Sustainable and data-driven airport operations

Optimisation models and machine learning approaches

Zoutendijk, M.

DOI
10.4233/uuid:1002070b-c29b-43e6-9b8a-095292cbff22

Publication date
2024

Document Version
Final published version

Citation (APA)
Zoutendijk, M. (2024). *Sustainable and data-driven airport operations: Optimisation models and machine learning approaches*. [Dissertation (TU Delft), Delft University of Technology].
https://doi.org/10.4233/uuid:1002070b-c29b-43e6-9b8a-095292cbff22

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

SUSTAINABLE AND DATA-DRIVEN AIRPORT OPERATIONS

OPTIMISATION MODELS AND MACHINE LEARNING APPROACHES

SUSTAINABLE AND DATA-DRIVEN AIRPORT OPERATIONS

OPTIMISATION MODELS AND MACHINE LEARNING APPROACHES

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof. dr. ir. T.H.J.J. van der Hagen,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen op donderdag 30 mei 2024 om 10:00 uur

door

Micha ZOUTENDIJK

Ingenieur Toegepaste Wiskunde,
Technische Universiteit Delft, Nederland,
geboren te Dordrecht, Nederland.

Dit proefschrift is goedgekeurd door de promotoren.

Samenstelling promotiecommissie:

Rector Magnificus,	voorzitter
Prof. dr. ir. J. M. Hoekstra,	Technische Universiteit Delft, promotor
Dr. M. A. Mitici,	Universiteit Utrecht, copromotor
Onafhankelijke leden:	
Prof. Dr. -Ing. P. Hecker	Technische Universität Braunschweig, Duitsland
Prof. Dr. -Ing. Habil. M. Schultz	Universität der Bundeswehr München, Duitsland
Prof. dr. D. Delahaye	Ecole Nationale de l'Aviation Civile, Frankrijk
Prof. dr. ir. M. Mulder	Technische Universiteit Delft
Dr. P. Wei	George Washington University, Verenigde Staten

This research was partly funded by the European Regional Development Fund (ERDF) with Grant No. KVW-00235.

Keywords: Airport Operations, Sustainability, Electric Taxiing, Fleet Assignment, Flight-to-Gate Assignment, Machine Learning, Probabilistic Prediction

Printed by: Ipskamp Printing

Front & Back: Open AI

ISBN 000-00-0000-000-0

An electronic version of this dissertation is available at
http://repository.tudelft.nl/.

CONTENTS

ACKNOWLEDGEMENTS

This dissertation marks the end of four years of research into improving the sustainability and efficiency of airport operations. I have enjoyed being introduced into this part of the aerospace world, which fortunately has a big dash of mathematics mixed in. During my time as a PhD candidate I've been granted the opportunity to meet many interesting, knowledgeable and inspirational people, to come in contact with airports, companies and institutions, and to disseminate research work at conferences and in journals. Now it is time to take these experiences with me and find a new challenge.

Naturally, this thesis would not have been possible without the excellent supervision of dr. Mihaela Mitici, whom I thank for her essential advice and suggestions during our meetings, for taking the time to review my writings, but mostly for guiding me on the path towards becoming an independent researcher. I thank my promotor prof. Jacco Hoekstra for helping me find the right focus in research, paper writing and working life in general. What I've learned from both of you will form an excellent foundation for any research I will do in the future.

I would like to thank my other colleagues, who have been a source of fun and inspiration. Simon van Oosterom, Ingeborg de Pater and Juseong Lee, thank you for being part of this small group helping each other with navigating the PhD process, but also following courses together, and going to operations research events and other gatherings. Hao Ma, Marie Bieber, Iordanis Tseremoglou, Ilias Parmaksizoglou and Haonan Li, thanks for making room 4.15 the most fun room of the section, and for all the good conversations. Marie Bieber (and Matthäus), Thomas Pioger and Malte von der Burg (and Carolina), thanks for creating our Austrian adventures. Thanks to all other ATO PhDs/PostDocs, Qichen Deng, Gülçin Ermiş, Chengpeng Jiang, Raissa Li, Federico Morlupo, Mahdi Noorafza, Prashant, Wenhua Qu, Matthieu Vert, Sahand Mohammadi Ziabari, and all PhD visitors, and of course the ATO professors and staff Marcia Baptista, Henk Blom, Alessandro Bombelli, Viswanath Dhanisetty, Marta Ribeiro, Paul Roling, Bruno Santos, Alexei Sharpanskykh and Nathalie Zoet. In particular I would like to thank Elise Bavelaar for helping me navigate the ATL project.

Later during the PhD I had the privilege to be received into another section as well. I would like to thank those in the C&S ATM section, Andrei Badea, Jan Groot, Fazlur Rahman, Marta Ribeiro, Esther Roosenbrand, Aidana Tassanbi, Andres Morfin Veytia, Sasha Vlaskin and visiting PhDs, as well as professors and staff Joost Ellerbroek, Bertine Markus and Junzi Sun for many interesting conversations and meetings, and activities such as offsites, birthdays and parties.

During my work, I have had the opportunity to cosupervise several MSc students: Rik Hendrickx, Robin Vervaat, Sarah Dutrieux and Bas Tindemans. It was an interesting experience to be able to help you with your research, which inspired me about mine. Thanks in particular to Rik, whose thesis forms the basis of Chapter 6 of this dissertation.

The dissertation not only marks the end of four years as a PhD candidate, but in fact of almost ten years at the TU Delft. I would like to thank everyone working at the TU Delft who has made this journey and the corresponding experiences and opportunities possible during this time, in particular my BSc and MSc thesis supervisors Wim van Horssen and Martin van Gijzen. I would like to thank my friends of Huiskerst Enzo for all the fun and our nice gatherings in the study period and the years since. I would like to thank my former roommates Massimo Achterberg and Marco Rozendaal for our time together in Rijswijk.

I would like to thank my parents and brother for their support and encouragement over the years. Last, I thank my beloved Jenny for being there for me throughout the PhD adventure.

Mike Zoutendijk
December 2023

SUMMARY

The aerospace industry annually provides transport for billions of passengers along trillions of kilometers. The industry is continuously aiming to provide these services in a more efficient and sustainable way. One possibility is to consider improving airside airport operations, both current types and those expected in the near future. Scheduling airport operations requires taking into account flight planning, airport layout, routing requirements and personnel planning. Current operational planning is characterised by application of linear programming tools for strategic planning, and manual adjustment for adaptive planning.

This dissertation aims to develop data-driven optimisation models, to increase the efficiency and sustainability of various airside airport operations, and to apply these models to airport case studies. The focus is first put on external electric taxiing, a new taxiing technique using electric towing vehicles (ETVs) to tow aircraft from gates to runways and vice versa. Many airports are considering to implement this technique, as it offers a large improvement in reducing their greenhouse gas emissions, noise levels and air pollution, which is an improvement for passengers, airport personnel, and local residents.

The first goal is to create a comprehensive overview of the operational aspects of external electric taxiing, by reviewing existing research work and industry sources. This overview includes the expected specifications of ETVs and the future procedures for electric taxiing movement. Electric taxiing introduces a new airside operation to the airport: ETV-to-aircraft scheduling. Studies on this new operation, as well as on vehicle routing, vehicle fleet sizing and battery charging optimisation models, which are needed for electric taxiing, are reviewed. The overview also includes the remaining research challenges to achieve large-scale ETV implementation in the next few decades.

The second goal is to develop an optimisation model to perform ETV-to-aircraft scheduling that takes into account realistic airport circumstances. A more efficient ETV-to-aircraft schedule, which allows more aircraft to be towed by an ETV fleet, will reduce airport emissions more. Some studies have already proposed ETV-to-aircraft scheduling models. However, they do not include all elements needed to make the model realistic and comprehensive, such as routing with conflict and collision avoidance, ETV charging and discharging, and airport surface movement specifications. Two more elements are added to this list in this work: airport electricity capacity and achieving a time-efficient model. Two models are developed for full-day ETV-to-aircraft scheduling, a Mixed-Integer Linear Programming (MILP) model and an Adaptive Large Neighbourhood Search (ALNS) model. Both models limit ETV charging to the electricity capacity of the airport. The ALNS model is able to create near-optimal full-day schedules for large fleet sizes within a few hours, for a large airport case study. The ALNS model is tested with various daily electricity capacity profiles, which shows the necessity of night charging and the effects of increasing amounts of charging during the day.

The third goal is to develop an optimisation approach to retain efficiency for electric taxiing in a real-time situation. The models developed for the second goal are applicable for strategic scheduling. During operation, disruptions to the strategic schedule will occur, and adaptive scheduling is required to continue operation. In this dissertation both a strategic and disrupted scheduling model are developed. The disrupted model reassigns delayed aircraft to ETV, aiming to minimize the changes to the original schedule. The model is used to create an adaptive schedule in a large airport case study using historical flight data. At the start of every half hour period, the disruptions due to flight delays of the next period are incorporated in a new schedule. The results show the efficacy of the disrupted model in minimizing schedule changes, which does not come at the expense of emission savings.

In addition to electric taxiing, this dissertation focuses on improving the efficiency and robustness of airside operations by predicting airport disruptions, to avoid additional use of resources and to provide a better service. Where the previous part consists of using models to react to flight delays, operations can also be improved by predicting them in advance. In existing works, delays are predicted by classification or as point prediction. In this dissertation, probabilistic prediction is applied to flight delay, using two machine learning algorithms: Mixture Density Networks and Random Forests Regression. In addition, metrics suited to probabilistic prediction are developed and used to evaluate the algorithm performance. In a small airport case study, the algorithms are shown to be able to predict delays within a Continuous Ranked Probability Score (CRPS) of eleven minutes.

The probabilistic prediction algorithms generate estimated delay distributions, which include extended uncertainty information. To illustrate the utility of the predictions for airport operations, they are applied in a probabilistic model aimed to increase the robustness of the flight-to-gate assignment problem. The proposed model is shown to reduce the number of gate-conflicted aircraft by up to 74% when compared to a deterministic flight-to-gate assignment model. The robustness of the assignment can be controlled with a model parameter.

Another method for predicting flight delays is binary classification, which is popular in literature. However, when posed as a binary problem, flight delay and also flight cancellation prediction suffer from a large data imbalance. This causes a distorted view when using metrics such as accuracy. This dissertation develops a systematic approach to binary prediction with imbalanced data, by considering a range of sampling ratios and various sampling techniques. Two machine learning algorithms are applied to a small airport historical flight dataset. The results underline the need to investigate the influence of varying data imbalance ratios on the performance of classification algorithms in various metrics.

Throughout this dissertation, the focus has been on improving the sustainability and efficiency of airport operations through data-driven approaches. These approaches include MILP models, heuristics and machine learning models. The developed models provide support for airport planners to improve current and future scheduling tasks. However, it remains future work to apply similar techniques to other airside operations and to further improve the realism and real-time usability of the current models. In addition, airports' spatial planners, air traffic controllers and ETV developers will play a

critical role in the further development and implementation of electric taxiing. Overall, this dissertation forms a starting point for airport planners aiming to use data-driven methods to improve the sustainability and efficiency of airports, to ensure more durable and reliable air transportation services.

SAMENVATTING

De luchtvaartindustrie verzorgt jaarlijks het vervoer van miljarden passagiers over afstanden van biljoenen kilometers. De industrie heeft zich ten doel gesteld om deze diensten op een efficiëntere en meer duurzame manier te leveren. Een van de mogelijkheden daartoe is het verbeteren van zowel bestaande als toekomstige luchthavenoperaties. Om deze te roosteren moet er rekening gehouden worden met de vluchtplanning, de ruimtelijke indeling van de luchthaven, eisen voor de routes en de personeelsplanning. De huidige operationele planning wordt gekenmerkt door het gebruik van lineaire programmeringstools voor de strategische planning, en van handmatige aanpassingen voor de adaptieve planning.

Dit proefschrift is gericht op het ontwikkelen van datagedreven optimalisatiemodellen, om de efficiëntie en duurzaamheid van verschillende luchthavenoperaties te verhogen, en om deze modellen toe te passen op casestudies van luchthavens. De focus wordt eerst gelegd op extern elektrisch taxiën, een nieuwe taxitechniek die gebruikt maakt van elektrische sleepvoertuigen (ETVs) om vliegtuigen van gates naar landingsbanen en andersom te slepen. Veel luchthavens overwegen deze techniek te implementeren, omdat het een grote bijdrage kan leveren aan de vermindering van de uitstoot van broeikasgassen, de geluidsoverlast en de luchtvervuiling, wat een verbetering is voor de passagiers, het luchthavenpersoneel, en de omwonenden.

Het eerste doel is om een uitgebreid overzicht van de operationale aspecten van extern elektrisch taxiën te maken, met behulp van wetenschappelijke publicaties en bronnen uit de industrie. Dit overzicht omvat de verwachte specificaties van ETVs en de toekomstige procedures voor de elektrische taxibewegingen. Voor de uitvoering van elektrisch taxiën is een ETV-naar-vliegtuig-roostering nodig. Publicaties worden bestudeerd aangaande deze nieuwe operatie, alsook aangaande voertuigroutering, de grootte van voertuigvloten, en optimalisatiemodellen voor het opladen van accu's, benodigd voor elektrisch taxiën. Het overzicht bevat ook de onderzoeksuitdagingen die opgelost dienen te worden om op grote schaal ETV-implementatie te bereiken in de komende decennia.

Het tweede doel is om een optimalisatiemodel te ontwikkelen dat een ETV-naar-vliegtuig-roostering kan maken en realistische luchthavenomstandigheden in acht neemt. Een efficiëntere ETV-naar-vliegtuig-roostering, waarin meer vliegtuigen gesleept kunnen worden door een ETV-vloot, zal de luchthavenemissies meer verminderen. In enkele onderzoeken zijn al modellen voor ETV-naar-vliegtuig-roostering voorgesteld. Deze bevatten echter niet alle elementen die nodig zijn om het model realistisch en volledig te maken, zoals routering met conflict- en botsing-vermijding, het opladen en ontladen van ETVs, en de specificaties van het verkeer op de luchthaven. In dit proefschrift worden hier twee elementen aan toegevoegd: de elektriciteitscapaciteit van de luchthaven, en het ontwikkelen van een tijdsefficiënt model. Er worden twee modellen voorgesteld om ETV-naar-vliegtuig-roostering voor een volledige dag te maken: een Mixed-Integer Linear Programming (MILP) model en een Adaptive Large Neighbour-

hood Search (ALNS) model. Beide modellen limiteren het opladen van ETVs tot de elektriciteitscapaciteit van de luchthaven. Het ALNS model is in staat om een bijna-optimale roostering voor elektrisch taxiën op een grote luchthaven te maken voor een volledige dag, binnen enkele uren. Het ALNS model is getest met verscheidene dagprofielen van electriciteitscapaciteit. De resultaten tonen de noodzaak van opladen gedurende de nacht aan, en illustreren de effecten van een toenemende mate van opladen gedurende de dag.

Het derde doel is om een optimalisatiemodel te ontwikkelen waarmee de efficiëntie van elektrisch taxiën wordt behouden tijdens de uitvoering van het rooster. De modellen die voor het tweede doel werden ontwikkeld zijn toepasbaar voor strategisch roosteren. Tijdens de uitvoering van het rooster kunnen verstoringen optreden, waardoor adaptief roosteren nodig is om met de uitvoering door te gaan. In dit proefschrift worden zowel een strategisch als verstoord roosteringsmodel ontwikkeld. In het verstoorde model worden vertraagde vliegtuigen waar nodig opnieuw aan een ETV toegewezen, en daarbij worden de veranderingen aan het originele rooster geminimaliseerd. Het model wordt gebruikt om een adaptatief rooster te maken voor een casestudie van een grote luchthaven met historische vluchtdata. Aan het begin van elke periode van een half uur worden de verstoringen door vluchtvertragingen van het komende halfuur verwerkt in een nieuw rooster. De resultaten laten de doeltreffendheid van het verstoorde model zien, waarin roosterveranderingen worden geminimaliseerd, zonder dat dit ten koste gaat van het verminderen van emissies.

Naast elektrisch taxiën focust dit proefschrift op het verbeteren van de efficiëntie en robuustheid van luchthavenoperaties, door het voorspellen van vluchtvertragingen, om het gebruik van extra middelen te voorkomen, en om betere service te leveren. In het vorige onderdeel werden modellen gebruikt om te reageren op vluchtvertragingen, maar operaties kunnen ook worden verbeterd door de vertragingen van tevoren te voorspellen. In de literatuur worden vertragingen veelal voorspeld door classificatie of als een puntvoorspelling. In dit proefschrift worden probabilistische voorspellingen toegepast op vluchtvertragingen, door het gebruik van twee machine learning algoritmen: Mixture Density Networks en Random Forests Regression. Daarnaast worden metrieken ontwikkeld die geschikt zijn voor probabilistische voorspellingen, en deze worden gebruikt om de prestaties van de algoritmes te evalueren. Door middel van een casestudie van een kleine luchthaven wordt aangetoond dat de algoritmes in staat zijn om vluchtvertragingen te voorspellen binnen een Continuous Ranked Probability Score (CRPS) van elf minuten.

De probabilistische voorspellingsalgoritmen genereren geschatte kansverdelingen van vertraging, waarbinnen uitgebreidere onzekerheidsinformatie omvat is dan bij slechts een voorspelde waarde. Om het nut van de voorspellingen voor luchtvaartoperaties aan te tonen, worden deze toegepast in een probabilistisch model, waarmee de robuustheid van het vlucht-naar-gate-toewijzingsprobleem kan worden verhoogd. Er wordt aangetoond dat het voorgestelde model het aantal vliegtuigen met een gate-conflict met 74% kan verminderen, vergeleken met een deterministisch vlucht-naar-gate-toewijzingsmodel. De robuustheid van de toewijzing kan worden geregeld met een modelparameter.

Een andere methode voor het voorspellen van vluchtvertragingen is door middel van binaire classificatie, een populaire methode in de literatuur. Echter, het binair voorspel-

len van vluchtvertragingen of vluchtannuleringen lijdt onder de grote onbalans in data. Dit veroorzaakt een verstoord beeld bij het gebruik van metrieken zoals nauwkeurigheid. In dit proefschrift wordt een systematische aanpak ontwikkeld voor binaire voorspellingen met data-onbalans, door het beschouwen van een reeks aan bemonsteringsverhoudingen, en verscheidene bemonsteringstechnieken. Twee machine learning algoritmes worden toegepast op een historische vluchtdataset van een kleine luchthaven. De resultaten onderstrepen de noodzaak om de invloed te onderzoeken van het variëren van de verhoudingen van de data-categorieën op de prestaties van classificatie-algoritmes, uitgedrukt in verschillende metrieken.

De focus van de onderzoeken beschreven in dit proefschrift is het verbeteren van de duurzaamheid en efficiëntie van luchthavenoperaties, door datagedreven benaderingen. Deze benaderingen omvatten MILP modellen, heuristieken en machine learning modellen. De ontwikkelde modellen bieden ondersteuning aan luchthavenplanners om de huidige en toekomstige roosteringstaken te verbeteren. Toekomstig onderzoek kan zich richten op het toepassen van vergelijkbare technieken op andere operaties aan de luchtzijde, en het verder verbeteren van het realisme en de operationele bruikbaarheid van de huidige modellen. Daarnaast zullen de omgevingsplanners van luchthavens, de luchtverkeersleiding en de ETV-ontwikkelaars een kritieke rol spelen in de verdere ontwikkeling en implementatie van elektrisch taxiën. Over het geheel genomen vormt dit proefschrift een beginpunt voor luchthavenplanners die door middel van datagedreven methodes de duurzaamheid en efficiëntie van luchthavens wensen te verbeteren, en daarmee een meer toekomstbestendige en betrouwbare luchtvaart mogelijk kunnen maken.

1

INTRODUCTION

1.1. RESEARCH BACKGROUND

During the last 120 years, the development of aviation has opened up enormous possibilities for fast transportation worldwide. In 2019, airlines transported 4.6 billion passengers along 8.7 trillion kilometers worldwide [73, 74]. To sustain this, the aviation sector requires a lot of resources; commercial airlines required 360 billion liters of jet fuel in 2019 [74], and a large airport such as Amsterdam Airport Schiphol requires 200 million kWh of electricity yearly [13]. This causes air travel to have a large CO_2 footprint. Therefore, organizations within the aviation sector aim to reduce their CO_2-emissions: the European Union by 90% in 2050 and the United States by 100% in 2050. In addition to increasing sustainability, the aviation sector is also seeking to make their operations as efficient and robust as possible, and to minimize the material and energy resources used. In short, the aim for airports is to improve efficiency and sustainability of airport operations.

Airport operations refer to logistic operations necessary to facilitate the aircraft throughput. Aircraft need to move from runways to gates and the other way around, but also to and from parking stands, de-icing platforms and other service locations. Typically, material, locations or personnel are assigned to an aircraft or another vehicle. Examples are flight-to-gate and flight-to-runway assignments and the assignment of refuelling or catering vehicles to a gate or aircraft.

Many types of airside operations have been standard practices for decades, but new developments can bring new operation types. A particular example is *external electric taxiing*: rather than taxiing using the jet engines, electric towing vehicles (ETVs) tow aircraft along the taxiways, as shown in Figure 1.1. Electric taxiing is expected to greatly reduce greenhouse gas emissions and noise, and improve air quality at the airport surface [70, 16, 65]. Implementing electric taxiing is therefore of interest to airports, since it greatly increases airport sustainability. This implementation introduces new scheduling problems, such as personnel-to-ETV assignment, ETV-to-aircraft assignment, and recharge scheduling [16]. The efficiency of airport operations expected to become main-

1

Figure 1.1: External electric towing vehicle towing an aircraft at Amsterdam Airport Schiphol [144].

stream in the future should be studied now, so that this foresight can be used in their implementation into the airport infrastructure, and to maximize emission savings.

