Durga Toshniwal. Deep learning framework to forecast electricity demand. *Applied Energy*, 238:1312–1326, 2019.

[BTR16] Paula Branco, Luís Torgo, and Rita P Ribeiro. A survey of predictive modeling on imbalanced domains. *ACM computing surveys (CSUR)*, 49:1–50, 2016.

[BTR19] Paula Branco, Luis Torgo, and Rita P. Ribeiro. Pre-processing approaches for imbalanced distributions in regression. *Neurocomputing*, 343:76–99, 2019.

[CBHK02] Nitesh V. Chawla, Kevin W. Bowyer, Lawrence O. Hall, and W. Philip Kegelmeyer. Smote: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, 16:321–357, 2002.

[CGCB14] Junyoung Chung, Caglar Gulcehre, KyungHyun Cho, and Yoshua Bengio. Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv preprint arXiv:1412.3555*, 2014.

[Cit25] Citizens Information Board. Citizens information electric vehicles, 2025. `https://www.citizensinformation.ie/en/travel-and-recreation/motoring/buying-or-selling-a-vehicle/electric-vehicles/` [Accessed: 2025-05-09].

[CKZ+21] Yuqing Chen, Yuqiong Kang, Yun Zhao, Li Wang, Jilei Liu, Yanxi Li, Zheng Liang, Xiangming He, Xing Li, Naser Tavajohi, and Baohua Li. A review of lithium-ion battery safety concerns: The issues, strategies, and testing standards. *Journal of Energy Chemistry*, 59:83–99, 2021.

[CME21] Magín Campos, Josep Mensión, and Miquel Estrada. Charging operations in battery electric bus systems at the depot. *Transportation Research Procedia*, 58:103–110, 2021.

[CSF+25] Carlos M. Costa, Manuel Salado, Chiara Ferrara, Riccardo Ruffo, Piercarlo Mustarelli, Rui Mao, Sheng Feng, Yuxiang Shang, Xiaochen Wang, Zhenkun Lei, Ruixiang Bai, Cheng Yan, Kwon Hyung Lee, Sang Woo Kim, Tae Hee Kim, Sang Young Lee, Long Kong, Qiang Zhang, Harsha Devnani, Shikha Gupta, James F. Rohan, Neil S. Curtis, Abhishek Lahiri, Yinghe He, and

S. Lanceros-Mendez. The wide range of battery systems: From micro- to structural batteries, from biodegradable to high performance batteries. *Progress in Materials Science*, 154, 2025.

[CWSB13] Laura Climent, Richard J. Wallace, Miguel A. Salido, and Federico Barber. A constraint programming approach to solve scheduling problems under uncertainty. *Proceedings of the 8th Workshop on Constraint Satisfaction Techniques for Planning and Scheduling Problems, COPLAS 2013*, None:28–37, 2013.

[CX13] Zhu Chao and Chen Xiaohong. Optimizing battery electric bus transit vehicle scheduling with battery exchanging: Model and case study. *Procedia-Social and Behavioral Sciences*, 96:2725–2736, 2013.

[CZZ24] Chenhao Cai, Leyao Zhang, and Jianguo Zhou. Dmpr: A novel wind speed forecasting model based on optimized decomposition, multi-objective feature selection, and patch-based rnn. *Energy*, 310, 2024.

[DAD18] Robinson Duque, Alejandro Arbelaez, and Juan F. Díaz. Online over time processing of combinatorial problems. *Constraints*, 23:310–334, 2018.

[DBEB20] James Dixon, Waqquas Bukhsh, Calum Edmunds, and Keith Bell. Scheduling electric vehicle charging to minimise carbon emissions and wind curtailment. *Renewable Energy*, 161:1072–1091, 2020.

[DDS92] Martin Desrochers, Jacques Desrosiers, and Marius Solomon. A new optimization algorithm for the vehicle routing problem with time windows. *Operations Research*, 40:342–354, 1992.

[Dep19] Department of Transport, Tourism & Sport, Irish Government. Report on diesel-and alternative-fuel bus trials, 2019. `https://assets.gov.ie/static/documents/report-on-diesel-and-alternative-fuel-bus-trials.pdf` [Accessed: 2025-05-19].

[Dep23a] Department of Communications, Climate Action & Environment, Irish Government. Climate action plan 2023, 2023. `https://assets.gov.ie/static/documents/climate-action-`

plan-2023-8b7dba92-f397-4163-8293-3aea656f6b5f.pdf [Accessed: 2025-05-09].

[Dep23b] Department of Transportation - Irish Government. Transport appraisal framework, appraisal guidelines for capital investments in transport, module 8 - detailed guidance on appraisal parameters, 2023. https://web.archive.org/web/20230613234715/https://www.gov.ie/pdf/?file=https://assets.gov.ie/260307/1a7ff966-44cf-4aed-b1a0-bc9c746d294a.pdf [Accessed: 2024-05-27].