Airport operational planning is often complicated by requirements that are specific to aviation. Larger airports aim to maximize the runway usage, leading to a fast-paced succession of aircraft to and from runways. This airport surface traffic needs to be routed along a limited amount of (often one-way) taxiways, which sometimes cross each other or cross runways. The amount of available ground support equipment (GSE) and the service and recharging/refuelling stations associated with them also form a limiting factor. The high runway usage implies that disruptions such as flight delays, personnel shortage or material breakdown can have a large impact on airport operations.

Strategic planning of airport operations is often done using linear programming (LP) techniques, and does not always take into account the abovementioned complicating requirements. During operation, when schedules are disrupted, they are typically adapted without using a model aiming for efficiency or robustness. In recent years, more research has been published where the efficiency or robustness of a scheduling problem is improved by a data-driven approach [171, 184, 136]. Such an approach can also be applied to airport scheduling problems, using flight schedule data or weather data.

1.2. RESEARCH GAPS

From the research background outlined above, several research gaps are identified.

Understanding the process of electric taxiing and identifying operational challenges
External electric taxiing has emerged as a new type of airside airport operation. Many airports have expressed interest in implementing it as one of their taxiing methods, because of the expected benefits for airport sustainability and noise and pollution reduction. Previous research has established the technical feasibility of the technique [39, 70, 102], and preliminary work has been performed in investigating the operational aspects [61, 16, 182, 33]. However, a comprehensive overview of these aspects is lacking.

First, such an overview should include the best way to shape the process of an operational tow with ETVs, and a list of the requirements posed by the technique to the airport infrastructure and ETV design. Many studies omit one or more elements of the towing

process in their scheduling or routing approaches. Since only one external electric taxiing system is operational, reliable ETV specifications are minimally available, leading to variations in the ETV specifications considered in literature.

Second, all operational challenges expected to appear with electric taxiing implementation, should be identified. Specifically, routing and scheduling problems, charging planning and charging capacity pose challenges to implementation. In addition, the approaches tried in literature to address these challenges or comparable ones in other fields should be considered.

This research gap is addressed in Chapter 2.

Developing models to improve electric taxiing sustainability, efficiency and robustness In order to reduce the airport emissions of greenhouse gases, but also the use of resources such as fuel and manpower, the aerospace industry continuously strives to improve efficiency and sustainability of airport operations. The efficiency and sustainability of electric taxiing operations are largely determined by the used ETV-to-aircraft schedule.

Although there has been a lot of research in scheduling problems in other fields [21, 58, 67, 189, 141], few methods for creating an ETV-to-aircraft schedule have yet been developed, due to the novelty of the technique. These works often lack one or more elements that are needed to make the model realistic and comprehensive: an electric taxiing scheduling model should take into account airport infrastructure, vehicle properties, a mix between certified and non-certified aicraft, and realistic charging constraints. Models should be useful for airport planners, and thus have limited runtime, despite scheduling problems being NP-hard. Models should be generally usable and extendable to other airports, ETV fleet mixes and ETV specifications.

The electricity demand generated by a fleet of ETVs is expected to be very large [155, 52, 112]. Charging a fleet capable of towing all aircraft at an airport requires the same power as provided on average by 1-2 wind turbines [28]. This implies that the electricity capacity of an airport should be taken into account when implementing electric taxiing. Variations in electricity capacity during a day are expected to influence the optimal towing and charging schedules, depending on airport and ETV properties.

The robustness of electric taxiing operations is the extent to which a towing schedule can be kept when faced with disruptions. In addition to scheduling in advance, it is valuable to have an extension to the scheduling model for application in real time operations, where changes to the schedule due to disruptions such as flight delays and cancellations can be minimized.

This research gap is addressed in Chapters 3 and 4.

Predicting and managing disruptions & applying disruption prediction to operational planning There are many types of disruptions that can influence a carefully optimised airport operational schedule. Flight delays and cancellations are the most obvious disruption types and are most thoroughly studied [161]. Predicting delays and cancellations, even with limited accuracy, can not only bring the financial benefits of increased on-time performance, but also improve the usability and robustness of operational schedules, which in turn reduces the necessary resources for operations.

Many authors have aimed to predict delays and cancellations based on historical flight data, and in recent years an increasing amount has made use of various types of machine learning approaches [32, 90, 164, 106, 31, 178, 154, 95]. However, all works predicting delays make use of point or class predictions. An alternative is to create probabilistic predictions [22, 57], which come with more extensive information on the uncertainty of the predictions. Such information provides additional insight to airport planners when adapting assignments.

An example of an operation type that could benefit from probabilistic predictions is the flight-to-gate assignment. When an aircraft making use of a gate is delayed, a conflict may occur with another aircraft scheduled to use the same gate. Delay predictions can help alleviate such conflicts, as shown in literature [171]. However, no studies have yet applied probabilistic predictions to minimize such conflicts.

The earlier mentioned classification approaches to flight delay prediction can be effective for airport planners to separate flights in general classes, e.g. "on-time", "small delay", "large delay" and "cancelled". However, a classification approach suffers from the fact that delay and cancellation datasets are heavily imbalanced, which distorts performance measures. This issue is not addressed in the current literature. A general approach for handling imbalanced classification data is needed, so that it can be applied to flight delay prediction. Such an approach should be based on the expected effects of using various sampling techniques and performance metrics.

This research gap is addressed in Chapters 5 and 6.

1.3. RESEARCH OBJECTIVES

Considering the research gaps defined above, the main goal and research objectives of this dissertation can be identified.

The main goal of this dissertation is to develop data-driven optimisation models to increase the efficiency and sustainability of various airside airport operations.

The objectives of this dissertation are:

Objective 1
Identify the current research status on efficient large-scale application of electric taxiing at airports.

Objective 2
Develop an optimisation model to perform ETV-to-aircraft scheduling that takes into account realistic airport circumstances.

Objective 3
Develop an optimisation approach to retain efficiency for electric taxiing when confronted with delays.

Objective 4
Improve the robustness of airside operations by predicting flight delays and cancella-

tions.

Objective 4 is divided into three subobjectives:

Objective 4.1
Perform probabilistic prediction of flight delays using a data-driven approach.

Objective 4.2
Establish a general approach to flight delay and cancellation classification with imbalanced data.

Objective 4.3
Apply the predictions to improve the robustness of flight-to-gate assignments.

1.4. RESEARCH METHODOLOGY

This section describes the methodology used to achieve the research objectives. Figure 1.2 shows how the objectives are addressed throughout the dissertation.

Chapters 2-4 address research **Objective 1 - 3**. First, an extensive literature search is performed on electric taxiing (**Objective 1**). The electric taxi systems currently under development are identified, and the differences and (dis)advantages of each system mentioned in literature are outlined. Then, the focus is put on external electric taxiing. The full electric towing process, as envisioned at future airports, is outlined, based on the specifications of ETVs and airports provided in existing studies and sources from industry. The operational challenges that lie ahead on the road to successful electric towing implementation are identified, and possible solutions are discussed. These challenges are categorised as routing, scheduling and charging challenges.

Second, a linear programming approach for full-day fleet scheduling of electric taxiing is developed (**Objective 2**). The approach includes constraints to model limited availability of electricity at airports. The linear programming approach is too complex to generate full-day towing schedules. Therefore, an Adaptive Large Neighbourhood Search approach is developed. Both approaches include an energy usage and charging model, conflict and collision avoidance, and realistic model parameters. Both approaches are applied to a large airport case study, by modelling the airport infrastructure, and making use of large datasets of historical flight schedules. The approaches are used to optimize charging and towing schedules given varying electricity availability and vehicle properties.

Last, the linear programming model for fleet scheduling, created for **Objective 2**, is extended to create a model that is used for fleet reassignment for electric taxiing operations (**Objective 3**). The model contains additional constraints to minimize changes to an existing fleet assignment, when creating an updated assignment due to observed flight delays. Several metrics are introduced to quantify change minimization. The model is applied to an airport case study, and uses flight schedules as input.

Chapters 5 and 6 address research **Objective 4**. First, data-driven algorithms are developed to perform flight delay prediction (***Objective 4.1***). Two large datasets are selected: one with flight schedule data and one with weather data. From these sets, features are selected or created that are expected to have prediction power for delay. Then, two probabilistic prediction algorithms, Mixture Density Networks and Random For-

est Regression, are introduced, and performance metrics that can quantify probabilistic predictions are selected or defined.

Second, the probabilistic predictions obtained for ***Objective 4.1*** are applied to improve the robustness of flight-to-gate assignments (***Objective 4.3***). A deterministic model, as typically applied in literature, is first defined. After this, the model is extended to consider aircraft gate presence as probabilistic, rather than deterministic. This probabilistic aircraft presence is shown to be obtainable from the probabilistic delay predictions. The robustness of the probabilistic gate assignment model is quantified by the number of flight-to-gate conflicts, and can be controlled through an overlap probability.

Last, a systematic approach is developed to predict flight delays and cancellations using binary classification (***Objective 4.2***). The predictions are made with machine learning techniques, which suffer from the data imbalance typical for events such as delays and cancellations. The developed approach considers finding the most suitable performance metric and sampling methodology, given a prediction problem. The approach is demonstrated by applying it in combination with several machine learning classification methods and various metrics, on historical data for a large airport.

1.5. DISSERTATION OVERVIEW

This dissertation is divided in seven chapters. Figure 1.2 provides an overview. Chapters 2-4 concern developing optimisation models to improve sustainability, efficiency and robustness of electric taxiing operations. In Chapter 2 the background, processes and operational management aspects of electric taxiing are investigated, and research challenges are identified. Chapters 3 and 4 aim to explore some of these challenges. In Chapter 3 models are developed to perform ETV-to-aircraft assignment, taking into account limited electricity capacity at airports. In Chapter 4 the assignment model is extended to include disruption management. Chapters 5 and 6 revolve around predicting such disruptions, to increase robustness and efficiency in operational assignments: in Chapter 5 machine learning techniques are applied to probabilistically predict flight delay given historical flight data. In Chapter 6 a systematic approach is developed to predict flight delays and cancellations given highly imbalanced datasets. In Chapter 7 conclusions are drawn based on an overview of the results of all chapters, and recommendations for future research are provided.

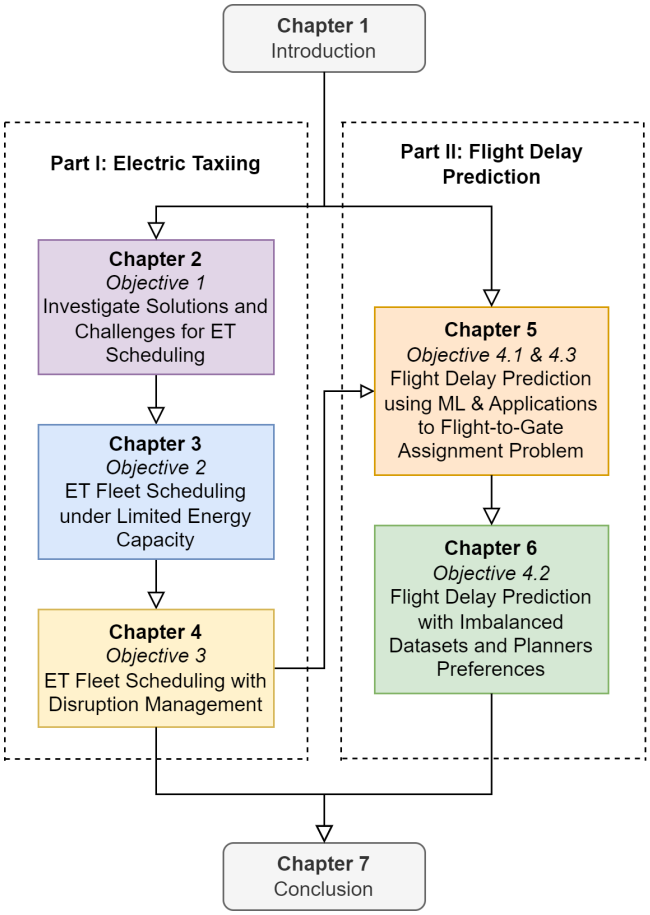


Figure 1.2: Overview of chapters and objectives in this dissertation.

2

An Investigation of Operational Management Solutions and Challenges for Scheduling Electric Towing of Aircraft

This chapter reviews existing work on operational aspects of electric towing of aircraft, and discusses management solutions. First, the varying electric taxi systems currently under development, and their implementation progress at airports, are discussed. The current specifications of Electric Taxiing Vehicles (ETVs) and the procedures needed to perform electric taxiing movements are outlined. The management needs for implementing ETVs at an airport are discussed, by reviewing existing mathematical models for ETV fleet management: dedicated vehicle routing models, ETV to flight assignment models, fleet sizing models and battery charging optimisation models. Last, the remaining research challenges are identified. This chapter summarizes the main research directions needed to support large-scale ETV implementation in the next few decades.

This chapter is based on the following research article:
Zoutendijk, M., Mitici, M., & Hoekstra, J. M. (2023). "An investigation of operational management solutions and challenges for electric taxiing of aircraft." in *Research in Transportation Business & Management*, 49, 101019 [192].

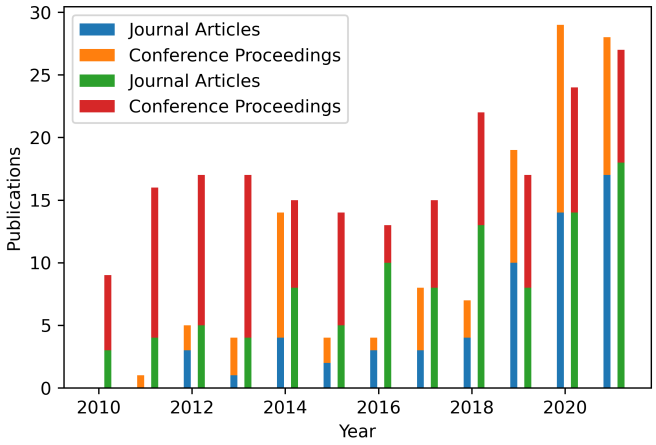


Figure 2.1: Research output containing the phrase "electric taxiing" (blue and orange, left columns) and the phrase "airport surface movement" (green and red, right columns), as indexed by Scopus [152] (accessed 11-03-2022).

2.1. INTRODUCTION

In 2017, the CO_2-emissions of the aviation sector accounted for 3.8% of world emissions. By 2050, the European Green Deal aims to reduce aviation emissions by 90%, compared to 1990 [47]. In 2021 the United States set the goal of achieving net-zero greenhouse gas emissions from the aviation sector by 2050 [49]. To achieve these goals, a large amount of research has focused on electric flying or flying using sustainable fuels such as hydrogen. However, 7% of total flight fuel use, 43% of HC emissions, 41% of CO emissions, and 12% of NOx emissions are attributed to aircraft taxiing at airports, rather than the flight phase, according to Turgut, Usanmaz, and Rosen [167]. Electric aircraft taxiing is expected to significantly reduce these emissions. In fact, the research output dedicated to electric aircraft taxiing has increased steadily in recent years, see Figure 2.1.

The current standard is to taxi with one or both of the aircraft's jet engines at roughly 7% power [17, 70]. However, this is a very fuel-inefficient way of taxiing [103]. Electric taxiing systems (ETS) are therefore a promising solution. When using an ETS, the jet engines of the aircraft are not powering the taxiing movement, thus reducing fuel consumption and emissions. ETSs are classified into two types: on-board systems and external systems. On-board systems are integrated into the aircraft and provide electric power to the nose or main landing gear when taxiing. An external system consists of a fleet of electric taxiing vehicles (ETVs) that tow the aircraft along taxiways.

One external ETS, the TaxiBot [157], has been certified for use with a significant number of aircraft types and is operating at a number of airports, while on-board solutions and other green airport solutions are still in the development phase. Lukic et al. [103] wrote a detailed technical review on the varying ETSs, and Hospodka [70] wrote a detailed cost-benefit analysis for the introduction of external ETSs from the aircraft per-

spective. However, to our knowledge no review on the operational management aspects of ETSs has been written. Although there has been research into some of these aspects separately, there is little understanding of the overall challenges to the effective implementation of this emerging technology.

This chapter reviews the research on operational management problems of *external* ETSs, and identifies the challenges that need to be addressed in the near future. To this end, the methods, assumptions and results of varying approaches are discussed, on three separate topics: the vehicle routing problem, the fleet scheduling assignment problem, and the charging aspect of electric taxiing. These topics comprise the main operational management challenges that airports face when implementing external electric taxiing. The main contribution of this work is that it aims to provide a clear view on the achievements and challenges within these topics. Together with the reviews from other viewpoints, it forms a general overview on this many faceted subject, that can aid the industry and academia in moving towards more effective and speedy implementation of ETSs. This is crucial for timely accomplishment of the goals that have been set for the reduction in aviation emissions.

In addition, this chapter discusses the differences between varying ETSs and their current implementation progress, and provides a detailed description of the taxiing process with an external ETS. For this, the specifications of ETVs and the requirements for operation are identified.

The initial selection of research contributions to consider in this work has been done by querying search engines such as Google Scholar and Scopus using the terms 'electric taxiing' and 'airport surface movement'. Afterwards, references from these works, as well as papers that cited these works, were added to the collection. These contributions were used as basis to write Section 2.2, which introduces and compares promising ETS concepts, and Section 2.3, which outlines the procedures associated with maintaining and operating a fleet of towing vehicles for electric taxiing. Then, the contributions that specifically concern the operational management aspects of external electric taxiing were selected for further review. These contributions all pertain to one or more of three areas: vehicle routing, fleet assignment and charging infrastructure. Section 2.4 reviews these works from the perspective of these three areas and identifies operational management problems that remain to be solved. Finally, Section 2.5 summarizes the findings and the recommendations for future research.

2.2. ELECTRIC TAXI SYSTEMS CURRENTLY UNDER DEVELOPMENT

In this section we illustrate the various electric taxiing concepts that are currently under development in the industry, as well as their advantages and disadvantages. This information serves as the basis needed to review the operational management aspects of electric taxiing in later sections. For an extensive review of the technical aspects of various electric taxiing systems, please refer to Lukic et al. [103].

2.2.1. ON-BOARD ETS

We will discuss three of the most promising on-board electric taxi solutions:

a) Main Landing Gear systems

A main landing gear system consists of electric motors placed in the main landing gear (MLG) of an aircraft. The installation of such a system would increase the aircraft weight by roughly 400 kg [103]. Advantages of an on-board system placed in the MLG are that a large torque can be attained, and that the turning radius of the aircraft becomes smaller than during regular operation [131, 55, 132, 86, 72]. Furthermore, the traction on the airport surface is expected to be sufficient, since the MLG carries 90% of the aircraft weight [103]. A disadvantage is that implementing a motor within the MLG is very challenging since the presence of the brakes both limits space within the MLG and provides an unwanted heat source [102].

Safran and Honeywell were developing an MLG system, but they stopped developing the project in 2016. This system was able to provide a power of 120 kW and during demonstrations in 2013 an A320 aircraft equipped with the system was able to attain a taxiing speed of 37 km/h [102].

b) Nose Landing Gear systems

A nose landing gear system consists of two electric motors placed on the rim of the nose landing gear (NLG) of an aircraft. These motors increase the aircraft weight by about 140 kg [103]. An advantage of such a system is that it allows for easy manoevrability and therefore a simplified and faster turnaround and pushback process [71, 102]. A disadvantage is that the system is powered by the APU, which has limited power. This has the following drawbacks: a) the system might not be able to provide enough traction under adverse operating conditions, b) the maximum taxiing speed that can be obtained with this ETS is 17 km/h, and c) it is unlikely that the system can be applied to wide-body aircraft in the future, according to Lukic et al. [103].

An NLG system is currently being developed by WheelTug. This system has been in the process of certification by the FAA and EASA for several years. Production and operation are expected to start soon and the WheelTug company has received orders from at least 20 airlines, but as of 2022, the system is still in the testing phase [102].

c) Hydrogen powered systems

In a collaboration with Lufthansa Technik, the German Aerospace Centre (DLR) is developing an on-board solution in the NLG consisting of a permanent magnet synchronous motor which is to be powered by on-board hydrogen fuel cells [140]. This ETS has been shown to be able to perform electric taxiing with narrow-body aircraft at a top speed of 25 km/h [102] and is reported to reduce the aircraft emissions by up to 27% [131]. Disadvantages are that the needed magnets are relatively expensive and that the ETS produces a power of only 50 kW, which makes it yet unsuitable for actually towing aircraft [132]. Furthermore, it is still challenging to store hydrogen on board an aircraft, due to the high energy content and flammability [163]. Last, hydrogen is currently still difficult and expensive to synthesize [173].

Several studies have analyzed the economical and environmental impact of on-board systems, e.g. [70, 102, 39] and [118]. Hospodka [70] calculate fuel and CO_2 emissions savings when using a 300 kg on-board ETS on an A320 aircraft: taxiing electrically reduces

Figure 2.2: TaxiBot in operation at Frankfurt Airport [157].

the needed amount of taxiing fuel by 80%. After subtracting the increased fuel need due to the added weight of the ETS, this figure is reduced to 75%, which corresponds to 0.65 tons of CO_2 per flight. Dzikus et al. [39] find that 99% of flights in the US airspace would save fuel when equipped with a 200 kg on-board ETS, on average 3% of total fuel. Nicolas [118] shows that fuel reduction depends on the combination of flight time and total taxi time: e.g. in their model, an A320 equipped with an on-board ETS with 14 min total taxi time will not experience fuel savings for flight lengths over 2400 km.