[Dep24a] Department of Communications, Climate Action & Environment, Irish Government. Ireland's integrated national energy and climate plan 2021-2030, 2024. https://assets.gov.ie/static/documents/irelands-integrated-national-energy-and-climate-plan-2021-2030.pdf [Accessed: 2025-05-08].

[Dep24b] Department of Transport and Department for Infrastructure, Irish Government. All-island strategic rail review, 2024. https://assets.gov.ie/static/documents/final-report-of-the-all-island-strategic-rail-review.pdf[Accessed: 2025-05-27].

[Dep24c] Department of Transportation - Irish Government. Transport appraisal framework, appraisal guidelines for capital investments in transport, module 8 - detailed guidance on appraisal parameters, 2024. https://www.gov.ie/pdf/?file=https://assets.gov.ie/299622/e179ef16-9af0-41bf-af27-311d2f752935.pdf [Accessed: 2025-03-31].

[DS17] Rahul Dey and Fathi M. Salemt. Gate-variants of gated recurrent unit (gru) neural networks. *Midwest Symposium on Circuits and Systems*, 2017-August:1597–1600, 2017.

[ECLR19] Tomislav Erdelic, Tonči Carić, and Eduardo Lalla-Ruiz. A survey on the electric vehicle routing problem: Variants and solution approaches. *Journal of Advanced Transportation*, 2019, 2019.

[EG22] Adam N. Elmachtoub and Paul Grigas. Smart "predict, then optimize". *Management Science*, 68:9–26, 2022.

[Eir25] Eirgrid. System and renewable data summary report v16, 2025. `https://cms.eirgrid.ie/sites/default/files/publications/System-and-Renewable-Data-Summary-Report-v16.xlsx` [Accessed: 2025-11-24].

[EMSB22] Miquel Estrada, Josep Mensión, Miquel Salicrú, and Hugo Badia. Charging operations in battery electric bus systems considering fleet size variability along the service. *Transportation Research Part C: Emerging Technologies*, 138, 2022.

[Env25] Environmental Protection Agency. Latest emissions data, 2025. `https://www.epa.ie/our-services/monitoring--assessment/climate-change/ghg/latest-emissions-data/` [Accessed: 2025-05-09].

[ETFM19] Nader A El-Taweel, Hany EZ Farag, and Moataz Mohamed. Integrated utility-transit model for optimal configuration of battery electric bus systems. *IEEE Systems Journal*, 14:738–748, 2019.

[FAC23] James Fitzpatrick, Deepak Ajwani, and Paula Carroll. Learning to prune electric vehicle routing problems. *International Conference on Learning and Intelligent Optimization*, 14286 LNCS, 2023.

[FAMES23] Ahmed Foda, Hatem Abdelaty, Moataz Mohamed, and Ehab El-Saadany. A generic cost-utility-emission optimization for electric bus transit infrastructure planning and charging scheduling. *Energy*, 277, 2023.

[FDAH+22] Shahid Farah, Wood David A, Nisar Humaira, Zameer Aneela, and Eger Steffen. Short-term multi-hour ahead country-wide wind power prediction for germany using gated recurrent unit deep learning. *Renewable and Sustainable Energy Reviews*, 167, 2022.

[GC23] Julie Gallagher and Siobhán Clarke. Energy efficient route prediction for solar powered vehicles. *Green Energy and Intelligent Transportation*, 2, 2023.

[GIK23] K. Gkiotsalitis, C. Iliopoulou, and K. Kepaptsoglou. An exact approach for the multi-depot electric bus scheduling problem with time windows. *European Journal of Operational Research*, 306:189–206, 2023.

[GIOS14] Diarmuid Grimes, Georgiana Ifrim, Barry O'Sullivan, and Helmut Simonis. Analyzing the impact of electricity price forecasting on energy cost-aware scheduling. *Sustainable Computing: Informatics and Systems*, 4:276–291, 2014.

[GMH18] Marc Gallet, Tobias Massier, and Thomas Hamacher. Estimation of the energy demand of electric buses based on real-world data for large-scale public transport networks. *Applied Energy*, 230:344–356, 2018.

[GN23] Pranav Gairola and N. Nezamuddin. Optimization framework for integrated battery electric bus planning and charging scheduling. *Transportation Research Part D: Transport and Environment*, 118, 2023.

[HCEQ19] Carl H. Häll, Avishai (Avi) Ceder, Joakim Ekström, and Nils Hassan Quttineh. Adjustments of public transit operations planning process for the use of electric buses. *Journal of Intelligent Transportation Systems: Technology, Planning, and Operations*, 23:216–230, 2019.

[HDLP22] Hao Hu, Bo Du, Wei Liu, and Pascal Perez. A joint optimisation model for charger locating and electric bus charging scheduling considering opportunity fast charging and uncertainties. *Transportation Research Part C: Emerging Technologies*, 141, 2022.

[HLLL20] Shahram Hanifi, Xiaolei Liu, Zi Lin, and Saeid Lotfian. A critical review of wind power forecasting methods-past, present and future. *Energies*, 13, 2020.

[HLS22] Yi He, Zhaocai Liu, and Ziqi Song. Integrated charging infrastructure planning and charging scheduling for battery electric bus systems. *Transportation Research Part D: Transport and Environment*, 111, 2022.

[HLS23] Yi He, Zhaocai Liu, and Ziqi Song. Joint optimization of electric bus charging infrastructure, vehicle scheduling, and charging management. *Transportation Research Part D: Transport and Environment*, 117, 2023.

[HMMY21] Stephen P. Holland, Erin T. Mansur, Nicholas Z. Muller, and Andrew J. Yates. The environmental benefits of transportation electrification: Urban buses. *Energy Policy*, 148, 2021.

[HNBM23] Charitha Buddhika Heendeniya, Lorenzo Nespoli, Marco Belliardi, and Vasco Medici. Exploring trade-offs in public bus electrification under stochastic conditions. *Energy Informatics*, 6, 2023.

[HWCF21] Chenyu Hou, Jiawei Wu, Bin Cao, and Jing Fan. A deep-learning prediction model for imbalanced time series data forecasting. *Big Data Mining and Analytics*, 4:266–278, 2021.

[HZ24] Jiacai Han and Pan Zeng. Short-term power load forecasting based on hybrid feature extraction and parallel bilstm network. *Computers and Electrical Engineering*, 119:109631, 2024.

[IBM] IBM. Ibm ilog cplex optimization studio 22.1 documentation. `https://www.ibm.com/docs/en/icos/22.1.1?topic=concepts-branch-cut-in-cplex` [Accessed: 2025-03-15].

[IK19] Christina Iliopoulou and Konstantinos Kepaptsoglou. Integrated transit route network design and infrastructure planning for online electric vehicles. *Transportation Research Part D: Transport and Environment*, 77:178–197, 2019.

[IL19] Adrita Islam and Nicholas Lownes. When to go electric? a parallel bus fleet replacement study. *Transportation Research Part D: Transport and Environment*, 72:299–311, 2019.

[IOS12] Georgiana Ifrim, Barry O'Sullivan, and Helmut Simonis. Properties of energy-price forecasts for scheduling. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 7514 LNCS:957–972, 2012.

[Iri24] Irish Government. Climate action plan, 2024. `https://www.gov.ie/pdf/?file=https://assets.gov.ie/296414/7a06bae1-4c1c-4cdc-ac36-978e3119362e.pdf` [Accessed: 2025-03-20].

[JBZW22] Jinhua Ji, Yiming Bie, Ziling Zeng, and Linhong Wang. Trip energy consumption estimation for electric buses. *Communications in Transportation Research*, 2, 2022.

[JCCL12] Zeng Jin, Peng Chao, Li Chao, and Nie Lei. Passenger ticket theory and empirical research based on valuation of travel time. *World Automation Congress Proceedings*, None, 2012.

[JES19a] Amra Jahic, Mina Eskander, and Detlef Schulz. Charging schedule for load peak minimization on large-scale electric bus depots. *Applied Sciences (Switzerland)*, 9, 2019.

[JES19b] Amra Jahic, Mina Eskander, and Detlef Schulz. Preemptive vs. non-preemptive charging schedule for large-scale electric bus depots. *Proceedings of 2019 IEEE PES Innovative Smart Grid Technologies Europe, ISGT-Europe 2019*, None, 2019.

[JK19] Maroš Janovec and Michal Kohání. Exact approach to the electric bus fleet scheduling. *Transportation Research Procedia*, 40:1380–1387, 2019.

[JZ22] Mengyan Jiang and Yi Zhang. A branch-and-price algorithm for large-scale multidepot electric bus scheduling. *IEEE Transactions on Intelligent Transportation Systems*, 24:15355–15368, 2022.

[JZZ+18] Mengyan Jiang, Yi Zhang, Canrong Zhang, Kai Zhang, Gechu Zhang, and Ziyao Zhao. Operation and scheduling of pure electric buses under regular charging mode. *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, 2018-November:1894–1899, 2018.

[JZZ21] Mengyan Jiang, Yi Zhang, and Yi Zhang. Multi-depot electric bus scheduling considering operational constraint and partial charging: A case study in shenzhen, china. *Sustainability*, 14:255, 2021.

[Kar21] Sašo Karakatič. Optimizing nonlinear charging times of electric vehicle routing with genetic algorithm. *Expert Systems with Applications*, 164, 2021.

[KB21] Mouna Kchaou-Boujelben. Charging station location problem: A comprehensive review on models and solution approaches. *Transportation Research Part C: Emerging Technologies*, 132, 2021.

[KDKGD+21] Ashutosh Kumar Dubey, Abhishek Kumar, Vicente García-Díaz, Arpit Kumar Sharma, and Kishan Kanhaiya. Study and analysis of sarima and lstm in forecasting time series data. *Sustainable Energy Technologies and Assessments*, 47, 2021.

[KKIM21] Andreas Kanavos, Fotios Kounelis, Lazaros Iliadis, and Christos Makris. Deep learning models for forecasting aviation demand time series. *Neural Computing and Applications*, 33:16329–16343, 2021.