2.2.2. EXTERNAL ETS

An external electric taxiing solution consists of a fleet of electric towing vehicles that can connect to the NLG of an aircraft and perform pushback and taxiing movements. Currently, there is a similar system that has completed the development stage and is operational at airports: the TaxiBot, developed by Israel Aerospace Industries (IAI) [79]. These vehicles are currently powered by diesel engines, which are to be replaced by electrically powered versions within several years. A TaxiBot can produce a power of 500 kW and achieve a taxiing speed of 43 km/h for narrow-body aircraft [103]. Currently, only the narrow-body towing truck, with 8 wheels and a cost of 1.5 million USD, is operational. Soon, the wide-body vehicle, with 12 wheels and a cost of 3 million USD, is expected to become operational [7]. An example of a narrow-body towing vehicle in operation is given in Figure 2.2. The NLG of the aircraft is clamped onto the ETV while towing.

Several studies regarding external ETSs have discussed the expected effects of ETSs on the operational costs, fuel use and emissions [36, 39, 170, 63, 88, 129, 103], noise reduction [65] and operational safety [20]. These works typically investigate the feasibility of an external ETS at an airport, and are often based on average taxiing times and distances. Most do not take into account the variable demand or the precise routing and

scheduling involved at the operational stage. Khammash, Mantecchini, and Reis [88] show that the introduction of 4 ETVs at Lisbon Airport (LPPT) can reduce CO_2-emissions by more than 18% and lead to cost savings for both the airport and airlines. Similar results are obtained by Postorino, Mantecchini, and Gualandi [129] for Bologna Airport (LIPE). When comparing diesel-powered dispatch towing to regular taxiing, Deonandan and Balakrishnan [36] find that taxiing fuel use can be reduced by 75% and that taxiing emissions can be reduced by for instance 70% for CO_2. Dzikus et al. [39] expand on this by specifically considering the economic viability of towing short-haul flights. Vaishnav [170] and Guo, Zhang, and Wang [63] consider more cost factors such as the operation and maintenance costs for towing vehicles, compare electric taxing to other solutions such as single engine taxiing, and consider many different airports. In general, studies find that implementing external electric taxiing a) will lead to significant fuel savings in all cases, increasing with increasing taxiing distance, and b) can lead to an increase in taxiing times, especially when airports are congested.

2.2.3. Differences between the ETSs

Table 2.1 outlines the main differences between on-board and external ETSs. From a management perspective the most important difference between the on-board and external ETS is that both the need for investment and the responsibility lie with the airline or the airport, respectively. From the airport perspective, a large advantage of on-board systems is that no investment is required on their part. In order to adopt the external system, the airport needs to change the airside infrastructure, add charging infrastructure and manage the towing vehicle fleet. External ETSs require only a short pilot training for the airline and no changes to the aircraft. This is because the NLG is situated on a rotatable turret on the back of the ETV (see Figure 2.2), so that the latter can be controlled by the pilot as they would control the NLG during regular taxiing or pushback, as shown by Schiphol [144].

On-board taxiing systems are not yet operational. In contrast, the narrow-body external electric taxiing system has been certified for multiple aircraft types (comprising 70+% of worldwide commercial airline flights [157, 102]). The first demonstration took place in 2013 at Frankfurt Airport [5], and the first airport to use the ETS for operational towing was New Delhi Airport in 2019 [6]. The ETS is currently in use for testing and non-operational towing at Frankfurt, New Delhi, Amsterdam and Bangalore airports [102, 78, 6]. Several airports are planning to move to operational towing in the near future: Bangalore Airport was planning this for 2023 and Schiphol Airport for 2024. New Delhi Airport is planning to expand their ETV fleet to 15 vehicles by 2025 [76, 117].

Table 2.1 also summarizes the operational differences between the on-board and external electric taxiing systems. During regular taxiing the aircraft taxiing speed is 56 km/h [146]. Roling, Sillekens, and Curran [134] show that the minimum taxiing speed needed to prevent airport surface congestion at Amsterdam Airport Schiphol is 32 km/h. Therefore, the use of the on-board system is expected to increase the taxi time substantially. Nevertheless, the average pushback time is expected to be reduced, since no pushback vehicle needs to be connected to and disconnected from the aicraft [4, 122]. When operating the external system, a connecting and disconnecting procedure is still required, but now to the ETV, rather than the pushback vehicle. This procedure requires

Table 2.1: Comparison between on-board and external ETSs.

Investment	On-board ETS	External ETS
Acquisition costs per system	Undisclosed	USD 1.5M (NB), 3.0M (WB)
Adjustment aircraft	Install system at NLG or MLG	N/A
Adjustment airport	Not required	Management fleet of ETVs and charging and routing infrastructure
Suitable aircraft types	NB	All
Implementation progress		
First demonstration (manufacturer, aircraft, airport)	2005 (WheelTug, B767, KMZJ)	2013 (IAI, B737, EDDF)
First operational ETS (manufacturer, aircraft, airport)	N/A	2019 (IAI, A320, VIDP)
Certified aircraft type (year)	Ongoing	B737 (2014), A318-21 (2017)
Airports (nr ETVs in use, year)	N/A	EDDF (1, 2014), VIDP (2, 2019), EHAM (2, 2020), VOBL (1, 2021)
Operational aspects		
Taxiing speed	17 km/h	42 km/h
Additional operations	N/A	Connecting/Disconnecting ETV to/from aircraft
Engine warm-up	During taxiing/ near runway	During taxiing/ near runway
Engine cool-down	At gate	At gate
Electricity source	APU	ETV battery
Charging	APU generator	Charging stations
Added aircraft weight	140 kg	N/A
Environmental effects		
Fuel saving (B737)	85%	50-85%

three minutes, according to Schiphol [144].

Last, both ETSs are expected to greatly reduce the needed taxiing fuel, and consequently, the taxiing emissions. Tests with the external ETS at Amsterdam Airport Schiphol with a Boeing 737 resulted in taxiing fuel savings of 90%, which reduces to 50-85% when taking into account (dis)connecting and engine warm-up. Schiphol [145] shows that taxiing with an on-board ETS saves 85% of the fuel. The APU powering an on-board ETS is charged during flight. On the other hand, external ETSs charge on the ground and thus require charging infrastructure at the airport. Since the aircraft is not modified when using an external ETS, any fuel savings also directly contribute to weight reduction of the aircraft, and therefore to further fuel and emission savings during the flight.

2.3. ETV fleet management procedures

In the previous section we have observed that the external ETS is the system that is at the most advanced implementation stage. In this and following sections we focus on external ETSs.

Implementing an external ETS implies integrating a fleet of electric taxiing vehicles into the regular airport surface traffic. This poses several management challenges. This section outlines the management procedures for electric towing operations, and the roles played by the ETVs, the airport, the aircraft, and Air Traffic Control (ATC).

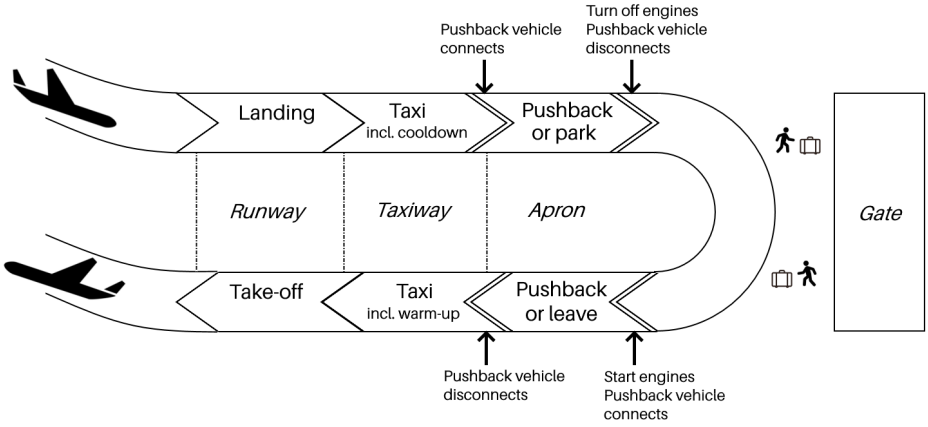
Figure 2.3 shows the regular aircraft taxiing procedure and the procedure for towing aircraft using ETVs. When considering *regular taxiing* (Figure 2.3a), the aircraft lands on the runway, and starts taxiing. The jet engines cool down during taxiing. After arriving at the apron, the aircraft parks at a gate, or a pushback vehicle is connected to the aircraft, and pushes it into a parking position. For taxi-out, the procedure is reversed, and the jet engines warm up while taxiing.

When considering *electric taxiing* (Figure 2.3b), the aircraft connects to an ETV directly after landing. The ETV tows the aircraft to a parking position at the apron. For an aircraft departure, an ETV tows the aircraft from the parking position to the runway. Here, the ETV is disconnected and the aircrafts jet engines are warmed up for take-off. Below a detailed description of electric taxiing with an ETV is given.

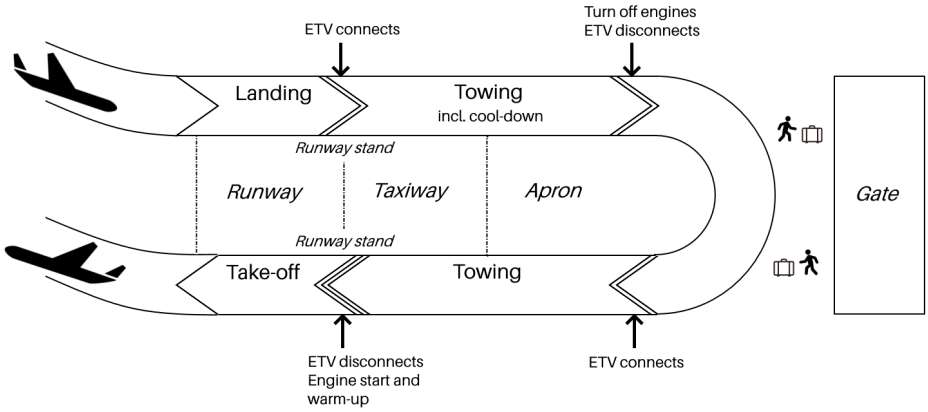
Non-towing ETV: Before an arriving aircraft lands, an ETV is on its way towards the runway. It can come from a previous task at a gate or a runway, an ETV depot, or an ETV charging station. For most airports it is expected that ETVs will use service roads for non-towing movements, to avoid a large increase of traffic on the taxiways that needs to be regulated [146, 180]. A typical maximum speed on these service roads is 30 km/h [143, 116].

Landing and connecting: After an arriving aircraft lands on a runway, it needs to connect to the ETV. The connection process takes roughly three minutes [144], but some runways receive arriving aircraft at a rate faster than one per three minutes. Therefore, it is expected that airports should designate a separate space near the runway exit for the connection and engine cool-down processes to take place, without interfering with the runway traffic [122, 103, 142]. In Figure 2.3b this space is indicated as a *runway stand*. Ideally the runway stand would consist of a paved area separate from the taxiway, large enough for maneuvering. In that case, self-taxiing aircraft would be able to pass by connecting aircraft, avoiding blockage of the taxiway near the runway exit [142].

Towing: After connecting, the ETV tows the aircraft along the taxiways. During regular taxiing, trailing aircraft need to keep a safe separation distance from a leading aircraft to avoid its jet blast. This distance directly influences the throughput of aircraft on the ground. A typical value for a safe separation distance is 200m [133, 84, 158]. During electric taxiing, there is no jet blast. However, the reaction time of the pilot and the braking distance of the aircraft still need to be accounted for: the separation distance is expected to remain necessary, but smaller. Last, ATC retains its task of conflict avoidance on the taxiways, but now with changed taxiing speeds and separation distance.



(a) Regular taxiing



(b) Electric taxiing

Figure 2.3: The taxiing process for a turnaround (arrival and departure) for regular taxiing and electric taxiing with an ETV.

Table 2.2: Operational specifications for electric taxiing.

Speed on taxiways	42 km/h
Speed on service roads	30 km/h
Connecting & disconnecting time	3 min
Engine warm-up and cool-down time	3-5 min
Minimum separation distance	200 m or less

Parking: Upon arriving at the apron, the ETV tows the aircraft into parking position, without the need for a pushback truck. When the aircraft is in position at the gate, the ETV disconnects from the aircraft in 3 minutes [144].

Engine warm-up: When the aircraft is ready to depart, an ETV connects to it at the gate, and tows it to a runway stand. The main difference with the arrival procedure is that the jet engines of a departing flight need to warm up before take-off. During normal operation the engines are warmed up during the regular taxiing procedure, while during electric taxiing, the jet engines are not used [157].

There are several possibilities for the location of the engine warm-up:

i) Engine warm-up at the runway stand. This minimizes the amount of time the engines are running before take-off, and therefore minimizes the engine emissions.

ii) Engine warm up during towing. This could raise safety concerns: if there are problems with the engine during towing, then the taxiway traffic will be disrupted.

iii) Engine warm-up at the apron, as is the case for regular taxiing. A disadvantage is that the engines produce emissions during the entire towing process.

The engine warm-up time, sometimes referred to as ESUT (Engine Start-Up Time), is typically estimated between 3 min [144, 137] and 5 min [102, 39].

Table 2.2 summarizes the relevant operational specifications of the ETV, the aircraft, and the airport.

2.4. MANAGEMENT CHALLENGES FOR ELECTRIC TOWING VEHICLES

The economic, environmental, and technical aspects of implementing external electric taxiing solutions have been reviewed by Hospodka [70] and Lukic et al. [103]. Complementary to these studies, this section reviews existing work on the operational management of external ETSs, using three main challenges as starting points for identifying challenges: the routing of ETVs, ETV fleet assignment and electric infrastructure.

2.4.1. ETV VEHICLE ROUTING PROBLEM

Given the daily flight schedule, airport planners assign a route along the taxiways for each departing/arriving aircraft. Departing aircraft taxi from gate to runway, and arriving aircraft from runway to gate. The routes need to planned in such a way that conflicts are avoided and taxi time is minimized.

Table 2.3: Assumptions and approaches to VRPs and FSAs used in literature on electric taxiing.

Study	Objectives	Problem formulation	Conflict avoidance
Sirigu et al. [156]	shortest path	simulation	no
Baaren and Roling [16]	minimize taxiing fuel	MILP	no
Zaninotto et al. [182]	minimize taxi time, conflicts	simulation	penalties
Soltani et al. [160]	minimize taxiing fuel & delays	MILP	yes
Salihu et al. [137]	minimize taxiing costs	simulation	yes
Oosterom et al. [124]	minimize number of ETVs	MILP+Greedy	MILP
Study	**Airport**	**Number of movements**	**Fleet size range**
Sirigu et al. [156]	LIMF	N/A	N/A
Baaren and Roling [16]	EHRD&EHAM	39&1430	0 to 42
Zaninotto et al. [182]	LMML	36	Unconstrained
Soltani et al. [160]	CYUL	205	0 to 20
Salihu et al. [137]	CYUL	644	10 to 30
Oosterom et al. [124]	EHAM	913-1258	38 to 50

The *vehicle routing problem* (VRP) aims to answer the question: "Which is the optimal route to take for a certain vehicle to reach an ordered list of destinations?" for every vehicle in a fleet. This problem appears in many types of delivery or collection problems, such as for postal companies or robot planning in warehouses [107]. Often, additional constraints are involved, such as time windows, loading and unloading or vehicle capacity.

The problem of obtaining optimal taxiing routes for taxiing aircraft from gates to runways or vice versa is also usually posed as a VRP. When considering electric taxiing, such a VRP can be extended with charging constraints for the ETVs.

Table 2.3 provides an overview of methodologies and assumptions used in literature for electric taxiing. In this section we compare these approaches.

GRAPH REPRESENTATION OF AIRPORT LAYOUT

All studies considered in Table 2.3 use a graph representation of the airport surface. The edges represent the taxiways and service roads, and the nodes represent intersections, gates or gate groups, runway entrances and exits, runway stands and ETV depots. For example, Soltani et al. [160] use multiple runway entrance and exit nodes, and Zaninotto et al. [182] use runway stands and ETV depots.

Figure 2.4 shows a schematic representation of an airport with six gates and two runways. In this example, towing vehicles and other ground support equipment are not allowed to drive on the taxiways. Therefore it should be possible to travel between any combination of runway and gate via both taxiways and service roads. Furthermore, multiple runway entrance and exit points are reachable.

The layout of the airport and the possible runway configurations influence the per-

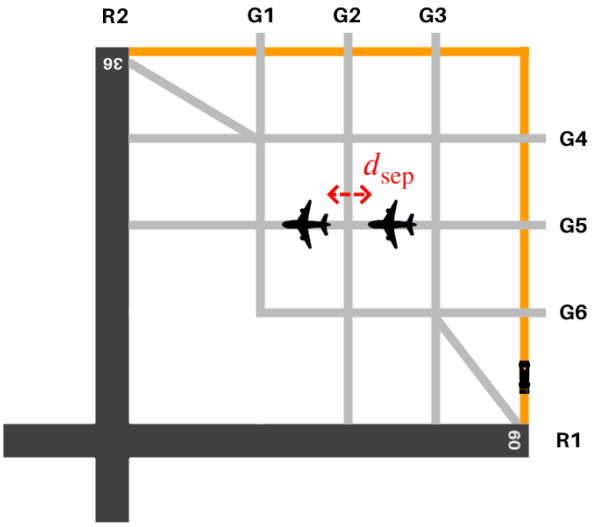


Figure 2.4: Schematic representation of a sample airport. Wide dark gray lines represent the runways R1 (09) and R2 (36). Light gray lines represent the taxiways, where aircraft may taxi, and orange lines represent service roads, where ETVs may drive. Gates G1-6 are indicated. The separation distance d_{sep} between two aircraft has to be respected.

formance of an ETV fleet. Zaninotto et al. [182] consider one of the runways of LMML, but in later work also apply their algorithm to LFBO, LLBG and KDFW [180]. Soltani et al. [160] and Salihu et al. [137] use three runways with 20 entrance/exit nodes of CYUL. Baaren and Roling [16] consider EHRD, which has 1 runway with a 2km taxiing route, and EHAM, which uses 2 or 3 runways with a regularly changing configuration, and a longest taxiing route of 11 km.

Modelling approach

Given these inputs, routing can be performed on the airport surface. Sirigu et al. [156] consider the algorithms (Modified) Hopfield Neural Networks, Dijkstra and A* to find the shortest taxiing routes. Soltani et al. [160], Baaren and Roling [16] and Oosterom et al. [124] formulate the routing and scheduling problem as an MILP. All possible routes between gates and runways are calculated in advance, and the usage of a route is included in a decision variable.

On the other hand, Zaninotto et al. [182] and Salihu et al. [137] develop a simulation, in which each movement is scheduled sequentially, and routing is performed using Dijkstra's algorithm. Except for Zaninotto et al. [182], who divide the optimization time into 20-second time windows, all approaches in Table 2.3 use continuous time values.

Routing conditions

Conflict avoidance for surface movement: Including conflict avoidance in taxi route planning makes it more realistic, and can help identify problems such as traffic jams at taxiway intersections. Zaninotto et al. [182] include conflict avoidance by introducing

penalties for using already occupied edges for a new route, and found a trade-off between minimizing the number of conflicts and increasing the taxi time. Soltani et al. [160] include eight constraint sets in their MILP formulation to ensure conflict avoidance on all edges and nodes in the routing solution. Oosterom et al. [124] create and solve a separate MILP formulation for routing vehicles, which enforces conflict avoidance through constraints, before solving the MILP formulation for scheduling. The performance of the latter is compared to that of a greedy algorithm. Salihu et al. [137] enforce conflict avoidance in their simulation by respecting the separation distance and following a first-in-first-out procedure at intersections. Regarding the movement of unloaded ETVs, Soltani et al. [160] and Salihu et al. [137] assume they travel on the taxiways, but clear the way for aircraft. Baaren and Roling [16], Oosterom et al. [124] and Zaninotto et al. [182] assume these ETVs will use the service roads, except when no other route is available. In short, in the reviewed literature, conflict avoidance for unloaded ETVs is assumed to not be required.

ETV movement between tasks: An ETV that is not towing or charging is waiting for its next towing task. The behavior of the ETV during this time constitutes a management choice:

i) In case all routes for an ETV have been determined in an optimization model, one can choose to let the vehicle proceed to the starting point of its next task, as in Baaren and Roling [16].

ii) The ETVs remain idle at the location where they performed their task, as in Salihu et al. [137].

iii) In case there are multiple ETV depots: one can choose to have ETVs return to one of the depots, as in Zaninotto et al. [182]. This can be used as a technique to pursue a good spread of ETVs on the airport surface. One can for example select the closest depot, or the depot with the least amount of other idle ETVs.

Start and end time of the taxiing procedure: In order to keep to the schedule, it is important to minimize deviation from the scheduled taxiing start and end time. Soltani et al. [160] create an upper and lower bound for the start and end time of all taxiing movements in their MILP model. Baaren and Roling [16] assume no en-route delays occur and set the start and end of taxiing to a fixed moment in time. Oosterom et al. [124] calculate these times to ensure conflict avoidance. On the other hand, Zaninotto et al. [182] investigate to what extent deliberately delaying the start of taxiing by a fixed amount can help decrease the number of routing conflicts.

ASSUMPTIONS ABOUT MODEL PARAMETERS

The specifications of aircraft and ETVs have a large influence on routing. An example is the assumed connecting and disconnecting time of ETVs. Zaninotto et al. [182] and Salihu et al. [137] assume one minute for these operations, while for example the TaxiBot requires three minutes [157].