[KLCF20] Bwo Ren Ke, Yu Hsun Lin, Hong Zhang Chen, and Shyang Chyuan Fang. Battery charging and discharging scheduling with demand response for an electric bus public transportation system. *Sustainable Energy Technologies and Assessments*, 40, 2020.

[KLL21] Adam Kisvari, Zi Lin, and Xiaolei Liu. Wind power forecasting – a data-driven method along with gated recurrent neural network. *Renewable Energy*, 163:1895–1909, 2021.

[KSA+21] Matloob Khushi, Kamran Shaukat, Talha Mahboob Alam, Ibrahim A. Hameed, Shahadat Uddin, Suhuai Luo, Xiaoyan Yang, and Maranatha Consuelo Reyes. A comparative performance analysis of data resampling methods on imbalance medical data. *IEEE Access*, 9:109960–109975, 2021.

[LA20] Cesar Loaiza and Alejandro Arbelaez. Reference road network, 2020. Unpublished Work Package 2 Milestone 2 of SMARTeBuses project.

[LAC20] Tao Liu and Avishai (Avi) Ceder. Battery-electric transit vehicle scheduling with optimal number of stationary chargers. *Transportation Research Part C: Emerging Technologies*, 114:118–139, 2020.

[Laj18] Antti Lajunen. Lifecycle costs and charging requirements of electric buses with different charging methods. *Journal of Cleaner Production*, 172:56–67, 2018.

[LCC+23] Xin Lai, Junjie Chen, Quanwei Chen, Xuebing Han, Languang Lu, Haifeng Dai, and Yuejiu Zheng. Comprehensive assessment of carbon emissions and environmental impacts of sodium-ion batteries and lithium-ion batteries at the manufacturing stage. *Journal of Cleaner Production*, 423, 2023.

[LGL+21] Kai Liu, Hong Gao, Zhe Liang, Meng Zhao, and Cheng Li. Optimal charging strategy for large-scale electric buses considering resource constraints. *Transportation Research Part D: Transport and Environment*, 99, 2021.

[Li14] Jing Quan Li. Transit bus scheduling with limited energy. *Transportation Science*, 48:521–539, 2014.

[LK81] J. K. Lenstra and A. H.G.Rinnooy Kan. Complexity of vehicle routing and scheduling problems. *Networks*, 11:221–227, 1981.

[LLF21] Xiaolei Liu, Zi Lin, and Ziming Feng. Short-term offshore wind speed forecast by seasonal arima - a comparison against gru and lstm. *Energy*, 227, 2021.

[LLX19] Lu Li, Hong K. Lo, and Feng Xiao. Mixed bus fleet scheduling under range and refueling constraints. *Transportation Research Part C: Emerging Technologies*, 104:443–462, 2019.

[LMH+23] Lek Keng Lim, Zarina Ab Muis, Wai Shin Ho, Haslenda Hashim, and Cassendra Phun Chien Bong. Review of the energy forecasting and scheduling model for electric buses. *Energy*, 263, 2023.

[LPP20] Ioannis E. Livieris, Emmanuel Pintelas, and Panagiotis Pintelas. A cnn–lstm model for gold price time-series forecasting. *Neural Computing and Applications*, 32:17351–17360, 2020.

[LSPS21] Jinwoo Lee, Heeseung Shon, Ilia Papakonstantinou, and Sanghoon Son. Optimal fleet, battery, and charging infrastructure planning for reliable electric bus operations. *Transportation Research Part D: Transport and Environment*, 100, 2021.

[LWL+20] Xinghua Li, Tianzuo Wang, Lingjie Li, Feiyu Feng, Wei Wang, and Cheng Cheng. Joint optimization of regular charging electric bus transit network schedule and stationary charger deployment considering partial charging policy and time-of-use electricity prices. *Journal of Advanced Transportation*, 2020, 2020.

[LYP+24] Xiaohan Liu, Sonia Yeh, Patrick Plötz, Wenxi Ma, Feng Li, and Xiaolei Ma. Electric bus charging scheduling problem considering charging infrastructure integrated with solar photovoltaic and energy storage systems. *Transportation Research Part E: Logistics and Transportation Review*, 187, 2024.

[LZ21] Bryan Lim and Stefan Zohren. Time-series forecasting with deep learning: A survey. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 379, 2021.

[Man20] Anna Manowska. Using the lstm network to forecast the demand for electricity in poland. *Applied Sciences (Switzerland)*, 10:1–16, 2020.

[MBEH22] Teresa McGrath, Luke Blades, Juliana Early, and Andrew Harris. Uk battery electric bus operation: Examining battery degradation, carbon emissions and cost. *Transportation Research Part D: Transport and Environment*, 109, 2022.

[MBT17] Nuno Moniz, Paula Branco, and Luís Torgo. Resampling strategies for imbalanced time series forecasting. *International Journal of Data Science and Analytics*, 3:161–181, 2017.

[MCM21] Elianne Mora, Jenny Cifuentes, and Geovanny Marulanda. Short-term forecasting of wind energy: A comparison of deep learning frameworks. *Energies*, 14, 2021.

[MDSG20] Jayanta Mandi, Emir Demirović, Peter J. Stuckey, and Tias Guns. Smart predict-and-optimize for hard combinatorial optimization problems. *AAAI 2020 - 34th AAAI Conference on Artificial Intelligence*, None:1603–1610, 2020.

[MS22] Sidra Mehtab and Jaydip Sen. Analysis and forecasting of financial time series using cnn and lstm-based deep learning models. *Lecture Notes in Networks and Systems*, 302:405–423, 2022.

[MSM10] Rajesh Matai, Surya Prakash Singh, and Murari Lal Mittal. Traveling salesman problem: an overview of applications, formulations, and solution approaches. *Traveling salesman problem, theory and applications*, 1:1–25, 2010.

[MTHA22] Jônatas Augusto Manzolli, João Pedro F. Trovão, and Carlos Henggeler Antunes. Electric bus coordinated charging strategy considering v2g and battery degradation. *Energy*, 254, 2022.

[MTZ+20] Ahmadreza Moradipari, Nathaniel Tucker, Tuo Zhang, Gustavo Cezar, and Mahnoosh Alizadeh. Mobility-aware smart charging of electric bus fleets. *IEEE Power and Energy Society General Meeting*, 2020-August, 2020.

[MXY+23] Anbo Meng, Zikang Xian, Hao Yin, Jianqiang Luo, Xiaolin Wang, Haitao Zhang, Jiayu Rong, Chen Li, Zhenbo Wu, Zhifeng Xie, Zhan Zhang, Chenen Wang, and Yingjun Chen. A novel network training approach for solving sample imbalance problem in wind power prediction. *Energy Conversion and Management*, 283, 2023.

[OBL+19] Tom O'Malley, Elie Bursztein, James Long, François Chollet, Haifeng Jin, Luca Invernizzi, et al. Kerastuner. `https://github.com/keras-team/keras-tuner`, 2019.

[OCPV24] Ciaran O'Connor, Joseph Collins, Steven Prestwich, and Andrea Visentin. Electricity price forecasting in the irish balancing market. *Energy Strategy Reviews*, 54, 2024.

[OF20] Emanuel Florentin Olariu and Cristian Frasinaru. Multiple-depot vehicle scheduling problem heuristics. *Procedia Computer Science*, 176:241–250, 2020.

[Oh11] Sang Hoon Oh. Error back-propagation algorithm for classification of imbalanced data. *Neurocomputing*, 74:1058–1061, 2011.

[OK20] Nils Olsen and Natalia Kliewer. Scheduling electric buses in public transport: Modeling of the charging process and analysis of assumptions. *Logistics Research*, 13, 2020.

[OKW22] Nils Olsen, Natalia Kliewer, and Lena Wolbeck. A study on flow decomposition methods for scheduling of electric buses in public transport based on aggregated time–space network models. *Central European Journal of Operations Research*, 30:883–919, 2022.

[ON15] Keiron O'shea and Ryan Nash. An introduction to convolutional neural networks. *arXiv preprint arXiv:1511.08458*, 2015.

[ORG+22] Vera O'Riordan, Fionn Rogan, Brian O Gallachóir, Tomás Mac Uidhir, and Hannah Daly. How and why we travel – mobility demand and emissions from passenger transport. *Transportation Research Part D: Transport and Environment*, 104, 2022.

[PDH+21] Shyam S.G. Perumal, Twan Dollevoet, Dennis Huisman, Richard M. Lusby, Jesper Larsen, and Morten Riis. Solution approaches for integrated vehicle and crew scheduling with electric buses. *Computers and Operations Research*, 132, 2021.

[Pin19] Robert S. Pindyck. The social cost of carbon revisited. *Journal of Environmental Economics and Management*, 94:140–160, 2019.

[PLL22] Shyam S.G. Perumal, Richard M. Lusby, and Jesper Larsen. Electric bus planning & scheduling: A review of related problems and methodologies. *European Journal of Operational Research*, 301:395–413, 2022.

[RCS+21] A. Rong, S. Chen, D. Shi, M. Zhang, and C. Wang. A review on electric bus charging scheduling from viewpoints of vehicle scheduling. *2021 IEEE International Conference on Industrial Engineering and Engineering Management, IEEM 2021*, None:1–5, 2021.

[RKW15] Josephine Reuer, Natalia Kliewer, and Lena Wolbeck. The electric vehicle scheduling problem: A study on time-space network based and heuristic solution. *Proceedings of the Conference on Advanced Systems in Public Transport (CASPT)*, pages 1–15, 2015.

[RM20] Rita P. Ribeiro and Nuno Moniz. Imbalanced regression and extreme value prediction. *Machine Learning*, 109:1803–1835, 2020.

[RN16] Stuart J Russell and Peter Norvig. *Artificial intelligence: a modern approach*. pearson, 2016.