Another routing parameter is the minimum separation distance between taxiing aircraft. Zaninotto et al. [182] use a minimum separation distance of 300 meters, based on

minimum clearance, pilot reaction time and braking distance. In contrast, Salihu et al. [137] use a distance of only 15 meters, based on Australian Civil Aviation regulations.

A third parameter is the taxiing speed. Sirigu et al. [156] assume a constant taxiing speed of 10 m/s. Baaren and Roling [16] and Oosterom et al. [124] use the specifications of the TaxiBot, i.e. 11.8 m/s. Salihu et al. [137] assume a regular aircraft taxiing speed of 7 m/s, and speeds of 4 m/s and 7 m/s for a towing and non-towing ETV, respectively. Zaninotto et al. [182], Oosterom et al. [124] and Baaren and Roling [16] include the acceleration and deceleration of ETVs in their simulation. The others keep to a constant velocity.

The value for separation distance and taxiing speed directly influence the taxiway capacity: they can make the difference between a model showing that electric taxiing leads to small increases in taxi time, such as for Baaren and Roling [16] or that it leads to large [137] or even unacceptable taxiing delays. Therefore, it is important to obtain realistic estimations for these parameters.

Last, the engine warm-up needs to be incorporated in the planning of a departing aircraft. Salihu et al. [137] and Baaren and Roling [16] incorporate this into their model but do not specify the exact time taken. Salihu et al. [137] assume warm-up occurs during taxiing and Baaren and Roling [16] and Oosterom et al. [124] assume warm-up occurs after taxiing.

Challenges

ETV manager to assist Air Traffic Control: Managing a fleet of ETVs increases the work load and responsibilities of Air Traffic Controllers: the Ground Controller, who manages the traffic on the taxiways [158], now also needs to route all ETVs and make sure conflicts are avoided. ATC will need to be aware which aircraft are taxiing by themselves and which are towed by a vehicle. A possible solution is to add a separate role, that of *ETV manager*, to the airport. The ETV manager can be involved both in the routing and scheduling of the ETV fleet, as well as monitoring of the actual movement and dealing with disruptions in the schedule. They should be in close contact with ATC to ensure smooth operation at the airport. Workload is also increased at the airport surface; each of the ETVs will need a driver. Airport planners will have to take into account the working times and breaks of the drivers when creating the towing schedule with its driving and charging periods.

Autonomous airport surface movement: Another solution to mitigate the increased ATC workload is to aim for autonomous routing and scheduling of all airport surface movement (ASM) [160]. For example, EUROCONTROL is working towards an Advanced Surface Movement Guidance and Control System (A-SMGCS), which is an automatic system that supports ATC in monitoring ASM operations by e.g. creating routes, monitoring possible conflicts, and operating stop bars and lights automatically [41]. An autonomous system can increase the safety, predictability and reliability of operations, by avoiding ground incidents, miscommunications and other human errors, and decrease delay and costs due to smart planning [103, 150, 149].

Several authors have been working towards autonomous ASM: an example is Zaninotto, Gauci, and Zammit [180], who simulate ASM by connecting various programmed mod-

ules, such as a *Vehicle Simulator*, *Path Planning with Dynamic Obstacles* and *Tow Trucks Optimisation System*. Such a simulation can form the underlying model for an autonomous system for ASM. Going even further, Morris et al. [113] apply self-driving vehicle technology to the problem of towing aircraft. In their model, towing vehicles drive by themselves, but are supervised by ATC in a Human-Machine Interface. Although the routing and scheduling is performed by an algorithm, resolving separation constraint violations remains the task of the controller. Okuniek and Beckmann [122] note that the successful implementation of A-SMGCS depends on the ability of aircraft to follow the required surface movement plan: an autonomous system of intercommunicating ETVs can contribute to this goal. For example, one could program the ETVs in such a way that they communicate with each other to avoid conflicts and enforce separation distances, but also to avoid unnecessary braking and speed changes. This is expected to help the ETVs to follow the most fuel-efficient driving strategy.

Airport routing guidelines and taxi times: The additional operations associated with electric taxiing, combined with the reduced maximum taxiing speed (discussed in Section 2.2 and 2.3) can lead to increased taxi times and congestion of airport surface movement. There are several management measures that airports can consider when aiming to increase the efficiency of ETV routing. As discussed earlier, airports that are expected to experience taxiway congestions due to unloaded ETV movements on their taxiways, might seek to construct wider or more service roads. The implementation of runway stands for arriving aircraft that are connecting to an ETV could alleviate congestion near arrival runways. An airport that aims to implement external towing but needs to limit total taxiing time might consider allowing ETVs to travel faster on the service roads. Another option is to investigate whether the taxiing separation distance can be reduced, since there is no jet blast from an aircraft when it is being towed. The challenge of increased taxi times and congestion is expected to be particularly important for airports to address, since they will aim to implement electric taxiing without having to reduce the throughput of aircraft.

2.4.2. ETV FLEET SCHEDULING ASSIGNMENT PROBLEM

The *fleet scheduling assignment problem* (FSA) is the problem of assigning vehicles to tasks in a travelling schedule. This problem appears for example in taxi fleet scheduling, and the assignment of aircraft to flight numbers. When considering electric taxiing, all towing tasks need to be assigned to a specific ETV. This can be formulated as an FSA, which can be optimized for varying objectives.

Figure 2.5 shows an example of fleet assignment on an airport, with three aircraft in three different situations.

MODELLING APPROACH

The models used in literature minimize the taxiing time [182], the taxiing fuel [16, 160], combine this into a taxiing cost [137], or minimize the number of used ETVs [124]. The linear programming models aim to find the best routes and assignments for all aircraft movements at once, while the simulation approaches move through the flight schedule

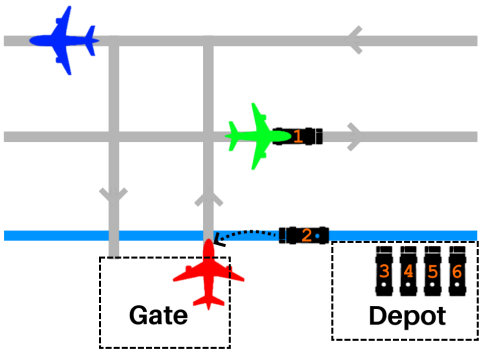


Figure 2.5: Schematic representation of fleet assignment on an airport. The red aircraft is waiting at a gate for ETV 1 to arrive, so that it can be towed towards a runway. The green aircraft is being towed by ETV 2. Other ETVs are waiting to be deployed at a depot. The blue aircraft is taxiing by itself. Taxiway directions are indicated by arrows.

and perform route planning and vehicle assignment for each aircraft sequentially. The constraints used are typically grouped as:

i) Assignment constraints, for example: an aircraft should have only one vehicle assigned to it (see Baaren and Roling [16] eq. 3);

ii) Route flow and route timing constraints, for example: the arrival time at a node is calculated with the edge speed and the departure time of the previous node (see Soltani et al. [160] eq. 6, 7);

iii) Collision avoidance constraints, for example: two aircraft that reach the same node from different edges must be separated by a separation time or distance (see Soltani et al. [160] eq. 13, 14);

iv) Energy or fuel constraints, for example: the energy required for an upcoming task should be smaller than the current state of charge of the ETV (see Baaren and Roling [16] eq. 10).

RESULTS

The existing studies performing fleet scheduling all have their own optimization objectives and corresponding results. Baaren and Roling [16] show that electric taxiing uses less fuel than regular taxiing in all cases. While towing all aircraft with ETVs provides the largest fuel and emissions savings, Salihu et al. [137] calculate that at CYUL, the most economical solution in their combined cost model is to only use electric taxiing for departing aircraft. Using their simulation model that balances taxi delay and routing conflicts, Zaninotto et al. [182] show that slowing aircraft to halve the number of aircraft conflicts in a schedule leads to an increase of 13% in taxi time at Malta International Airport (LMML), compared to a situation with unresolved conflicts. Furthermore, allowing aircraft up to four minutes waiting time before starting the towing procedure can also reduce the number of conflicts.

Fleet sizing: An important parameter which is often the subject of optimization or sensitivity analyses is the size of the ETV fleet. Zaninotto et al. [182] assume an infinite number of tow trucks, while Soltani et al. [160] find that the economic optimum for the fleet size at CYUL is 12 vehicles, when taking into account ETV operating costs, fuel and delay costs. Salihu et al. [137] arrive at an optimum of 16 vehicles for electric towing of departing aircraft, when taking into account the annual total taxi time and annual operating costs. When considering electric towing of all aircraft, they find an optimum of 26 vehicles. Baaren and Roling [16] show that introducing 5 towing vehicles at EHRD decreases the fuel use by 65% and introducing 24 towing vehicles at EHAM decreases the fuel use by 75%. Furthermore, they show that increasing the fleet size further is less cost effective due to decreasing marginal fuel savings. Oosterom et al. [124] use the fleet size as the objective of the MILP formulation, and find that it has a roughly linear relation to the number of flights. In general, these studies showcase the trade-off between additional sustainability benefits and increased operating costs that appears in the fleet sizing problem.

SCHEDULING MANAGEMENT DECISIONS

An important scheduling decision is which aircraft are to be towed and which aircraft will taxi by themselves. Baaren and Roling [16] choose regular taxiing if it is more fuel-efficient than electric taxiing, but find for both airports that this occurs in none of the cases. Furthermore, they have aircraft taxi regularly when their taxi time is smaller than the engine warm-up time. Unlike e.g. for delivery problems, not towing an aircraft still results in the aircraft participating in the airport surface movement. This means both the towed and self-taxiing aircraft are factors to consider in scheduling management decisions. Soltani et al. [160] select the self-taxiing option when the cost of the delay incurred by waiting for an ETV is larger than the fuel saving benefit. In Salihu et al. [137], all aircraft are towed, and it is shown that this leads to enormous costs and delays if the ETV fleet is not large enough.

Second, the wide-body aircraft have to be towed by the wide-body ETV and the narrow-body aircraft preferably by the narrow-body ETV. Baaren and Roling [16] incorporate these two types of ETVs in their optimization. Oosterom et al. [124] allow a heavy-wide body vehicle in addition.

Last, one needs to decide whether an ETV should be present at the aircraft at the scheduled taxiing starting time [16, 182, 160], or whether the ETV starts to move towards the aircraft at this time [137].

CHALLENGES

Based on the literature with regard to ETV fleet scheduling, the following challenges for future research are identified:

Robust scheduling and disruption management: When creating a schedule it is often assumed that all operations take place as planned. However, at execution disruptions can take place, such as flight delay, mechanical failures of an aircraft or ETV, or unavailability of a road, gate or runway. Airside disruptions present a relevant challenge specifically for this type of ground operations, since effective operation requires the vehicles to be present at the correct gate or runway at the correct (disrupted) time, while

the fleet is likely spread around the airport during operation. This is in contrast to other ground vehicles, which operate mainly in the gate areas. This means it might be more difficult to find an ETV that is near enough to perform a task it was not originally assigned to. Zaninotto, Gauci, and Zammit [180] introduce a probabilistic version of their vehicle movement simulation by varying the vehicle speed. When comparing this to the algorithm defined in their previous work, Zaninotto et al. [182], they obtain double the amount of vehicle conflicts (violations of separation distance) for LFBO, for arrival rates larger than 30 aircraft per hour. This illustrates that such disruptions can cause negative effects on the carefully optimized schedule objective. Soltani et al. [160] recommend to consider stochastic events such as weather conditions, de-icing operations and the reliability of ETVs.

Steps can be taken to reduce these effects both before operation (*robust scheduling*) and during the operation (*disruption management*). Robust scheduling can be performed by considering the effects of possible disruptions on a given schedule. For example, one can run a simulation of a given schedule, where disruptions occur with a given probability. These probabilities can be estimated or predicted based on earlier occurrences and other factors. Based on such a simulation, changes can be made to an FSA to make it more robust. Another option is to create a robust schedule from scratch, by incorporating constraints that guarantee the robustness into the scheduling problem. An example of robust scheduling in literature is Jamili [81], who creates an MILP and a Simulated Annealing heuristic for robust aircraft routing and scheduling using traffic on origin-destination pairs as input. Cadarso and Celis [24] consider stochastic demand figures and uncertain operating conditions in a robust planning model for flight timetables and fleet assignments, and show that the number of misconnected passengers can be reduced.

Disruption management is a continuous process: as soon as planners are aware of a disruption, they will need to make changes to the schedule. Towing routes might have to be deconflicted, and gate, runway or ETV assignments might have to change. A typical objective in disruption management is to minimize the number of changes needed to reach a feasible or locally optimal schedule again. More changes means an increased workload for personnel and increased uncertainty for passengers, and often leads to increased costs. Oosterom et al. [124] perform disruption management by testing their MILP and greedy algorithms in a 30-minute rolling horizon approach, and investigating which fraction of the amount of originally towed aircraft can still be towed. Using this approach, they find fractions of 94% (greedy) and 98% (MILP) for the busiest test day. Another example of disruption management in aerospace is Lee et al. [96], who develop an optimization model of disruption recovery for a network of airports, and integrate a stochastic queuing model of congestion therein. This approach reduces expected disruption recovery costs by 1 to 4%. Tang, Lin, and He [162] develop a dynamic model to simultaneously optimize vehicle schedules and electric fleet sizes of electric buses. The model incorporates road-traffic stochasticity to mitigate the breakdown of a vehicle.

Note that robust scheduling and disruption management have only limited capability to mitigate disruptive effects, due to the stochastic nature of disruptions. This means that it is likely that there will be departing or arriving aircraft that need an ETV at a time when their scheduled ETV is unavailable. One solution would be for the aircraft to per-

form self-taxiing. Another would be to maintain a group of separate ETVs that are not assigned to any aircraft, but tow aircraft for which no other ETV is available. On a large airport with many gates and runways entrances/exits that take long to drive to, such a spare ETV may take a long time to arrive at the aircraft. A trade-off is expected between the costs of extra delay for the aircraft and the costs of maintaining a larger group of ETVs for this purpose.

Technological developments: As the development of operationally deployable external ETSs progresses, more becomes clear about the technological specifications of the towing vehicles. Such specifications can be used in research to make models for scheduling and routing ETVs more realistic. For example, the list of aircraft types that have been certified for using the external ETS (as shown in Table 2.1) can be used to create a routing and scheduling model that represents an intermediate implementation situation where only a part of the aircraft fleet may be towed. Similarly, including both the narrow-body and wide-body ETVs and their specifications in a model introduces several unaddressed scheduling considerations: the fleet sizing problem with two types of vehicles, the utilization of either type, but also the influence of differing charging rates and electricity usage of the two types on the routing and scheduling.

2.4.3. CHARGING FOR ELECTRIC VEHICLES

Currently, the towing vehicles operating at the airports shown in Table 2.1 are diesel-powered. Eventually, all towing vehicles are expected to become actual ETVs, which regularly need to recharge their batteries. This can take considerably longer than refueling for vehicles operating on fossil fuels, so that the recharging time becomes an important part of the vehicle planning. Note that another possibility is to create a battery swapping system. Given a taxiing schedule for an airport, one can aim to find an optimal charging strategy, depending on e.g. the size and type of batteries in the ETV. The locations of charging stations influence both the routing and the charging schedule, and it is therefore vital to optimize their placement on the airport surface.

Most of the electric taxiing literature investigating VRPs and FSAs does not take into account charging for their routes and schedules. In this subsection, we review literature that considers the charging aspect of managing a fleet of electric vehicles (EVs), from both airport surface movement and other fields. Table 2.4 provides an overview of this literature.

OPTIMAL CHARGING STRATEGY

Since charging an EV can be a time-consuming process, it is important to find the best time for charging and the best charging method, while taking into account the requirements and specifications involved.

Charging period: When an EV is being charged, it can be charged to its capacity (*full recharge*), or for a fixed amount of time or time steps (*fixed charging time*). A third option is to charge until it is needed for operation (*partial recharge*). In Baaren and Roling [16], every vehicle is charged for a fixed time in between any two jobs, and it is assumed that the vehicle is fully charged after this. Similarly, Hiermann et al. [67] assume that a

Table 2.4: Literature on the management of charging a fleet of electric vehicles. GSE indicates Ground Support Equipment and MILP indicates Mixed Integer Linear Programming.

Study	Electric application	Model formulation
Hiermann et al. [67]	General	MILP
Schiffer and Walther [141]	General	MILP
Baaren and Roling [16]	Aircraft taxiing	MILP
Lin et al. [99]	Buses	MILP
Gulan et al. [62]	GSE	Monte Carlo simulation
Xiang et al. [175]	GSE	Sequencing algorithm
Oosterom et al. [124]	Aircraft taxiing	MILP
Study	**Charging strategy**	**Charging station placing method**
Hiermann et al. [67]	Full recharge	Bidirectional labelling
Schiffer and Walther [141]	Full & partial recharge	Included in MILP
Baaren and Roling [16]	Full recharge	Given
Lin et al. [99]	Full recharge	Locations selected in MILP
Gulan et al. [62]	Partial recharge	Given
Xiang et al. [175]	Partial recharge	Given
Oosterom et al. [124]	Partial recharge	Given

vehicle is recharged till full, when it arrives at a charging station, and Lin et al. [99] fully recharge electric buses overnight. On the other hand, Gulan et al. [62] and Xiang et al. [175] allow partial recharging. In these studies, every vehicle is given attributes such as the charging level, vehicle activity, vehicle type and availability for tasks. Based on these attributes and the tasks that need to be performed, a selection is made which vehicle will be charged during this time step and which will be sent to perform a task. Schiffer and Walther [141] allow both full and partial recharging, as do Oosterom et al. [124] who define the amount of charge through keeping track of the state of charge of a vehicle after towing an aircraft.

In addition to the actual charging period, a vehicle needs to travel to and from a charging station. Baaren and Roling [16], Oosterom et al. [124] and Hiermann et al. [67] include the routing of EVs to and from tasks and charging stations in their schedules. However, Gulan et al. [62] and Xiang et al. [175] do not take into account travelling between tasks and charging stations. Instead, a large time step of 15 minutes is taken in which vehicles are either charging or performing their duties.

Problem formulation and model inputs: Gulan et al. [62] perform a Monte Carlo simulation and Pareto front analysis to test combinations of input parameter values on their joint objective: minimizing the needed amount of electric vehicles and minimizing the amount of gas used by gas vehicles (the alternative to the EVs). These input parameters include the number of charging stations, the number of each type of vehicle and the maximum electrical load of the terminal. The GSE tasks are derived from synthetic flight schedules, and the simulation runs for three schedule days. Xiang et al. [175] create a

sequencing algorithm to perform a similar simulation with the goal of maximizing the usage of electric vehicles. The charging algorithm is an input to a larger model that investigates the costs of an airport energy microgrid including hydrogen, solar and battery energy sources. The authors used a year of historical flight data from Chengdu Airport (ZUUU) to find the GSE tasks that need to be performed and the electrical load needed at the airport.

In their MILP approach to the ETV routing and scheduling problem at EHAM and EHRD, Baaren and Roling [16] include constraints enforcing vehicles to charge in between tasks. Oosterom et al. [124] control the charging process by enforcing constraints regarding the state of charge of vehicles after towing an aircraft. Hiermann et al. [67] create an extensive model for vehicle routing with time windows, charging station placement and fleet sizing at the same time. Their goal is to cover a set of customers on the routes, while minimizing the number of needed EVs and their total travelled distance. They compare two different solution approaches: an MILP formulation, and a combination of Adaptive Large Neighbourhood Search and local search algorithms. A bidirectional labelling algorithm is used to determine the optimal placement of charging stations. These approaches are able to solve instances with 15 customers and 2 to 8 charging stations within a gap of 1% compared to best known results. Similarly, Schiffer and Walther [141] consider a model incorporating charging station placement, capacity constraints, time windows and recharging. Several objectives were considered in this model, such as minimizing travel distance, the number of needed vehicles and charging stations, and the total costs. The authors show that reducing the solution space to strengthen the model formulation ensures that more benchmark instances of 5, 10 or 15 customers can be solved to optimality, in a shorter computation time.

Battery specifications: Important specifications with regard to charging are the battery capacity, and the charging and depletion rate of the battery. Most studies shown in Table 2.4 do not specify battery capacity, recharging time or energy consumption rate. The medium towing vehicle introduced by Baaren and Roling [16] has a battery capacity of 840 kWh and a maximum power of 1400 kW. They find that an average tow of a medium aircraft would require 33 kWh at EHAM. The vehicles used by Oosterom et al. [124] have capacities ranging from 400 to 3200 kWh, and charging power ranging from 100 to 500 kW. Adegbohun et al. [1] note that fully charging the battery of a 50-100 kWh EV requires two to three hours, or 0.5 to one hour for fast charging. Modern electric pushback trucks, capable of towing fully loaded aircraft for short distances, have a battery capacity of up to 165 kW, and can be fully charged in under an hour assuming fast charging with a linear charging profile [115, 59]. Soares and Wang [159] envisions a 500kW fast-charging system for an airport, capable of recharging a 300 kW pushback truck battery pack in 40 minutes. As shown in Table 2.5, the varying charging rates found throughout literature have a large impact on the scheduling of an ETV fleet, and can determine whether ETVs can operate for a full day and charge overnight, or if they will need to be partially charged during the day.

Table 2.5: Battery specifications and their influence on ETV performance based on various studies. Values in the columns 'full charging time' and 'charging time for one tow' are calculated using the parameters from Baaren and Roling [16] and assuming a linear charging profile.