[RNI+22] S. Rafique, M. S.H. Nizami, U. B. Irshad, M. J. Hossain, and S. C. Mukhopadhyay. A two-stage multi-objective stochastic optimization strategy to minimize cost for electric bus depot operators. *Journal of Cleaner Production*, 332, 2022.

[Ros24] Andrea Rossi. *Modelling uncertainty in cloud workload forecasting: a Hybrid Bayesian Neural Network approach*. PhD thesis, University College Cork, 2024.

[RPDV20] Marco Rinaldi, Erika Picarelli, Andrea D'Ariano, and Francesco Viti. Mixed-fleet single-terminal bus scheduling problem: Modelling, solution scheme and potential applications. *Omega (United Kingdom)*, 96, 2020.

[RRHK20] Matthias Rupp, Christian Rieke, Nils Handschuh, and Isabel Kuperjans. Economic and ecological optimization of electric bus charging considering variable electricity prices and co<inf>2eq</inf> intensities. *Transportation Research Part D: Transport and Environment*, 81, 2020.

[RVBW06] Francesca Rossi, Peter Van Beek, and Toby Walsh. *Handbook of constraint programming*. Elsevier, 2006.

[RvdHLS18] Matthias Rogge, Evelien van der Hurk, Allan Larsen, and Dirk Uwe Sauer. Electric bus fleet size and mix problem with optimization of charging infrastructure. *Applied Energy*, 211:282–295, 2018.

[SB22] Sharath Kariya Shetty and Rajesh Buktar. A comparative study of automobile sales forecasting with arima, sarima and deep learning lstm model. *International Journal of Advanced Operations Management*, 14:366–387, 2022.

[SBA22] Uppala Meena Sirisha, Manjula C. Belavagi, and Girija Attigeri. Profit prediction using arima, sarima and lstm models in time series forecasting: A comparison. *IEEE Access*, 10:124715–124727, 2022.

[SCC+22] Yi Wei Sung, James C. Chu, Yao Jen Chang, Jou Chun Yeh, and Yen Hung Chou. Optimizing mix of heterogeneous buses and chargers in electric bus scheduling problems. *Simulation Modelling Practice and Theory*, 119, 2022.

[Sco21] Scottish Government. Energy statistics for scotland q4 2020 figures, 2021. `https://www.gov.scot/binaries/content/documents/govscot/publications/statistics/2018/10/quarterly-energy-statistics-bulletins/documents/energy-statistics-summary---march-2021/energy-statistics-summary---march-2021/govscot%3Adocument/Scotland%2BEnergy%2BStatistics%2BQ4%2B2020.pdf` [Accessed: 2025-03-20].

[SD20] Matteo Sangiorgio and Fabio Dercole. Robustness of lstm neural networks for multi-step forecasting of chaotic time series. *Chaos, Solitons and Fractals*, 139, 2020.

[SDMB+21] A. Shobana Devi, G. Maragatham, K. Boopathi, M.C. Lavanya, and R. Saranya. Long-term wind speed forecasting—a review. *Lecture Notes in Networks and Systems*, 130:79 – 99, 2021.

[SG15] Remy Spliet and Adriana F. Gabor. The time window assignment vehicle routing problem. *Transportation Science*, 49:721–731, 2015.

[SIKF24] Osamah Siddiqui, Haris Ishaq, Daniyal Ahmed Khan, and Hesham Fazel. Social cost-benefit analysis of different types of buses for sustainable public transportation. *Journal of Cleaner Production*, 438, 2024.

[SIMC13] Arunesh Kumar Singh, S Khatoon Ibraheem, Md Muazzam, and DK Chaturvedi. An overview of electricity demand forecasting techniques. *Network and complex systems*, 3:38–48, 2013.

[SMGS19] Jakub Szczesniak, Tobias Massier, Marc Gallet, and Anurag Sharma. Optimal electric bus charging scheduling for local balancing of fluctuations in pv generation. *2019 IEEE PES Innovative Smart Grid Technologies Asia, ISGT 2019*, None:2799–2804, 2019.

[SO17] Ons Sassi and Ammar Oulamara. Electric vehicle scheduling and optimal charging problem: complexity, exact and heuristic approaches. *International Journal of Production Research*, 55:519–535, 2017.

[SS12] L. Suganthi and Anand A. Samuel. Energy models for demand forecasting - a review. *Renewable and Sustainable Energy Reviews*, 16:1223–1240, 2012.

[SSHSP23] Daan R. Scheepens, Irene Schicker, Katerina Hlaváčková-Schindler, and Claudia Plant. Adapting a deep convolutional rnn model with imbalanced regression loss for improved spatio-temporal forecasting of extreme wind speed events in the short to medium range. *Geoscientific Model Development*, 16:251–270, 2023.

[SSS00] Robert H Shumway, David S Stoffer, and David S Stoffer. *Time series analysis and its applications*, volume 3. Springer, 2000.

[STLS22] Shuai Shao, Zhijia Tan, Zhiyuan Liu, and Wenlong Shang. Balancing the ghg emissions and operational costs for a mixed fleet of electric buses and diesel buses. *Applied Energy*, 328, 2022.

[Sus24] Sustainable Energy Authority Ireland. Energy in ireland 2024 report, 2024. `https://www.seai.ie/sites/default/files/publications/energy-in-ireland-2024.pdf` [Accessed: 2025-05-14].

[Sus25] Sustainable Energy Authority Ireland. Battery storage: Manual of consenting procedures, 2025. `https://singlepointofcontact.seai.ie/renewable-energy-guides/Battery-Storage-Manual-SEAI.pdf` [Accessed: 2025-11-24].

[SVG+17] Sreelekshmy Selvin, R. Vinayakumar, E. A. Gopalakrishnan, Vijay Krishna Menon, and K. P. Soman. Stock price prediction using lstm, rnn and cnn-sliding window model. *2017 International Conference on Advances in Computing, Communications and Informatics, ICACCI 2017*, 2017-January:1643–1647, 2017.

[SZS+24] Yu Sun, Qibo Zhou, Li Sun, Liping Sun, Jichuan Kang, and He Li. Cnn–lstm–am: A power prediction model for offshore wind turbines. *Ocean Engineering*, 301, 2024.

[TCW25] Zhirui Tian, Yujie Chen, and Guangyu Wang. Enhancing pv power forecasting accuracy through nonlinear weather correction based on multi-task learning. *Applied Energy*, 386:125525, 2025.

[THS+21] José F. Torres, Dalil Hadjout, Abderrazak Sebaa, Francisco Martínez-Álvarez, and Alicia Troncoso. Deep learning for time series forecasting: A survey. *Big Data*, 9:3–21, 2021.

[TLH19] Xindi Tang, Xi Lin, and Fang He. Robust scheduling strategies of electric buses under stochastic traffic conditions. *Transportation Research Part C: Emerging Technologies*, 105:163–182, 2019.

[TR09] Luis Torgo and Rita Ribeiro. Precision and recall for regression. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 5808 LNAI:332–346, 2009.

[Tra25] Transport Infrastructure Ireland. Luas cork, 2025. `https://www.luascork.ie/` [Accessed: 2025-05-27].

[TRPB13] Luís Torgo, Rita P. Ribeiro, Bernhard Pfahringer, and Paula Branco. Smote for regression. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 8154 LNAI:378–389, 2013.

[vOBvN12] N. van Oort, J. W. Boterman, and R. van Nes. The impact of scheduling on service reliability: Trip-time determination and holding points in long-headway services. *Public Transport*, 4:39–56, 2012.

[WCdJ16] Mark Wardman, V. Phani K. Chintakayala, and Gerard de Jong. Values of travel time in europe: Review and meta-analysis. *Transportation Research Part A: Policy and Practice*, 94:93–111, 2016.

[WGH11] Xiaochen Wang, Peng Guo, and Xiaobin Huang. A review of wind power forecasting models. *Energy Procedia*, 12:770–778, 2011.

[WLC18] Yusen Wang, Wenlong Liao, and Yuqing Chang. Gated recurrent unit network-based short-term photovoltaic forecasting. *Energies*, 11, 2018.

[WLL22] Yongxing Wang, Feixiong Liao, and Chaoru Lu. Integrated optimization of charger deployment and fleet scheduling for battery electric buses. *Transportation Research Part D: Transport and Environment*, 109, 2022.

[WLLJ22] Weitiao Wu, Yue Lin, Ronghui Liu, and Wenzhou Jin. The multi-depot electric vehicle scheduling problem with power grid characteristics. *Transportation Research Part B: Methodological*, 155:322–347, 2022.

[WLOKF18] Ran Wei, Xiaoyue Liu, Yi Ou, and S. Kiavash Fayyaz. Optimizing the spatio-temporal deployment of battery electric bus system. *Journal of Transport Geography*, 68:160–168, 2018.

[WLR+16] M. Wen, E. Linde, S. Ropke, P. Mirchandani, and A. Larsen. An adaptive large neighborhood search heuristic for the electric vehicle scheduling problem. *Computers and Operations Research*, 76:73–83, 2016.

[Wol20] Laurence A Wolsey. *Integer programming*. John Wiley & Sons, 2020.

[WS17] Fei Wu and Ramteen Sioshansi. A two-stage stochastic optimization model for scheduling electric vehicle charging loads to relieve distribution-system constraints. *Transportation Research Part B: Methodological*, 102:55–82, 2017.

[WSXW23] Xuedong Wang, Ziqi Song, Hongli Xu, and Hua Wang. En-route fast charging infrastructure planning and scheduling for battery electric bus systems. *Transportation Research Part D: Transport and Environment*, 117, 2023.

[WWHG24] Mengqin Wang, Yanling Wei, Xueliang Huang, and Shan Gao. An end-to-end deep reinforcement learning framework for electric vehicle routing problem. *IEEE Internet of Things Journal*, 2024.

[WZL+21] Yun Wang, Runmin Zou, Fang Liu, Lingjun Zhang, and Qianyi Liu. A review of wind speed and wind power forecasting with deep neural networks. *Applied Energy*, 304, 2021.