Study	Charging method	Charging rate	Full charging time	Charging time for one tow
Adegbohun et al. [1]	Regular charging	30 kW	28 h	66 min
Adegbohun et al. [1]	Fast charging	100 kW	8.4 h	20 min
Goldhofer [59]	Fast charging	165 kW	5.1 h	12 min
Soares and Wang [159]	Fast charging	500 kW	1.7 h	4 min

Electrical load on the network: Charging many powerful ETVs at the same time, for example with overnight charging, can be a burden on the electricity grid of an airport, especially if faster charging techniques are used. Adegbohun et al. [1] notice that fast DC charging of EVs at 50 kW and up can lead to unsustainable load spikes on the distribution grid, and could critically affect its reliability and stability. Silvester et al. [155] find that charging a thousand electric cars at EHAM would be equivalent to the total electricity peak load at the airport (2.5 MW). As can be deducted from Table 2.5, a relatively small fleet of ETVs can already be very demanding for the electricity network, depending on the charging rate. Baaren and Roling [16] calculate that electric towing for all aircraft will cost 90.4 MWh of energy at EHAM and 1.1 MWh at EHRD, without losses due to charging. For EHAM, this is equivalent to 36 hours of the 2013 peak load every day. Xiang et al. [175] take into account the available grid power and its costs in their optimization model for charging GSE at an airport, where it is used as an alternative for hydrogen fuel cell generation and battery storage systems. Lin et al. [99] include a decision variable in their MILP model to decide which charging station is connected to which power grid node, and include the maximum power such nodes can provide as a constraint.

Battery swapping: For some vehicles, such as electric cars [177, 1], electric aircraft [111, 138], and electric container transporters [149], battery swapping is being investigated as an alternative charging strategy, sometimes in combination with regular charging. Battery swapping allows one battery to be charged without being in the vehicle, while another is being used by the vehicle. The main benefit of battery swapping is that the refuelling time is comparable to that of fossil fuels, as opposed to battery charging, which can take multiple hours, depending on the used charging technology. Battery swapping can thus avoid long downtime due to charging. Another advantage of battery swapping is that peak loads on the power grid can be reduced because battery charging can be spread out during the day or night.

Yang and Sun [177] use heuristics to solve an MILP formulation for a location routing problem for battery swapping stations, for general EVs. Adegbohun et al. [1] describe the design and working of battery swapping stations for electric cars. Such facilities are already operational, for example in China for NIO cars [119].

Mitici et al. [111] investigate battery swapping for electric aircraft during turnaround,

by solving an MILP formulation to decide which batteries will be swapped, and consequently, at which charging station they will be charged. Other outputs are the fleet size, the aircraft to flight schedule, the number and location of charging stations, and the number of batteries needed. Using the combination of battery swapping and charging, three times more missions can be performed with electric aircraft than the fleet size. Salucci et al. [138] also identify the number of spare batteries needed as one of the key points for achieving smooth operations, and use simulation modelling to perform infrastructure planning for electric aircraft at airports.

Schmidt et al. [149] investigate charging strategies for charging automated guided vehicles in container terminals, and find that the best balance between high productivity, low costs and low waiting time is to use 1.6 batteries per vehicle in the charging system.

PLACEMENT OF CHARGING STATIONS

In the case that there are not enough charging stations at an airport, or they are not placed strategically, there will be vehicles that cannot perform their duties and are lining up at the charging stations. On the other hand, it is expensive to keep many charging stations operational if not all of them are used enough. This trade-off is a consideration when implementing a fleet of ETVs at an airport.

Establishing the number of charging stations: Gulan et al. [62] perform their analysis for a range of 27 to 80 electric ground support vehicles, combined with a range of 25 to 45 charging stations, for one airport terminal. Schiffer and Walther [141] show how to obtain a lower bound on the number of needed charging stations and EVs, to reduce the needed computational time for solving their MILP formulation. Hiermann et al. [67] find that in benchmark instances where normally a set of 21 charging stations was required to serve all EVs, optimizing the fleet mix leads to a situation where less then half of these stations are needed. Doctor et al. [37] makes use of discrete event simulation to determine the best number of charging stations at London Heathrow (EGLL) for electric aircraft. They consider fixed charging times of various lengths and illustrate the influence of an electric fleet on the airport throughput and turnaround times.

Deciding the location of charging stations: In the studies in Table 2.4 that consider electric taxiing or ground support equipment, the locations of charging stations are considered fixed. In other applications of EVs, the so-called location routing problem (LRP) has been investigated: Hiermann et al. [67] model both the choice of locations and the choice of the number of charging stations for EVs by inserting stations on given routes where needed, using a bi-directional labeling algorithm. Lin et al. [99] formulate an MILP model to select charging station locations for electric buses, from a list of candidate locations. The aim is total cost minimization, where factors such as facilities, transportation and grid power loss are considered. The authors use the model to select 12 charging stations from 30 candidate locations in Shenzhen, China, which hosts more than 16000 electric buses.

CHALLENGES

Charging stations on the airport surface: Table 2.4 shows that optimization regarding the charging stations has not been performed in the context of airport surface move-

ment. There the locations and amount of charging stations have been assumed given. However, the number of charging stations for ETVs can depend on many variables, such as the airport layout, runway usage, the fleet size, the vehicle electricity usage and the electric power available. A suitable location should be quickly reachable from the service road network, have sufficient space for multiple charging points, and it should be possible to connect the location to the airport electricity network with high-voltage cables [138, 37].

Multi-stage approach: As shown in Section 2.4.2, it is important to consider the development of ETV usage in the future. As more aircraft types become certified to be towed by ETVs, it is expected that airports will slowly increase the size of their ETV fleet, and consequently their need for the associated charging infrastructure. For example, Schiphol currently has a fleet of only three ETVs, but by 2030 it is envisioned that all aircraft can be towed by ETVs [144]. It is possible that a charging station location that is suitable for the ETV fleet in 2025 does not fit in the optimal charging station configuration for the ETV fleet in 2030. In order to make sure that charging stations do not need to be relocated, one can develop a multi-stage approach to the charging station placement problem for electric taxiing. Such an approach has been developed by Lin et al. [99] for electric buses, where the first stage was defined as the coming ten years, and the second stage as the twenty years thereafter. The authors obtained expected values for the number of buses and charging stations, the energy demand, and the station construction costs, for the two stages. The charging station placement problem was then solved for both stages simultaneously, prohibiting station relocation. The authors show that multi-stage optimization reduces the total cost by 17% when compared to single-stage optimization. When considering the charging station configuration for ETV, developing a multi-stage approach would require knowledge of the expected ETV fleet size and amount of charging stations during the coming years, but also other factors that might influence the charging network. For example, the introduction of the wide-body version of the ETV, which will likely have a different battery capacity and depletion rate, can be included in one of the stages.

Battery specifications: To construct a realistic charging model for ETVs, there are several factors that should be considered. In literature, the charging and depletion rate of EV batteries are most often assumed to be constant. In contrast, Goeke and Schneider [58] incorporate speed, gradient and load distribution in their model of EV energy consumption, and Mitici et al. [111] assume a bilinear charging profile for electric aircraft. When considering ETVs the batteries may also exhibit nonlinear charging behavior. The depletion rate will be influenced by vehicle speed and acceleration, but also by factors such as the outside circumstances, the weights of the aircraft that are being pulled, and the acceleration profile.

Energy demand: Silvester et al. [155] assert that achieving sufficient electrical distribution capacity is the largest bottleneck for successful operation of a fleet of EVs. Charging a fleet of ETVs at an airport is expected to require a large amount of energy. An approach to the demand for electricity for an ETV fleet should consider:

i) The power supply available at the airport during the entire day of operations. For example, Gulan et al. [62] propose a model describing a trade-off between available power supply and energy demand. Specifically, a trade-off between using gas-powered GSE and electric GSE is obtained.

ii) The charging protocol, e.g. overnight or daytime charging [99], full or partial charging [62], and fast or regular charging.

iii) The price of energy at different moments during the day/year, as well as the expected price of electricity and batteries.

When all of these aspects are integrated in a model, its results can be used to make management decisions with regard to investments in electrical infrastructure at an airport.

Alternative charging strategies: As outlined above, battery swapping technology can help avoiding ETV downtime, increasing the usability, and reducing peak loads on the power grid. Improved usability can in turn lead to a smaller fleet size, saving operating costs. On the other hand, depending on the amount of spare batteries used, it may lead to an increased peak energy load on the airport.

Another alternative charging strategy is to charge EVs using wireless power transfer technology (WPT). Rather than charging EVs at a set of charging stations, it is possible to use dynamic wireless charging (DWC) to charge EVs while driving along roads. The first commercial EV using DWC was deployed in 2009 [110], and since then a significant amount of research has been conducted towards implementing DWC for varying EVs. Many authors view DWC technology as a promising solution for future EVs [82]. Alwesabi et al. [10] develop an MILP model to determine the needed amount of electrical buses and needed wireless cable length to facilitate a given bus schedule. By placing the cables strategically along the bus routes, the electric buses can suffice with a battery of 18 kWh. Oliveira et al. [123] aggregate human factor data to determine the best location for DWC cables that serve electric taxis. Their solution involves placing cables under taxi ranks. DWC techniques could be of interest to airports as well, especially those that aim for electrification of all ground support equipment [62, 175]. Since ETVs and other GSE regularly drive along the same roads, strategically placing DWC cables under these roads can provide these vehicles with power for a significant part of their driving time. Therefore, implementing DWC technology for ETVs is expected to reduce or remove the need for charging stations at the airport, and possibly the needed battery size. This in turn may reduce the vehicle's weight, which is expected to lead to further improvement in energy use [159].

2.5. DISCUSSION AND CONCLUSION

Electric taxiing is expected to significantly contribute to the reduction of air traffic emissions, and has attracted increasing research attention. Although scientific reviews have been written from a technical and economic perspective, the existing literature has not been reviewed from an operational perspective. This study has reviewed the operational aspects of managing a fleet of electric taxiing vehicles (ETVs) at an airport and has identified challenges for future research.

In the past 10 to 15 years, multiple electric taxiing systems (ETSs) have been proposed to reduce airport emissions due to taxiing. The systems are commonly classified into on-board and external systems. On-board systems become part of an aircraft and do not require a fleet of vehicles to be routed on the airport surface. External systems require no changes to aircraft, can attain a large taxiing speed, are technically less difficult to implement, and are currently operational at airports on a small scale.

The electric taxiing procedure for an aircraft that makes a turnaround has been compared to regular operation in detail. In addition to the required changes in the procedure, time and space will have to be reserved for the connecting and disconnecting of vehicle and aircraft, for the engine-warm up of the aircraft, and for charging the vehicles. The ETV replaces the pushback truck, but will add to the traffic on the taxiways or service roads.

The challenges associated with the operational management of ETVs have been treated from the perspective of three main topics. The first topic is the routing of vehicles and aircraft. The vehicle routing problem for electric taxiing is different from regular VRPs, because of the separation requirements, the many one-way taxiways, the use of two types of roads for ETVs (taxiways and service roads) and the specific delay characteristics of air traffic. Some authors represent this problem with a simulation, allowing for sequential routing of aircraft. Others set up a Mixed Integer Linear Programming Model (MILP), which is then solved for a full day. The simulation approach allows for a more straightforward conflict avoidance, while the MILP approach requires many additional constraints. The assumed values of several key parameters differ across the literature: taxiing separation distance, speed, and engine warm-up time and place. Routing challenges identified are as follows: first, there are possibilities for airports to adjust the current rules and procedures on the airport surface to better facilitate swift electric taxiing. Second, the increased workload posed by the necessary management of the ETV fleet on Air Traffic Control must be addressed, for example with a dedicated ETV manager, or in the long term by implementing Advanced Surface Movement Guidance and Control Systems (A-SMGCS).

The second topic is the assignment of vehicles to aircraft. Typical objectives from literature are to minimize taxiing time or used fuel. Factors that have a large influence on these objectives are the size of the ETV fleet, the manner of conflict avoidance, and the instances where the aircraft taxi in the regular way. An important scheduling challenge is dealing with airside disruptions, such as flight delays, which can cause disruptions of the ETV schedule and increased workload of personnel, since the ETVs need travel time to arrive on time at gates and runways around the airport. The development of robust scheduling algorithms and disruption management procedures for the airport surface movement can reduce the effects of these airside disruptions. Another challenge is to solve the scheduling problem with realistic technological specifications of external ETSs, so that the expected performance can be modelled accurately, and possible bottlenecks can be identified. Airside disruptions present a relevant challenge for this type of ground operations, since effective operation requires the vehicles to be present at the correct gate or runway at the correct (disrupted) time, while the fleet is likely spread around the airport during operation. This is in contrast to other ground vehicles, which operate mainly in the gate areas.

The last topic is comprised by the complications due to the electrical aspect of ETVs. There have been but a few operational management approaches to ETVs that include this aspect. Some subjects that are interesting for ETVs have been treated for other types of electric vehicles: Typical objectives are to minimize the number of vehicles or charging stations, or the taxiing distance. Several characteristics of the problem such as the charging period, the influence on the electrical network and possibilities for battery swapping are topics of interest in current research. The main challenge is to apply the optimization problems with their characteristics as reviewed here to the problem with ETVs. Specifically, the location routing problem for charging stations has not been attempted for ETV fleets to the knowledge of the authors. Possible additions to these optimization problems are to devise a multi-stage approach for the electric infrastructure, to make sure the demand for electrical power can be met by the airport, and to use realistic battery specifications. Last, alternative charging techniques such as dynamic wireless charging can be of interest for the ETV charging problem.

Overall, we have seen that important research steps have been taken in the implemention of external electric taxi systems, but numerous research directions and challenges remain. Addressing these challenges will help the industry move to large-scale ETV implementation in the next decades and thereby hopefully significantly reduce airport ground emissions.

3

Fleet Scheduling for Electric Towing of Aircraft under Limited Airport Energy Capacity

One of the operational challenges regarding electric taxiing implementation identified in Chapter 2, is the fact that making the transition to electric taxiing operations is expected to significantly increase the electricity demand at airports. In this chapter, a mixed-integer linear programming (MILP) model to schedule electric vehicles for aircraft towing and battery charging is developed. This model considers limits for the supply of energy at an airport. In addition to the MILP model, an Adaptive Large Neighbourhood Search model is developed, to identify time-efficient scheduling solutions. The models are applied to a large airport case study, and conclusions are drawn regarding the optimal charging strategy given airport electricity capacity and electric vehicle properties. The models provide support for infrastructure planning of airports during the transition to aircraft electric taxiing.

This chapter is based on the following research article:
Zoutendijk, M. & Mitici, M., "Fleet Scheduling for Electric Towing of Aircraft under Limited Airport Energy Capacity", in *Energy, Special Issue "The Role of Smart Technologies in Energy Engineering"*, 294, p.130924 [190].

3.1. Introduction

The aerospace industry has committed to reducing net greenhouse gas emissions to zero in the USA and to 10% of 1990 emissions in the EU by 2050 [49, 47]. In addition to emissions produced while flying, the aerospace industry also produces ground-based emissions. Electric taxiing is a promising technique for reducing these emissions. In this work, the focus is on external electric taxi systems, where an electric towing vehicle (ETV) tows aircraft from gates to runways and vice versa. Electric towing vehicles are currently operational at several airports, and are under further development [157]. Their implementation is expected to reduce taxiing fuel use by up to 80% and thus reduce the airport emissions of greenhouse gases [36, 63, 170, 103, 145]. This is not only beneficial for the aerospace emission goals, but also improves air quality and reduces noise pollution for airport personnel, passengers and residents of airport surroundings [65].

The technical feasibility of the external Electric Taxiing System (ETS) has been investigated in literature, and demonstrated during early implementation. The next step is to move towards large scale implementation, and to find out how many towing vehicles are needed for seamless surface movement. This requires airport infrastructural planning and a strategy for operational management of large fleets. Vehicle operation needs to be scheduled, taking into account airport routing, flight schedules and electricity use. A model is needed that creates a daily towing schedule for ETVs to aircraft. Such a model can be used by airport operators to manage current or future ETV fleets at airports, and also by airport planners to consider the infrastructural requirements needed. Below we review existing studies addressing comparable problems in other domains.

Scheduling a fleet of electric vehicles

Several studies have developed models to schedule generic electric vehicles (EVs) for operations and charging. Hiermann et al. [67] consider a routing problem with a mixed fleet of EVs, where vehicles are assigned to customers with time windows. They optimize for the best fleet composition, choices of recharging moments and locations. Routing is done using labeling algorithms, and scheduling using branch-and-price, but also a heuristic based on Adaptive Large Neighbourhood Search (ALNS). Schiffer and Walther [141] solve a Mixed Integer Linear Programming (MILP) location routing problem to find charging station locations, while optimizing for traveling distance, fleet size and number of stations. Keskin and Çatay [87] apply ALNS to schedule EV tasks with time windows, and allow partial recharging. In addition to customer removal and insertion algorithms, the authors also introduce removal and insertion algorithms for visits to charging stations. Emde, Abedinnia, and Glock [40] develop heuristics to schedule a fleet of EVs performing round trips including recharging breaks. They aim to minimize the fleet size and maximize fairness in workload for EV operators. Their neighbourhood search heuristic based on operations 'push' and 'swap' performs best and can solve problem instances in a few minutes. Frey et al. [54] solve a vehicle routing and scheduling problem with customer time windows. In addition to a branch-price-and-cut algorithm introduced in earlier work, they introduce an ALNS approach. Their approach allows moving to infeasible solutions, and penalizes such solutions in the objective function. Several special removal operators based on the spatial arrangement of customer locations are introduced. Last, Foda et al. [51] develop a generic optimisation model for electric bus fleets,

with a very broad approach, taking into account a mixture of multiple objectives. Since the authors infer that all system parameters should be included in their model as system parameters, it outputs include charging schedules, but also battery properties and optimal charging infrastructure parameters.

ETV FLEET SCHEDULING APPROACHES

In recent years, several authors have started to research the routing and scheduling challenges connected to ETV implementation. An important difference between EV and ETV scheduling is the split between driving and towing: the speed, road usage, conflict avoidance and time constraints are different when driving an ETV compared to towing an aircraft. These aspects introduce additional constraints and complications that need to be taken into account when creating a realistic representation of airport surface movement with electric taxiing.

One of the first works concerning ETV scheduling is Baaren and Roling [16], who create an MILP model minimizing fuel consumption. The model generates trips that can be performed on one battery charge. The model is applied to two airports. The fuel savings and electricity costs are calculated, and the minimum fleet size is ascertained. Soltani et al. [160] develop an MILP model to assign diesel-powered towing vehicles to 205 aircraft, and introduce extra variables and constraints to ensure conflict and collision avoidance, while minimizing the sum of taxiing delay costs, maintenance costs and labor costs. It is found that the optimal number of vehicles for Montreal-Trudeau International Airport (CYUL) is twelve, reducing taxiing fuel consumption by 95%. Salihu, Lloyd, and Akgunduz [137] develop a discrete event simulation for scheduling a year of ETV operation, with the goal of modelling the taxiway congestion that can be expected from using electric taxiing. The taxi routes are calculated in advance. It is assumed that all charging can be done during the night. It is found that electric taxiing leads to a longer taxiing time, as the ETVs taxi slower than the aircraft, and the aircraft have to wait after requesting an ETV. Building on earlier work, Zaninotto, Gauci, and Zammit [181] create a real-time simulation to schedule ETVs to aircraft. Conflict avoidance is performed by slightly delaying aircraft (less than 3 min) where necessary, before selecting the ETV. The model minimizes taxiing delays and route lengths. The ETV state of charge and recharging are considered throughout the simulation. More than 80% of flights were assigned to an ETV, with a fleet size of 25% of the airfield hourly traffic. The tow truck utilisation time was 30%. Also building on their earlier work, Oosterom, Mitici, and Hoekstra [125] develop an MILP and a greedy model to dispatch multiple types of ETVs to aircraft. Deciding which ETVs charge when is based on the residual state-of-charge of the ETVs. The objective of the models is to minimize the fleet size required to tow all flights in the daily flight schedule. A 5% optimality gap is obtained by the greedy model compared to the MILP approach. Applying both models in a rolling horizon approach and using a flight schedule including historical flight delays, the authors show that 95% of flights can still be towed by ETVs. Last, Ahmadi and Akgunduz [3] develop a one hour rolling horizon MILP model with the goal of showing the best fleet size to be purchased by airports. The model minimizes taxiing delay, total taxiing time and fuel consumption. Aircraft can either be towed or taxi in the regular way. The rolling horizon approach allows for solving realistic size problems in foreseeable time, and accommodating for flight disruption during the day. Charging stations and charging times are not considered in the model.

Transport electrification at airports

The transport industry has far-reaching ambitions for electrification, e.g. the EU will only allow zero-emission vehicles from 2035 onwards [48]. There has been research interest into the electrical infrastructure requirements to meet such ambitions.

Some authors investigate the option of supplying the electrical grid with energy from idle EVs, i.e. electric vehicle-to-grid delivery (V2G), such as Mahmud, Hossain, and Ravishankar [104] and Li et al. [98]. In a review paper, Uddin et al. [168] identify this as one of the three major strategies considered in literature for peak load shaving, along with demand side management, and making use of energy storage systems (ESS). However, rather than using EV charge to mitigate demand peaks, such peaks can also occur when charging these EVs. Solutions suggested in literature include ESSs, battery swapping [1] and time-based electricity pricing [174]. Generally, it is expected that the electricity demand will remain a bottleneck for certain electrification developments. For example, Forrest et al. [52], a large scale study of energy infrastructure requirements for EVs in California, find that smart charging technology such as V2G and smart energy storage facilities can help provide in energy demands at off-peak hours. However, they stress that such techniques would still require a large excess of renewable energy generation in the first place. Moon et al. [112] estimate future electricity demand due to EVs in South Korea and determine where and at which time demand peaks will appear. They predict that current power grid infrastructure in parts of the country may not be able to cover the predicted demand.