[XYL23] Xiaoming Xu, Yanhong Yu, and Jiancheng Long. Integrated electric bus timetabling and scheduling problem. *Transportation Research Part C: Emerging Technologies*, 149, 2023.

[YCW+22] Cheng Hong Yang, Bo Hong Chen, Chih Hsien Wu, Kuo Chang Chen, and Li Yeh Chuang. Deep learning for forecasting electricity demand in taiwan. *Mathematics*, 10, 2022.

[YLLY20] Enjian Yao, Tong Liu, Tianwei Lu, and Yang Yang. Optimization of electric vehicle scheduling with multiple vehicle types in public transport. *Sustainable Cities and Society*, 52, 2020.

[YY21] Şule Yıldırım and Barış Yıldız. Electric bus fleet composition and scheduling. *Transportation Research Part C: Emerging Technologies*, 129, 2021.

[YZC+21] Yuzhe Yang, Kaiwen Zha, Ying Cong Chen, Hao Wang, and Dina Katabi. Delving into deep imbalanced regression. *Proceedings of Machine Learning Research*, 139:11842–11851, 2021.

[ZAdAR22] Letícia A.L. Zaneti, Nataly Bañol Arias, Madson C. de Almeida, and Marcos J. Rider. Sustainable charging schedule of electric buses in a university campus: A rolling horizon approach. *Renewable and Sustainable Energy Reviews*, 161, 2022.

[ZCL24] Feifeng Zheng, Runfeng Cao, and Ming Liu. Stochastic fast charging scheduling of battery electric buses with energy storage systems design. *Computers and Industrial Engineering*, 191, 2024.

[ZCZX23] Ailing Zeng, Muxi Chen, Lei Zhang, and Qiang Xu. Are transformers effective for time series forecasting? *Proceedings of the AAAI conference on artificial intelligence*, 37, 2023.

[ZL21] Peng Zhuang and Hao Liang. Stochastic energy management of electric bus charging stations with renewable energy integration and b2g capabilities. *IEEE Transactions on Sustainable Energy*, 12:1206–1216, 2021.

[ZLC+17] Bendong Zhao, Huanzhang Lu, Shangfeng Chen, Junliang Liu, and Dongya Wu. Convolutional neural networks for time series classification. *Journal of Systems Engineering and Electronics*, 28:162–169, 2017.

[ZLZ+22] Aijia Zhang, Tiezhu Li, Yue Zheng, Xuefeng Li, Muhammad Ghazanfar Abdullah, and Changyin Dong. Mixed electric bus fleet scheduling problem with partial mixed-route and partial recharging. *International Journal of Sustainable Transportation*, 16:73–83, 2022.

[ZMO22] Yu Zhou, Qiang Meng, and Ghim Ping Ong. Electric bus charging scheduling for a single public transport route considering nonlinear charging profile and battery degradation effect. *Transportation Research Part B: Methodological*, 159:49–75, 2022.

[ZMOW24] Yu Zhou, Qiang Meng, Ghim Ping Ong, and Hua Wang. Electric bus charging scheduling on a bus network. *Transportation Research Part C: Emerging Technologies*, 161, 2024.

[ZWQ21] Le Zhang, Shuaian Wang, and Xiaobo Qu. Optimal electric bus fleet scheduling considering battery degradation and non-linear charging profile. *Transportation Research Part E: Logistics and Transportation Review*, 154, 2021.

[ZWQ22] Ziling Zeng, Shuaian Wang, and Xiaobo Qu. On the role of battery degradation in en-route charge scheduling for an electric bus system. *Transportation Research Part E: Logistics and Transportation Review*, 161, 2022.

[ZWQ24] Ziling Zeng, Tingsong Wang, and Xiaobo Qu. En-route charge scheduling for an electric bus network: Stochasticity and real-world practice. *Transportation Research Part E: Logistics and Transportation Review*, 185, 2024.

[ZWW+24] Yu Zhou, Hua Wang, Yun Wang, Bin Yu, and Tianpei Tang. Charging facility planning and scheduling problems for battery electric bus systems: A comprehensive review. *Transportation Research Part E: Logistics and Transportation Review*, 183, 2024.

[ZWWL22] Yu Zhou, Hua Wang, Yun Wang, and Rui Li. Robust optimization for integrated planning of electric-bus charger deployment and charging scheduling. *Transportation Research Part D: Transport and Environment*, 110, 2022.

[ZXZL20] Guang Jing Zhou, Dong Fan Xie, Xiao Mei Zhao, and Chaoru Lu. Collaborative optimization of vehicle and charging scheduling for a bus fleet mixed with electric and traditional buses. *IEEE Access*, 8:8056–8072, 2020.

[ZZZZ24] Liwan Zhou, Di Zhang, Le Zhang, and Jizhong Zhu. Dynamic gust detection and conditional sequence modeling for ultra-short-term wind speed prediction. *Electronics (Switzerland)*, 13, 2024.