The electricity demand at airports without considering transport electrification stems mainly from HVAC systems and lighting, report Ortega Alba and Manana [127], who describe the main airport energy sources and consumers, and suggest ways to reduce electricity consumption at airports. Uysal and Sogut [169] apply a holistic architecture-based approach to airport energy demand and report large potential savings for light and thermal management in terminal buildings. In Ortega Alba and Manana [126] energy demand patterns at Santander Airport (LEXJ) are characterized and analyzed on a daily and yearly basis, by studying their electrical load profiles. Current peak electricity demand for this medium-sized airport does not exceed 600 kW.

However, when accounting for transport electrification, more and more systems start to make their appeal on the airport electricity capacity. Electric vehicles charging at airport parking spaces, electrified ground support equipment [62], and electric aircraft [37] are all expected to contribute to the future electricity demand, while capacity may not increase as quickly as necessary. Capacity bottlenecks could appear due to the electricity infrastructure at the airport, but also due to the local power grid. It suggests that it is wise to take into account the possibility of operating an airport that might reserve only limited electricity capacity for electric towing, while it is transitioning to sustainable airport surface movement.

In order to gain insight into the possibilities of electric taxiing under such circumstances, it is necessary to create a scheduling and charging model that can adhere to electricity capacity constraints. The model should closely monitor state-of-charge of ETVs and decide when and where they are charged. The schedule created by this model should cover a full day of operation, to find the best schedule given specific daily electricity capacity profiles for ETV charging. In addition, to emulate realistic airport scenarios,

the model should incorporate practical considerations such as a realistic airside airport representation, ETV connecting/disconnecting times and conflict avoidance.

To the best of our knowledge, there is no previous work addressing the issue of *the effect of limited electricity capacity at airports on the expected emission benefits of electric towing*. In this chapter we propose two models to assign ETVs to aircraft that take limited electricity capacity at airports into account. The main contributions of this chapter are:

- We propose an MILP formulation for the ETV-to-aircraft assignment problem, that takes into account a limited electricity capacity at an airport. The model also tracks the state of charge and electricity demand of every ETV throughout the day, which is omitted in the majority of previous works.
- We propose an ALNS method including tailored removal/insertion heuristics to obtain time-efficient solutions of the full-day ETV-to-aircraft assignment problem.
- We apply our models for a large airport, and for various ETV electricity capacity profiles. We investigate the effects of having a limited electricity capacity on the ETV operations for these profiles.

The remainder of this chapter is divided as follows: Section 3.2 introduces the ETV scheduling problem and the input data and parameters required. In Section 3.3 a method to calculate the emissions savings from the electric towing distance is described. Section 3.4 states the linear programming formulation describing the scheduling problem and Section 3.5 describes how this problem is solved using ALNS. Local search heuristics and notation for the ALNS algorithm are defined, as well as the local search framework. In Section 3.6 both models are applied to various instances of the ETV scheduling problem, to determine their efficacy, as well as the influence of ETV electricity capacity profiles and ETV battery properties on the results. Section 3.7 provides concluding remarks and future research directions.

3.2. PROBLEM DESCRIPTION AND FORMULATION

In this section the problem of scheduling ETVs to tow aircraft is defined, and the model developed to create the schedule is presented. The schedule consists of a *towing schedule* and a *charging schedule*. The former defines for each vehicle when it tows which aircraft, and the latter defines for each vehicle when and where it will be charged. An arriving or departing aircraft either performs regular taxiing (using the jet engines) or electric taxiing (being towed by an ETV). The ETVs tow aircraft, and charge at charging stations to replenish their battery. A limited amount of electricity capacity is assumed available at the airport for the charging of ETVs. The scheduling problem is therefore extended with constraints that track the electricity demand and capacity at discrete time steps. The problem is formulated as a Mixed Integer Linear Programming (MILP) model.

AIRPORT LAYOUT

We assume a taxiing system consisting of service roads (used only by ETVs at speed v^S) and taxiways (used by aircraft or ETV + aircraft combinations at speed v^X). The taxiing system is represented by a graph G, which is the union of a taxiway graph $G^X = (N^X, E^X)$

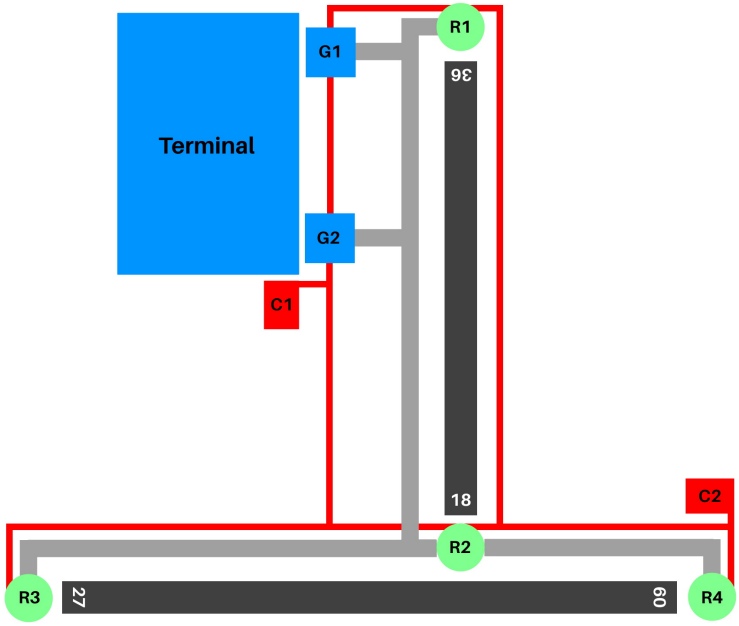


Figure 3.1: Schematic overview of an example airport. Two runways are shown in black. Gray lines indicate taxiways (E^X) and thin red lines indicate service roads (E^S). Runway entry/exit points R1-R4 are indicated with green circles. Gate group nodes G1-2 are indicated with blue squares. Charging stations C1-2 are indicated with red rectangles.

and a service road graph $G^S = (N^S, E^S)$. The edges correspond to the service roads and taxiways, and the nodes correspond to junctions, gate groups or runway entrance/exit points. The airport is assumed to have N_{cs} charging stations for the ETVs, one of which also has the function of ETV depot n_{dp}. Figure 3.1 shows an example airport with the road types indicated, as well as nodes for gates, runways and charging stations.

Aircraft and ETV routing

The routes taken by the aircraft and vehicles are calculated in advance. Distances from any node m to n are calculated using Dijkstra's shortest path algorithm, and are indicated on G^X with $d^X(m,n)$, and on G^S with $d^S(m,n)$. The schedule is created for a time period P, which spans the interval $[t^s, t^e]$. The set of $N^{F,S}$ aircraft arriving or departing at the airport within this period form the set A^S. From the flight schedule, the scheduled landing time (SLDT) of arriving flights and the scheduled off-block time (SOBT) of departing flights are collected. They correspond to the pick-up time t_a^p of aircraft a, the moment an ETV starts towing the aircraft. The moment an ETV stops towing the aircraft

is referred to as the drop-off time t_a^d, and is calculated as:

$$t_a^d = t_a^p + d^X(n_a^p, n_a^d)/v^X = t_a^p + t^X(n_a^p, n_a^d), \quad \forall a \in A^S, \tag{3.1}$$

with $t^X(n_a^p, n_a^d)$ the towing time for aircraft a, and n_a^p and n_a^d its pick-up and drop-off nodes, respectively. They correspond to the gate or runway where the pick-up or drop-off takes place, which are taken from the flight schedule. The towing distance for any aircraft $a \in A^S$ is denoted with the shorthand $d_a^X = d^X(n_a^p, n_a^d)$.

All ETVs start and end the period P at the depot n_{dp}. For every towing task, an ETV v assigned to aircraft a travels to the pick-up node n_a^p, and arrives at $t_a^p - t^c$. There the ETV connects to the aircraft, which takes a time of t^c. Between t_a^p and t_a^d the ETV tows aircraft a, and upon arrival the ETV disconnects from the aircraft until $t_a^d + t^c$. Then the ETV travels towards its next towing task or a charging station. The travel time for a non-towing ETV from node m to n is given as:

$$t^S(m,n) = d^S(m,n)/v^S, \quad \forall m,n \in N^S \tag{3.2}$$

The fleet of ETVs is denoted as V and has size N^V. A more detailed description of the electric towing procedure can be found in Chapter 2 of this dissertation.

Aircraft route conflicts need to be avoided. We say an aircraft has a route conflict when a) an aircraft is using the same node or edge as an another aircraft at the same time, or b) an aircraft is violating the minimum separation time t_{sep}: the taxiing time between it and an aircraft taxiing in front of it. All routes of aircraft in A^S are deconflicted before making the ETV-to-aircraft schedule, as follows:

The routes for all aircraft in A^S are calculated. Then, we iterate through the time period P in time steps of size t_{sep}. At any time step, for every aircraft $a \in A^S$, we check whether another aircraft is planned to a) occupy the same node as aircraft a or b) approach aircraft a on a one-directional edge in the set of taxiway edges E^X. In that case aircraft a is held at its previous node until it can proceed without a conflict with any other aircraft and the separation time is respected.

It is assumed that there is no taxiing delay for the aircraft, except due to deconflicting routes. For the ETVs it is assumed that deconflicting is not required, since the service roads are assumed to provide enough opportunities for passing other vehicles, and no separation time is required. For the same reasons, other vehicles using the service roads are assumed not to influence the ETV operations.

CHARGING AND CONSUMPTION MODULE FOR ELECTRIC TOWING VEHICLES (ETVS)

The state of charge (SOC) of each ETV v is tracked during the time period P. This requires a battery capacity Q, determined by the specifications of the ETV, a charging rate P^c, determined by the specifications of the ETV and the charging station used, and a discharging rate. The discharging rate can be obtained by considering the power needed to drive an ETV, and to tow an aircraft. This power can be found using the mass, velocity and rolling resistance of the vehicle and aircraft. The power consumed by an ETV during towing at speed v^X is denoted as P^X, and the power consumed during driving at speed v^S is denoted as P^S. In this work the ETVs are assumed to travel at constant speed, without accounting for acceleration and deceleration. The energy consumption of the ETV for any (part of a) route is found using the power and the time taken to traverse the route.

In addition to the characteristics of the ETV, we define the electricity capacity of the airport charging network. The electricity that is available at the airport for various electrical processes throughout the day is referred to as the *electricity capacity profile*. The processes unrelated to ETV charging require a part of the electricity capacity. Subtracting this demand from the electricity capacity profile yields the *electricity capacity profile for ETVs*. This is the airport electricity capacity that is specifically available to charge ETVs. In order to keep track of the electricity demand at different times of day we divide the time period P in $N^T = 144$ time steps t of length $\Delta = 10$ min. The start of each time step is the time associated with the time step, e.g. $T_0 = t^s$ and $T_{N^T-1} = t^e - \Delta$. The electricity capacity for ETVs at time step t is then denoted as C_t.

3.3. EMISSIONS SAVED BY ELECTRIC TAXIING

Several studies determine the amount of fuel or emissions spent using aircraft taxiing and other airport surface movement, e.g. [36, 39, 16]. In this section, we derive the amount of emissions saved per kilometer of self-taxiing that is replaced by electric towing. We consider the following assumptions:

- When calculating the emissions avoided by using electric towing, we consider only the emissions avoided by consuming less jet fuel. In this work we focus on CO_2 emissions.
- We consider towing vehicles that are electric battery-powered.
- The amount of emissions per km taxiing varies with the size of the aircraft. In this section, we calculate a value based on narrow-body aircraft, which is a lower bound for the actual value considering the mix of aircraft at the airport.
- When calculating the energy spent by ETVs, we consider only the energy spent while towing aircraft, since towing aircraft takes many times more energy than driving. This can be seen by comparing towing and driving power in Section 3.6.

GROSS TAXI EMISSIONS SAVED

We are interested to find an estimation for the emissions saved by towing an aircraft for 1 km instead of it self-taxiing that distance. We start with finding the fuel spent while self-taxiing an aircraft.

From the work of Zhang et al. [187] we obtain that an A320 aircraft spends 88.0 kg of jet fuel when taxiing a 2.5 km route at Shanghai Pudong Airport (ZSPD). In this calculation, acceleration, deceleration and idling have been included. Using these values leads to a value of 35.2 kg of jet fuel per km of taxiing for a narrow-body aircraft.

It is known that using jet fuel in an aircraft engine leads to 3.16 kg CO_2 per kg of jet fuel, see e.g. Nojoumi, Dincer, and Naterer [120]. This leads to an emission saving of 111 kg CO_2 per km of taxiing.

$$C^G_{\text{km}} = C_{\text{kgCO}_2} F_{\text{tow}} / d = 111 \text{ kg CO}_2, \tag{3.3}$$

where C^G_{km} is the gross amount of emissions saved per km taxiing, C_{kgCO_2} the amount of emissions per kg jet fuel, F_{tow} the average amount of jet fuel used during one towing event, and d the taxied distance.

NET TAXI EMISSIONS SAVED

From the flight schedule for 27-12-2021 at Schiphol [147] we obtain an average taxiing distance of 3.79 km. From the energy consumption described in Sec 3.2, we find that the energy needed for an average tow of a narrow-body aircraft is 19.9 kWh. From Scarlat, Prussi, and Padella [139] we obtain that the carbon intensity of electricity generation in Europe was 0.334 kgCO2/kWh in 2019. This means that during the generation of the electricity needed to tow an aircraft for 1 km, 1.75 kg CO_2 is emitted. In summary, under the assumptions described above, we obtain that for each km of electric towing, 109 kg CO_2 is saved when comparing the process of electric taxiing to self-taxiing. In this chapter, we will use this value to calculate emission savings obtained from distances travelled using electric towing.

$$C^N_{\text{km}} = C^G_{\text{km}} - C^E_{\text{km}} = C^G_{\text{km}} - \frac{E_{\text{tow}} E_{\text{el}}}{d_{\text{av}}} = 111 - 1.75 = 109 \text{ kg } CO_2, \tag{3.4}$$

where C^N_{km} is the net amount of emissions saved per km taxiing, C^E_{km} the amount of emissions spent generating electricity per km taxiing, E_{tow} the energy needed for an average narrow-body aircraft tow, d_{av} the average towing distance per aircraft movement, and E_{el} the carbon intensity of electricity generation.

TOTAL FLIGHT EMISSIONS

In addition to estimations of the absolute amount of emissions saved using this technology, it is insightful to obtain an estimate for the effect of optimizing the emission savings of electric towing on the total flight emissions. An example of total fuel consumption of a narrow-body, medium-haul flight is given in Kollmuss [93]: an A320 aircraft on a flight between Los Angeles and New York uses 11.6 tons of jet fuel. If we assume two taxiing events similar to the one at Shanghai Airport, and note again that CO_2 emissions are directly proportional to fuel consumption through the factor 3.16, we obtain that using electric taxiing saves 1.5% of the total flight emissions for this flight example:

$$\frac{C_{\text{saved}}}{C_{\text{total}}} = \frac{2F_{\text{tow}}}{F_{\text{total}}} = \frac{2 \cdot 88.0}{11.6 \cdot 10^3} = 1.5\%, \tag{3.5}$$

with $\frac{C_{\text{saved}}}{C_{\text{total}}}$ the fraction of saved total flight emissions, and F_{total} the total fuel consumption spent on the indicated flight.

3.4. MATHEMATICAL FORMULATION FOR VEHICLE-TO-AIRCRAFT SCHEDULING

In this section we formulate the mixed-integer linear program (MILP) for ETV-to-aircraft assignment, adapted from Oosterom and Mitici [124]. Note that a glossary with terms and notation used in this and other chapters can be found at the end of the dissertation.

For completeness, the assumptions applicable for the scheduling models in this chapter are summarized below:

- Aircraft deconfliction is performed in advance of scheduling.
- Taxiing and towing happens on taxiways and ETV driving happens on service roads.

- There is one ETV type for all aircraft.
- Any aircraft can perform electric taxiing or regular taxiing.
- The ETVs have a linear charging profile.
- All ETVs are recharged to capacity before the end of the time period P.
- The electricity capacity for ETVs is limited to C_t.

Now, we proceed to the MILP formulation. First, we introduce a set of artificial aircraft A^e, with one aircraft for each vehicle. The aircraft $a \in A^e$ have $n_a^p = n_a^d = n_{\text{dp}}$, $d_a^X = 0$ and $t_a^d = t_a^p = t^e$. Such artificial aircraft are necessary to enforce a state of charge value for every vehicle v at the end time t^e. All artificial aircraft will always be scheduled for electric towing, at no cost. The set of all aircraft will be denoted as $A = A^S \cup A^e$ with size $N^F = N^{F,S} + N^V$. Furthermore, we introduce notation describing various quantities of energy, required in the MILP:

$$q^X(a) = P^X t^X(n_a^p, n_a^d) \qquad \forall a \in A, \tag{3.6}$$
$$q^S(n,m) = P^S t^X(n,m) \qquad \forall m,n \in N^S, \tag{3.7}$$
$$q^S(a,b) = q^S(n_a^d, n_b^p) \qquad \forall a,b \in A, \tag{3.8}$$
$$q_f^S(a) = q^S(n_{\text{dp}}, n_a^p) \qquad \forall a \in A, \tag{3.9}$$
$$q^C(a,b) = \min_{i \le N_{\text{cs}}}\{q^S(n_a^d, n_{\text{cs},i}) + q^S(n_{\text{cs},i}, n_b^p)\} \qquad \forall a,b \in A, \tag{3.10}$$
$$q_1^C(a) = \min_{i \le N_{\text{cs}}}\{q^S(n_{\text{cs},i}, n_a^p)\} \qquad \forall a \in A, \tag{3.11}$$
$$q_2^C(a) = \min_{i \le N_{\text{cs}}}\{q^S(n_a^d, n_{\text{cs},i})\} \qquad \forall a \in A, \tag{3.12}$$
$$t^C(a,b) = \max(t_b^p - t_a^d - t^S(n_a^d, n_b^p) - 2t^c, 0) \qquad \forall a,b \in A, \tag{3.13}$$

with $q^X(a)$ the energy needed to tow aircraft a on G^X, $q^S(n,m)$ the energy needed by an ETV to travel from node n to m on G^S, $q^S(a,b)$ the energy needed by an ETV to travel from the dropoff point of aircraft a to the pickup point of aircraft b on G^S, $q_f^S(a)$ the energy needed by an ETV to travel from the depot n_{dp} to the pickup point of aircraft a, $q^C(a,b)$ the minimal energy needed by an ETV to travel from the dropoff point of aircraft a to the pickup point of aircraft b on G^S, via a charging station $n_{\text{cs},i}$, $q_1^C(a)$ the energy needed by an ETV to travel from the closest charging station to the pickup point of aircraft a, $q_2^C(a)$ the energy needed by an ETV to travel from the dropoff point of aircraft a to the closest charging station, and $t^C(a,b)$ the time between towing consecutive aircraft a and b that is freely available to the ETV towing them.

Using $t^C(a,b)$ and the set of aircraft A, we define:

$$A_a^{\text{out}} = \{b \in A : t^C(a,b) > 0\} \qquad \forall a \in A, \tag{3.14}$$
$$A_a^{\text{in}} = \{b \in A : t^C(b,a) > 0\} \qquad \forall a \in A, \tag{3.15}$$
$$A_a^{\text{PC}} = \{b \in A_{\text{out}}(a) : q^C(a,b) - q^S(a,b) < P^c(t^C(a,b) - t_{\min}^C)\} \qquad \forall a \in A. \tag{3.16}$$

Here A_a^{out} is the set of aircraft that can be towed by an ETV after it tows aircraft a, A_a^{in} is the set of aircraft that can be towed by an ETV before it tows aircraft a, and A_a^{PC} is the set of aircraft that can be towed by an ETV after it tows aircraft a and for which there is at least t_{min}^C time in between for effective charging, which is charging that occurs after the energy loss due to the rerouting to the charging station has been replenished. The time t_{min}^C is called the *minimum charging time*. Note that $A_a^{\text{PC}} \subseteq A_a^{\text{out}}$.

For brevity, define v_a as the vehicle v that tows aircraft a. Furthermore, we define $M \in \mathbb{R}$ as a large number. We consider the following decision variables:

$$x_{ab} = \begin{cases} 1 & \text{if } a, b \in A \text{ are towed consecutively} \\ 0 & \text{else} \end{cases} \tag{3.17}$$

$$x_a^f = \begin{cases} 1 & \text{if } a \in A \text{ is the first aircraft an ETV tows} \\ 0 & \text{else} \end{cases} \tag{3.18}$$

$$x_a^l = \begin{cases} 1 & \text{if } a \in A \text{ is the last aircraft an ETV tows} \\ 0 & \text{else} \end{cases} \tag{3.19}$$

$$q_a \in [q^X(a), Q] \quad \text{ETV state of charge at the start of towing } a \in A \tag{3.20}$$

$$c_a = \begin{cases} 1 & \text{if after towing aircraft } a, \text{ the ETV travels to a charging station and is charged} \\ 0 & \text{else} \end{cases} \tag{3.21}$$

$$c_a^t \in [0, Q/P^c] \quad \text{charging time of ETV } v_a \tag{3.22}$$

$$c_a^s \in P \quad \text{start time of charging of ETV } v_a \tag{3.23}$$

$$\alpha_{at} = \begin{cases} 1 & \text{if charging of ETV } v_a \text{ starts earlier than timestep } t \\ 0 & \text{else} \end{cases} \tag{3.24}$$

$$\beta_{at} = \begin{cases} 1 & \text{if charging of ETV } v_a \text{ finishes later than timestep } t \\ 0 & \text{else} \end{cases} \tag{3.25}$$

$$\gamma_{at} = \begin{cases} 1 & \text{if ETV } v_a \text{ is charged during timestep } t \\ 0 & \text{else} \end{cases} \tag{3.26}$$

$$y_a = \begin{cases} 1 & \text{if } a \in A \text{ is towed by an ETV} \\ 0 & \text{if } a \in A \text{ is taxiing by itself} \end{cases} \tag{3.27}$$

The objective function and constraints are given by:

$$\max_{x,q,y,c} \quad \sum_{a\in A} d_a^X y_a, \tag{3.28}$$

$$\text{s.t.} \quad x_a^f + \sum_{b\in A_a^{\text{in}}} x_{ba} = y_a \qquad \forall a \in A, \tag{3.29}$$

$$\sum_{b\notin A_a^{\text{in}}} x_{ba} = 0 \qquad \forall a \in A, \tag{3.30}$$

$$x_a^l + \sum_{b\in A_a^{\text{out}}} x_{ab} = y_a \qquad \forall a \in A, \tag{3.31}$$

$$\sum_{b\notin A_a^{\text{out}}} x_{ab} = 0 \qquad \forall a \in A, \tag{3.32}$$

$$q_a \le x_a^f (Q - q_f^S(a)) + Q(1 - x_a^f) \qquad \forall a \in A, \tag{3.33}$$

$$q_a \ge x_a^f (Q - q_f^S(a)) - Q(1 - x_a^f) \qquad \forall a \in A, \tag{3.34}$$

$$q_b \le q_a - x_{ab}(q^X(a) + q^S(a,b)) + Q(1 - x_{ab}) \qquad \forall a \in A, b \in A_a^{\text{out}} \setminus A_a^{\text{PC}}, \tag{3.35}$$

$$q_b \ge q_a - x_{ab}(q^X(a) + q^S(a,b)) - Q(1 - x_{ab}) \qquad \forall a \in A, b \in A_a^{\text{out}} \setminus A_a^{\text{PC}}, \tag{3.36}$$

$$q_b \le q_a - x_{ab}(q^X(a) + q^C(a,b)) + (1 - c_a)(q^C(a,b) - q^S(a,b)) + P^c c_a^t + Q(1 - x_{ab}) \qquad \forall a \in A, b \in A_a^{\text{PC}}, \tag{3.37}$$

$$q_b \ge q_a - x_{ab}(q^X(a) + q^C(a,b)) + (1 - c_a)(q^C(a,b) - q^S(a,b)) + P^c c_a^t - Q(1 - x_{ab}) \qquad \forall a \in A, b \in A_a^{\text{PC}}, \tag{3.38}$$

$$N^V \ge \sum_{a\in A} x_a^f, \tag{3.39}$$

$$y_a = 1 \qquad \forall a \in A^e, \tag{3.40}$$

$$q_a \ge q^e \qquad \forall a \in A^e, \tag{3.41}$$

$$c_a^s \ge t_a^d + t^c + q_2^C(a)/P^S - M(1 - c_a) \qquad \forall a \in A, \tag{3.42}$$

$$c_a^s + c_a^t \le t_b^p - t^c - q_1^C(b)/P^S + M(1 - x_{ab}) + M(1 - c_a) \qquad \forall a \in A, b \in A_a^{\text{out}}, \tag{3.43}$$

$$c_a^t \ge t_{\min}^C c_a \qquad \forall a \in A, \tag{3.44}$$

$$c_a^t \le \sum_{b\in A_a^{\text{PC}}} x_{ab} t^C(a,b) \qquad \forall a \in A, b \in A_a^{\text{PC}}, \tag{3.45}$$

$$c_a \ge M^{-1} c_a^t \qquad \forall a \in A, \tag{3.46}$$

(3.47)

$$c_a^s \leq P_t + \Delta + M(1-\alpha_{at}) \qquad \forall a \in A, t \in P, \tag{3.48}$$
$$c_a^s \geq P_t + \Delta - M\alpha_{at} \qquad \forall a \in A, t \in P, \tag{3.49}$$
$$c_a^s + c_a^t + M(1-\beta_{at}) \geq P_t \qquad \forall a \in A, t \in P, \tag{3.50}$$
$$c_a^s + c_a^t - M\beta_{at} \leq P_t \qquad \forall a \in A, t \in P, \tag{3.51}$$
$$\gamma_{at} \geq \alpha_{at} + \beta_{at} - 1 \qquad \forall a \in A, t \in P, \tag{3.52}$$
$$\gamma_{at} \leq \alpha_{at} \qquad \forall a \in A, t \in P, \tag{3.53}$$
$$\gamma_{at} \leq \beta_{at} \qquad \forall a \in A, t \in P, \tag{3.54}$$
$$P^c \sum_{a \in A} \gamma_{at} \leq C_t \qquad \forall t \in P. \tag{3.55}$$

The objective (3.28) is to maximize the total distance towed by the ETV fleet. Constraints (3.29) and (3.31) ensure that every aircraft a that is towed by an ETV, is either the first (last) to be towed by an ETV, or has another aircraft b preceding (following) it. Constraints (3.30) and (3.32) ensure that any aircraft b cannot precede (follow) aircraft a, if $b \notin A_a^{\text{in}}$ (A_a^{out}). Constraints (3.33) and (3.34) set the state of charge of all vehicles at the start of their first towing task. Constraints (3.35)-(3.38) set the new state of charge of a vehicle after towing aircraft a. Constraints (3.35)-(3.36) concern towing an aircraft without charging afterwards and Constraints (3.37)-(3.38) concern towing and charging. Constraint (3.39) enforces the fleet size of N^V. Constraints (3.40)-(3.41) define the artificial flights in A^e. Constraints (3.42)-(3.43) set the earliest time for a charging period to start and the latest time for it to end. Constraints (3.44)-(3.45) set the bounds for the charging time, based on the minimum charging time t_{min}^C and the maximum possible charging time for the aircraft a and b. Constraint (3.46) sets the charging indicator c_a to 1 if the charging time c_a^t is strictly positive. Finally, Constraints (3.48)-(3.54) keep track of the time steps during which an ETV is being charged, so that this can be limited to the electricity capacity for ETVs at time step t with Constraint (3.55).

The MILP is solved using Gurobi. Recall that the goal is to create an ETV-to-aircraft schedule for a full day. The number of constraints in the model is bounded by $9N^F + 6N^{F2} + 2N^V + 7N^F N^T + N^T$. The MILP model is expected to have a large runtime. An effective way to reduce the number of constraints is to adapt Equations (3.14)-(3.15) in such a way that an aircraft b that is scheduled more than e.g. a few hours later than aircraft a will not appear in A_a^{out}. Nevertheless, the number of constraints grows roughly quadratically with the number of flights N^F. Solving the MILP model is feasible for time periods P of a few hours, but becomes intractable for periods longer than 8 hours. Another approach is necessary, and such an approach will be introduced in the following section.

3.5. AN ALNS APPROACH TO ELECTRICITY CAPACITATED ETV SCHEDULING

The MILP formulation introduced in Section 3.4 is not able to provide a solution for problem instances of a full day. To find such solutions, we present here a heuristics-based approach to the ETV-to-aircraft assignment problem. The approach is based on the framework of Adaptive Large Neighborhood Search (ALNS), originally developed by

Ropke and Pisinger [135].

ALNS algorithms work by removing a relatively large part of a given solution, and then building a new solution with new values. Which part of the solution is removed or inserted is governed by several removal and insertion heuristics, respectively. It is desirable to explore new solutions without getting stuck in local minima. The local search framework ensures that not all candidate solutions obtained are accepted.

3.5.1. ADAPTING ALNS FOR ETV TO AIRCRAFT SCHEDULING

The ALNS algorithm cannot be used directly for the ETV-to-aircraft scheduling problem. Some alterations and definitions are necessary, and are discussed in this section.

First, we define a solution s, representing a towing schedule, as a vector of the decision values introduced in Sec (3.2). The objective value associated with s is denoted as $f(s)$, and calculated as in Equation (3.28). Algorithm (1) shows the procedure followed in the ALNS algorithm, adapted from Ropke and Pisinger [135] and Pisinger and Ropke [128]. In this section, the steps in the algorithm are clarified.

Algorithm 1 Adaptive Large Neighborhood Search for ETV-to-aircraft assignment

Require: s^i initial solution (Sec (3.5.1)), N^-, N^+ number of removal and insertion heuristics, $\sigma_1, \sigma_2, \sigma_3$ score values, N^{seg} segment length.

$s^b = s^l = s^i, \quad s^{\text{all}} = \{s^l\}$.
$\bar{\pi}^- = \pi^- = [1/N^-, ..., 1/N^-], \quad \bar{\pi}^+ = \pi^+ = [1/N^+, ..., 1/N^+]$.
for $k = 1, 2, ..., N_{\text{it}}$ **do**
    Select removal and insertion heuristic h_i^+ and h_j^- using π^+ and π^- (Sec (3.5.2)).
    $s^c = h_i^+(h_j^-(s^l))$
    if $f(s^c) > f(s^b)$ **then**
        $s^b = s^l = s^c. \quad s^{\text{all}} = s^{\text{all}} \cup s^l$.
        $\bar{\pi}_i^+ \mathrel{+}= \sigma_1, \quad \bar{\pi}_j^- \mathrel{+}= \sigma_1$
    else if s^c not in s^{all} and $f(s^c) > f(s^l)$ **then**
        $s^l = s^c. \quad s^{\text{all}} = s^{\text{all}} \cup s^l$.
        $\bar{\pi}_i^+ \mathrel{+}= \sigma_2, \quad \bar{\pi}_j^- \mathrel{+}= \sigma_2$
    else if s^c not in s^{all} and s^c is accepted (Sec (3.5.3)) **then**
        $s^l = s^c. \quad s^{\text{all}} = s^{\text{all}} \cup s^l$.
        $\bar{\pi}_i^+ \mathrel{+}= \sigma_3, \quad \bar{\pi}_j^- \mathrel{+}= \sigma_3$
    end if
    if $\text{mod}(k, N^{\text{seg}}) = 0$ **then**
        Calculate π^- and π^+ (Sec 3.5.2)
    end if
end for

FEASIBLE STEPS

The removal and insertion heuristics select certain aircraft from a given solution s to remove or insert. The aircraft are then removed or inserted sequentially, if the resulting solution is feasible. Not all aircraft can be readily removed from the solution or

added to it, without breaking some of the Constraints (3.29)-(3.55), most pertinently Constraints (3.41), (3.44), and (3.55). This means that when applying a heuristic, it should be known for every aircraft whether it may or may not be removed or inserted from the current schedule. We view the situation after removing or inserting any single aircraft as a new *solution step*. These steps are not stored during the ALNS algorithm. While applying the removal and insertion requests from the heuristic, the algorithm first calculates whether it is allowed for an aircraft to be inserted or removed from the current solution step, and if it is not, the aircraft is skipped. If it is allowed, the insertion or removal is executed, and the relevant decision variables are changed. At any point during the algorithm, the set of aircraft that are allowed to be removed is denoted as A^{rem}, and the set of aircraft that are allowed to be inserted into the schedule of vehicle v is denoted as A_v^{ins}.

When an aircraft a is added to the towing schedule of an ETV, the total energy required by that ETV will increase. To make sure that Constraint (3.41) is respected, the vehicle should recharge longer. An existing charging period is selected and lengthened where allowed, with respect to Constraint (3.55). If there is no charging period that can be lengthened, the addition of a is not allowed. Similarly, when removing an aircraft a, a charging period should be shortened. In that case it is important that Constraint (3.44) is still respected for each charging period.

MOVING CHARGING PERIODS

Note that the removal and insertion heuristics can only lengthen and shorten existing charging periods. It is not possible to move a charge period within the time period, or split or merge charging periods. For this reason, the model might not be able to reach all feasible (and possibly better) solutions. To make this possible, the model includes a procedure to change charging periods within schedules.

For every pair of aircraft that are towed consecutively, we name the period of time in between the *task gap*. A task gap can contain a charging period. When (a part of) a charging period is moved from one task gap to another, the former is named the *donator* and the latter the *receiver*. For every existing charging period, we calculate the amount of charge that may be moved to a different task gap. This amount is named the *available charge*. Here we take into account the minimum charging time $t^C_{\min}$ of both the receiver and donator task gaps, as well as the bounds for the state of charge q_a of the vehicle at its arrival at any of the aircraft it will tow. From all existing charging periods with positive available charge, one is selected randomly. The available charge is moved from the donator to the receiver by changing the associated decision variables. If none of the charging periods have positive available charge, no changes are made. In every ALNS iteration k, this procedure is performed $\lceil 0.2N^V \rceil$ times between using the removal and insertion heuristic and then once more after using the insertion heuristic. It was found that increasing the occurrence of the moving charge procedure did not improve the ALNS solution finding process.

Another way of changing the charging periods would be to introduce charging station removal and insertion heuristics, in addition to the current aircraft removal and insertion heuristics. Such an approach was taken by Keskin and Çatay [87]. In this work, the above method was selected, as it allows the moving of charging periods during the removal and insertion of aircraft, where more options for charge movement are possible, rather than only in between applying heuristics.

INITIAL SOLUTION

Before the algorithm starts using the removal and insertion heuristics, it needs an initial solution s^i. The initial solution should be a feasible solution to Problem (3.28)-(3.55), and it should be possible to move to other solutions by adding and removing aircraft and changing charging periods. Note that e.g. the solution with only the artificial aircraft (A^e) and no charging is feasible, but other solutions are unreachable from this solution due to Constraint (3.44).

Given the capacity profile for ETVs, period P and flight schedule, a valid initial solution can readily be constructed:

- For every vehicle, assign several aircraft that need to be towed near the start of the time period P, and together need more than the minimum charging time $t^C_{\min}$ to be replenished.
- For every vehicle, calculate the needed charging time to replenish the battery. Set this charging in such a place that Constraint (3.55) is respected: start by putting charging periods for each vehicle near the end of the time period P, moving backwards in time when required by this constraint.

The initial solution thus constructed can serve as input for the removal and insertion heuristics, which will be introduced next.

3.5.2. LOCAL SEARCH HEURISTICS

The choice of local search heuristics for removal and insertion is an essential part of the ALNS algorithm. Selecting a diverse set of heuristics contributes to the ability of the algorithm to explore the solution space and to escape local minima [128]. The heuristics in the set are being used throughout the run of the algorithm, and are selected with weighted random selection.

HEURISTIC WEIGHTS

The weight updates are performed as described by Pisinger and Ropke [128]. The total number of ALNS iterations is divided into equally sized segments. Two sets of scores exist: the *segment score* $\bar{\pi}^-_{i,j}$ and the *overall score* $\pi^-_{i,j}$ for removal heuristic i and iteration segment j. Similarly, we have scores $\bar{\pi}^+_{i,j}$ and $\pi^+_{i,j}$ for the insertion heuristics. The segment score has an initial value of 0 at the beginning of every segment. During the segment, the segment score is increased with scores that depend on the quality of the candidate solution s^c:

- The candidate solution is a new best solution: increase segment score by σ_1
- The candidate solution has a larger objective function than the latest solution s^l: increase segment score by σ_2
- The candidate solution has a smaller objective function than the latest solution but is accepted by the local search framework: increase segment score by σ_3
- The candidate solution is not accepted by the local search framework: the segment scores remain the same.

Here $\sigma_1 \geq \sigma_2 \geq \sigma_3$. At the end of segment j the overall scores are calculated with:

$$\pi^-_{i,j+1} = \rho \frac{\bar{\pi}^-_{i,j}}{a^-_{i,j}} + (1-\rho)\pi^-_{i,j}, \tag{3.56}$$

with ρ the reaction factor and $a^-_{i,j}$ the amount of times removal heuristic i has been selected in segment j. The overall insertion scores are calculated analogously. Finally, the overall scores π^- and π^+ are normalized, and the result forms the updated heuristic weights.

REMOVAL HEURISTICS

In previous research regarding vehicle routing problems and scheduling problems, many removal and insertion heuristics have been developed. Some heuristics used in this work are taken or adapted from existing literature. In a typical application of ALNS for routing and scheduling problems, tasks can be inserted at other times and in different orders compared to previous solution. By contrast, in our application, the time of a towing tasks always remains the same. In addition, removal and insertion is complicated by restrictions regarding the state of charge and allowed charging moments of vehicles. This means that, in contrast to works such as Frey et al. [54], we opted to disallow moving through infeasible regions of the solution space. Repairing an infeasible solution is expected to be difficult to automate and to be costly in computation time, given the many constraints in this problem.

We use six removal heuristics:

1. Random removal (from Ropke and Pisinger [135]). Select N^{rem} aircraft randomly from the removable aircraft A^{rem} and remove these from the schedule.

2. Vehicle removal. For every ETV, select one aircraft that is currently in its towing schedule and in A^{rem}, and remove that aircraft from the schedule.

3. Cluster removal (from Pisinger and Ropke [128]). Consider the ETV that tows the fewest aircraft during the schedule from all ETV in the fleet. Remove as many aircraft as possible from the schedule of that vehicle. This will leave a few aircraft and one charging period with charging time near the minimum t^C_{min}. The idea of this heuristic is that if the schedule for an ETV is stuck in a suboptimal position, one can take out everything for that ETV and start over.

4. Time-oriented removal (from Pisinger and Ropke [128]). Select randomly a time period of one hour in P. Remove as many aircraft as possible for which the towing period lies within this interval. This heuristic allows for a reorganization of the schedule around a time period, which can help resolve possible charging capacity conflicts in that period.

5. Worst removal (after Ropke and Pisinger [135]). For every aircraft in the schedule and in A^{rem}: calculate the amount of time in the task gap between the aircraft and its successor, and between its predecessor and itself (where applicable). Note that this is the time $t^C(a,b)$ defined in Equation(3.13). Then select the aircraft with the

largest average task gap length for removal. Repeat until N^{rem} aircraft have been removed. This heuristic is based on the notion that aircraft with more task gap length before and after their scheduled towing time can easier be assigned a more efficient place in the schedule.

6. Time-based worst removal. For every aircraft in the schedule and in A^{rem}, calculate the travelling energy q^{travel} saved by removing the aircraft from the schedule. For example, when towing an aircraft b after towing aircraft a and before towing aircraft c:

$$q_b^{\text{travel}} = q^S(a,b) + q^S(b,c) - q^S(a,c). \tag{3.57}$$

Then select the aircraft with the largest saved travelling energy for removal. Repeat until N^{rem} aircraft have been removed. This heuristic aims to optimize the route of an ETV and avoid large detours.

INSERTION HEURISTICS

We use four insertion heuristics. For all insertion heuristics, the procedure described below is repeated until no more aircraft can be inserted. All selected insertion heuristics are of the parallel category; they build on the routes of the entire problem at the same time, not one ETV at once.

1. Random insertion (from Ropke and Pisinger [135]). Consider all aircraft and vehicle combinations in the insertable aircraft for vehicle v (A_v^{ins}). Select one randomly and insert the aircraft into the schedule of that vehicle.

2. Greedy towing distance insertion. Consider the vehicle v^{g} that has the most total task gap length in its schedule. Let S_v^{pair} be the set of aircraft pairs towed consecutively by vehicle v. Then v^{g} is selected as in Eq (3.59).

$$S_v^{\text{pair}} = \{\{a,b\} \in [A \times A] : a \text{ is towed by } v \wedge x_{ab} = 1\} \quad \forall v \in V \tag{3.58}$$

$$v^{\text{g}} = \max_{v \in V} \sum_{\{a,b\} \in S_v^{\text{pair}}} t^C(a,b) \tag{3.59}$$

Consider all aircraft in $A_{v^{\text{g}}}^{\text{ins}}$. Insert the aircraft with the largest towing distance.

3. Greedy tight insertion. Iterate through all vehicles in the fleet. Given vehicle v, consider all aircraft in A_v^{ins}. Insert the aircraft for which the task gap length $t^C(a,b)$ between its predecessor and itself or between itself and its successor is minimized. The goal is to put aircraft as close together as possible in the towing schedule of an ETV, so that more aircraft might be added to the schedule. This heuristic is the opposite of removal heuristic 5.

4. Time-based best insertion. Iterate through all vehicles in the fleet. Given vehicle v, consider all aircraft in A_v^{ins}, and calculate the travelling energy needed to insert the aircraft into the schedule, as for the time-based worst removal heuristic (nr. 6). Then select the aircraft with the smallest needed travelling energy for insertion. Like its counterpart in the removal heuristics, this heuristic aims to optimize the route of an ETV by inserting aircraft that lead to the smallest detours.

Note that for all removal and insertion heuristics, removing/inserting an aircraft will only be performed if it does not result in an infeasible solution (see also Sec 3.5.1). Furthermore, aircraft in A^e are never considered for removal/insertion, as they are always in the schedule.

3.5.3. LOCAL SEARCH FRAMEWORK

The local search framework is the mechanism that decides whether a candidate solution s^c is accepted or rejected. After applying a removal and insertion heuristic the candidate solution is checked using Simulated Annealing and Tabu Search.

Simulated Annealing [91] is a method that aims to guide the local search towards the global optimum. At the start of the search, it allows for more exploration, while near the end of the search, exploration is restricted, and improvement of the objective value is increasingly required for a solution to be accepted.

During the search, Simulated Annealing makes use of the *temperature* parameter T. At the beginning, the temperature has value $T = T^{\text{st}}$. At every iteration of the ALNS algorithm, the temperature is decreased as $T := Tc$, with $0 < c < 1$ the *cooling rate*. To set the value of T^{st}, the method uses the *start temperature control parameter* w. At the start of the algorithm, a candidate solution s^c may be w% worse than the latest solution s^l to be accepted with probability 0.5. As the algorithm progresses, the acceptance probability for a given T and w decreases, following Equation (3.60):

$$p^{\text{accept}} = e^{-(f(s^l)-f(s^c))/f(s^l)/T} \tag{3.60}$$

Setting c and w correctly for a given problem instance requires trial and error. Model parameters that have a large influence on the optimal values for c and w are the length of the time period P, the fleet size N^V and the number of flights N^F.

In addition to the check performed by Simulated Annealing, we make sure that the candidate solution s^c has not been accepted before, i.e. we incorporate Tabu Search [56].

3.6. RESULTS

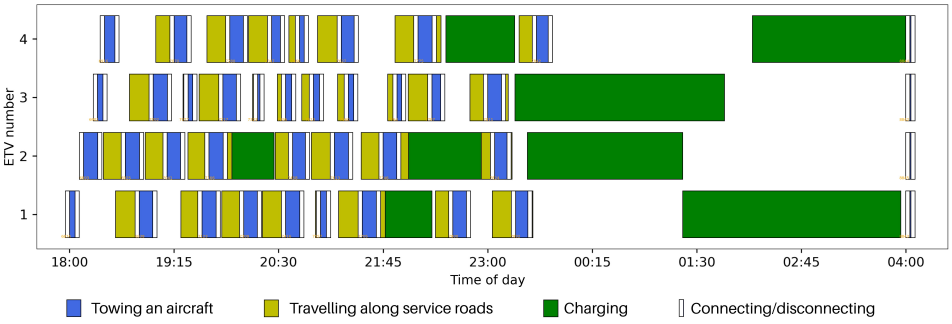
In this section results are presented that are obtained by applying the models introduced in Sec 3.4 and 3.5 to historical flight schedules at Amsterdam Airport Schiphol. The schedules are obtained from the Schiphol Flight API within the Schiphol Developer Center [147]. For each flight, the schedules contain the gate number and scheduled and actual arrival/departure times. The runway for each flight is found from runway use data [101]. A general overview of parameters contained in the models is provided in Table 3.1. The values indicated in the table are the standard values for our approach, and are used unless otherwise indicated.

Table 3.1: Parameters used in the ETV-to-aircraft scheduling models and their standard values.

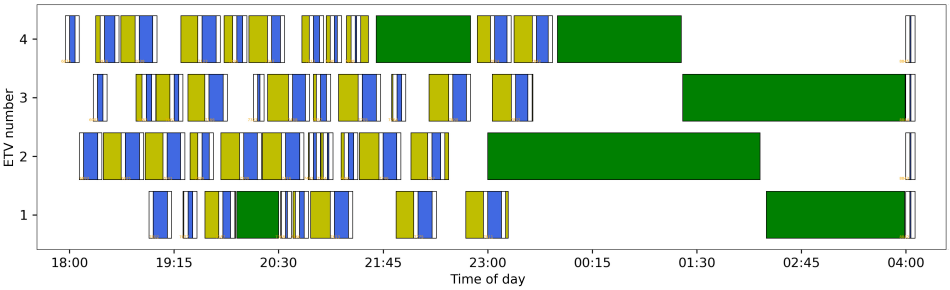
Symbol	Name	Value	Unit	Source
	ETV properties			
P^c	Charging rate	100	kW	[124]
P^X	Towing power	222	kW	[193]
P^S	Driving power	20.5	kW	[193]
Q	Battery capacity	400	kWh	[124]
v^s	Speed on service roads	30	km/h	[143, 116]
v^x	Speed on taxiways	42	km/h	[157]
t^c	Connecting/Disconnecting time	3.0	min	[144]
	Airport infrastructure			
t_{sep}	Separation time	20	s	[193]
t^C_{min}	Minimum charging time	30	min	
Δ	Time step	10	min	
N_{cs}	Number of charging stations	3	-	
	ALNS algorithm			
N^{it}	Number of ALNS iterations	1000	-	
N^{seg}	Nr of iterations in ALNS segment	10	-	
N^{rem}	Nr of aircraft removed in heuristic	N^V	-	
	Simulated Annealing			
σ_1	Global solution reward	2	-	
σ_2	Previous solution reward	0.6	-	
σ_3	Accepted solution reward	0.2	-	
ρ	Reaction factor	0.25	-	
w	Start temperature control parameter	1.05	-	
c	Cooling rate	0.997	-	

3.6.1. MODEL COMPARISON WITH SMALL SCALE PROBLEM INSTANCES

In order to illustrate the performance of the MILP and ALNS models, several problem instances have been defined, summarized in Table 3.2. All instances refer to the flight schedule of 27-12-2021 at Amsterdam Airport Schiphol. The capacity profile for each instance in Table 3.2 is given by: 100% capacity during 04:00-06:00 and 23:00-04:00, 200 kW capacity elsewhere. Parameters not appearing in this table have the values as indicated in Table 3.1. The values of N^{it}, c and w have been manually tuned for each instance to obtain the best performance.



(a) An optimal schedule with 40 aircraft generated by the MILP model.
Total towing distance 200.2 km, CO_2 emissions saved 21.9 ton, and total charging energy 1.16 MWh.



(b) A schedule with 45 aircraft generated using the ALNS algorithm for 4000 steps.
Total towing distance 191.2 km, CO_2 emissions saved 20.9 ton, and total charging energy 1.07 MWh.

Figure 3.2: ETV-to-aircraft schedules generated using the MILP model and ALNS algorithm, for instance 2 in Table 3.2. The ALNS algorithm attains a 4.5% gap compared to the MILP solution for this run.

Table 3.2: Parameter values for problem instances ran with the models: time period, fleet size, number of aircraft, number of iterations, cooling rate, start temperature control parameter and minimum charging time. Times indicated are on 27th (to 28th) of December 2021.

Instance	P	N^V	N^F	N^{it}	c	w	t^C_{min}
1	08:00-14:00	4	49	2500	1.10	0.9993	20
2	18:00-04:00	4	57	4000	1.25	0.9985	30
3	04:00-04:00	2	80	4000	1.25	0.9980	30
4	04:00-04:00	4	138	4500	1.25	0.9985	30
5	04:00-04:00	6	193	3500	1.25	0.9985	30

Figure 3.2 shows a side-by-side comparison of example results for the ALNS and MILP models, for a run of instance 2. These schedules show for each ETV when it is towing an aircraft, travelling along service roads, connecting/disconnecting to/from an aircraft, or charging at a charging station. For this instance, 40 (out of 57 possible) aircraft are towed by 4 ETVs in the optimal solution. In the solution created by ALNS, more aircraft are towed, but the total towing distance, and therefore the emission savings, are lower. By optimizing for total towing distance, the models are incentivized to schedule

towing tasks with longer towing distance, rather than as many short tasks as possible.

The ALNS solution shown is the best solution found after 4000 iterations. The selection probability at a certain iteration for each removal or insertion heuristic is a measure for its effectiveness in improving the solution during the previous iterations. Figure 3.3 shows the selection probabilities of all removal and insertion heuristics, listed in Section 3.5.2, for twenty runs of problem instance 2. Observing the average selection probabilities provides insight in the usefulness of each heuristic. For example, random removal and insertion (nrs. 1) perform well, despite their simple nature. Greedy insertion (nr. 2) also performs well, especially for schedules with larger fleet size. This likely happens because it pushes to insert the tasks with the larger towing distance. On the other hand, time-based removal (nr. 6) and time-based insertion (nr. 4) perform worst. On many occasions, the selection probability of time-based removal converges towards zero. The average towing distance of towing tasks added by time-based insertion is the smallest of all insertion heuristics. Similarly, the average towing distance of towing tasks removed by time-based removal is the largest of all removal heuristics. Rather than achieving their goal of minimizing detours in the ETV route, the heuristics seem to avoid longer towing tasks. This is because a longer towing task is more likely to constitute a large detour for an ETV than a shorter towing task.

We now examine the performance of the MILP and ALNS models for all instances introduced in Table 3.2. Note that the number of flights N^F indicated in Table 3.2 is not the total number of flights passing through the airport during the time period P. The number of flights eligible for towing has been reduced to accommodate the MILP model: with a smaller N^F value, the solution time is greatly reduced, since fewer decision variables and constraints are generated (see Sec 3.4). In addition, if the amount of scheduled aircraft in an optimal solution is close to the number of total available flights, the scheduling problem becomes easier than with a larger total of available flights, due to combinatorics. These problem instances allow us to compare both models equally. Nevertheless, the MILP model does not find the optimal solution for instance 3-5. For comparison with the ALNS model, the best incumbent solution and best bound are used.

Table 3.3 shows the objective values (total towing distance of the schedule) obtained with the MILP and ALNS models. For the MILP model, the optimal solution, best incumbent solution and best bound are indicated where applicable. The runtime for the instances without an optimal solution indicates the time the optimization process was continued for, not the time at which the best incumbent solution was found. For the ALNS model, 20 runs were performed for each problem instance. The best solution, the mean best solution (average of the best solutions of all runs) and the gap of the best ALNS solution with respect to the MILP solution (optimal or best incumbent) are given.

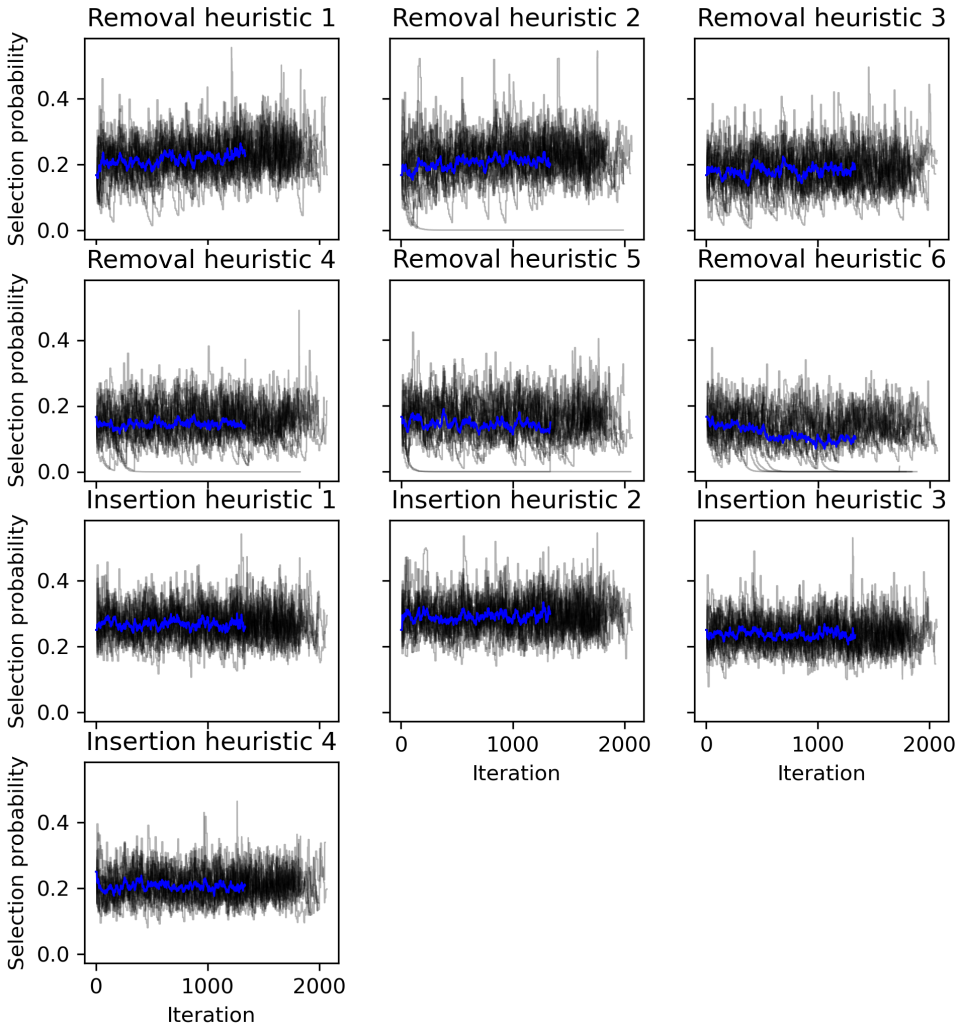


Figure 3.3: The evolution of selection probabilities for all heuristics, for 20 runs with 2500 steps of the scheduling problem for 18:00-04:00 on 27-12-2021. The average is indicated by the blue line.

Table 3.3: Total towed distance values and runtimes for solutions obtained by the MILP and ALNS models, when applied to the problem instances introduced in Table 3.2.

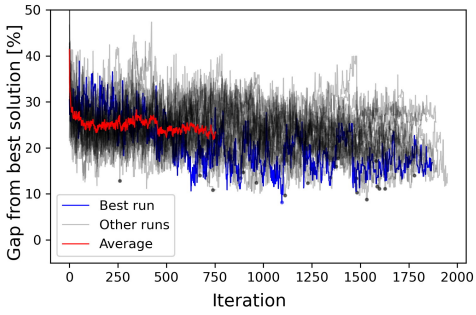
	MILP			
Instance Unit	Opt. sol. [km]	Best sol. [km]	Best bound [km]	Runtime [min]
1	141.3	-	-	1.2
2	200.2	-	-	0.1
3	-	265.0	275.2	190
4	-	508.2	527.0	188
5	-	712.7	719.1	241
	ALNS			
Instance Unit	Best sol. [km]	Mean best sol. [km]	Gap w.r.t MILP [%]	Runtime [min]
1	129.8	123.0	8.1	1.5
2	198.8	188.3	0.7	2.6
3	247.8	243.7	6.5	7.5
4	496.8	479.2	2.2	17
5	688.0	674.8	3.5	25

In addition, Figure 3.4 shows graphs of convergence for the 20 runs with the ALNS model for all five instances. The average and the best run are highlighted, as well as the best result for every run. The convergence is illustrated as a gap value relative to the MILP solution (either the optimal or best incumbent). For several of the instances in the figure it seems using more iterations than shown would improve the solution further. However, this is not the case, since in fact many rejected solutions are produced between and after the shown accepted solutions, which are not included in the figure.

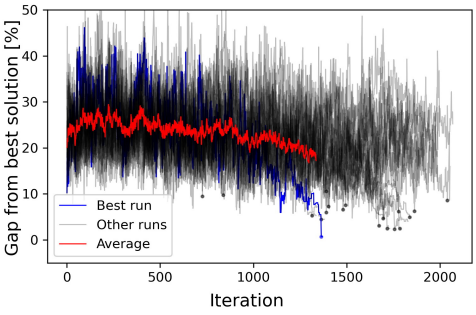
Table 3.3 and Figure 3.4 show that when the period of interest P and number of flights N^F are made small enough, the MILP model finds the optimal solution fast, and is superior to the ALNS model. However, when the period of interest P comprises a full day, and the ETV fleet size N^V starts to (slightly) increase, the MILP model is unable to find the optimal solution within several hours. The best solution found through the ALNS model is close to the best incumbent solution of the MILP model. These results show that for a time frame of more than a few hours, the ALNS provides results of quality close to that of the MILP model, in a much shorter time. Note that creating a schedule with the ALNS model for any of the instances 1-5 with all aircraft that pass through the airport during their respective time period P (rather than only the selection now used) will not increase the runtime.

3.6.2. THE IMPACT OF THE ELECTRICITY CAPACITY AT THE AIRPORT ON THE ETV TOWING SCHEDULES

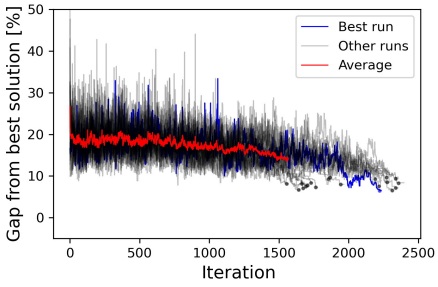
The variation of electricity capacity throughout a day of operation directly influences the ETVs' ability to recharge. In a successful ETV schedule the ETVs charge in such a way that they can perform towing tasks in as much of the time period P as possible, without running out of charge during a time when little or no charging is allowed. In other words,



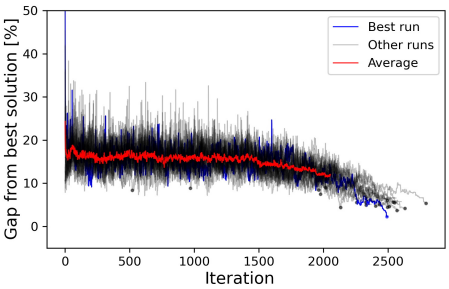
(a) Convergence for 08:00-14:00 on 27-12-2021, for 4 ETVs



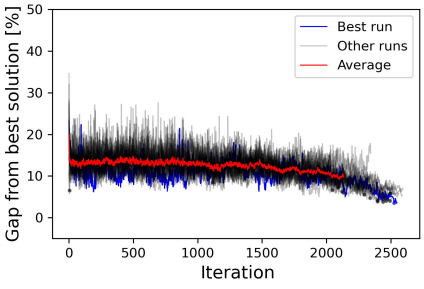
(b) Convergence for 18:00-04:00 on 27-12-2021, for 4 ETVs.



(c) Convergence for 04:00-04:00 on 27-12-2021, for 2 ETVs.



(d) Convergence for 04:00-04:00 on 27-12-2021, for 4 ETVs.



(e) Convergence for 04:00-04:00 on 27-12-2021, for 6 ETVs.

Figure 3.4: Convergence of ALNS algorithm applied to various problem instances. The run attaining the largest objective value is indicated in blue, the average is indicated in red. The best result of every run is indicated by a dot.

the ETV utilization is high. We aim to investigate the influence of the electricity capacity profile for ETVs on the schedules generated by the model.

The ETV capacity profiles are given by:

Capacity profile A: night capacity. 0% capacity during 04:00-23:00 and 100% capacity during 23:00-04:00.

Capacity profile B: overall low capacity. 40% capacity during 04:00-04:00.

Capacity profile C: no capacity during rush hour. 0% capacity during 07:00-10:00 and 16:00-19:00, 40% capacity during 10:00-16:00 and 19:00-23:00 and 100% capacity during 04:00-07:00 and 23:00-04:00.

Capacity profile D: low capacity during day. 0% capacity during 06:00-23:00 and 100% capacity during 04:00-06:00 and 23:00-04:00.

Capacity profile E: full capacity. 100% capacity during 04:00-04:00.

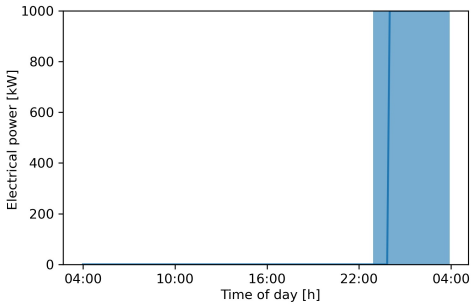
The profiles represent possible charging capacity situations at the airport. Note that larger charging capacity for ETV charging implies smaller charging capacity for the other processes at the airport. For example, in profile C, no capacity during rush hour implies that all capacity has been taken by other processes.

Using the ALNS model, ETV schedules have been created for each of these capacity profiles, and for several fleet sizes. In all cases, the time period P is given by a full day (between 04:00 and 04:00), in which 955 aircraft movements are planned. Figure 3.5 provides a visualization of the five capacity profiles considered, for a fleet size of 10 ETVs. In addition, the electricity demand from this fleet for a scheduling solution obtained with ALNS is shown. Note that $P^c = 100$ kW, so that every 100 kW on the y-axis translates to one charging ETV at the time given on the x-axis.

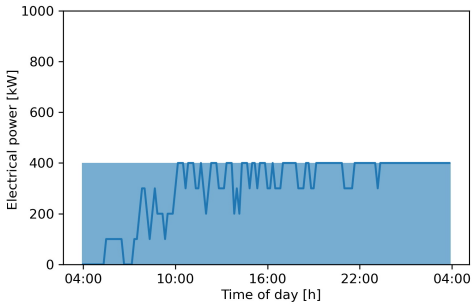
In addition, Figure 3.6a shows an overview of objective values obtained when creating ETV schedules for fleet sizes that are representative of an airport aiming to fully implement electric taxiing. For each combination of the capacity profiles A-E and fleet size value in $\{5, 10, 20, 30\}$, a boxplot is shown, which summarizes the objective values of five runs. Average runtimes for 1000 iterations of the ALNS algorithm are given by 45.1, 64.3, 125 and 199 min for fleet sizes 5, 10, 20 and 30 ETVs, respectively.

By examining Figure 3.5 and Figure 3.6a, as well as the schedules generated by all runs that have been summarised in Figure 3.6a, we draw conclusions regarding the ETV electricity capacity profiles.

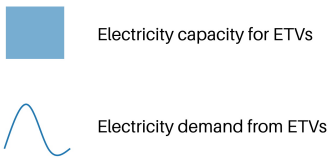
- We see that for every fleet size considered, only allowing night charging leads to an average decrease in total towed distance of 46% compared to the best performing profile, profile E. This shows that an ETV would need almost twice the current battery capacity to keep towing aircraft for the full day.
- For the second charging profile, we see that the limited capacity of 40% is used to the fullest from roughly 10:00 until the end at 04:00. For several ETVs the utilization time is reduced, because the ETV runs out of charge and cannot charge earlier due to the decreased capacity. The average decrease in objective value is 13%.
- The third charging profile simulates a situation where little or no electricity capacity can be reserved for charging ETVs during busy periods. The aircraft in the schedule between 16:00-19:00 are atypical: for small fleet size (5 or 10), aircraft

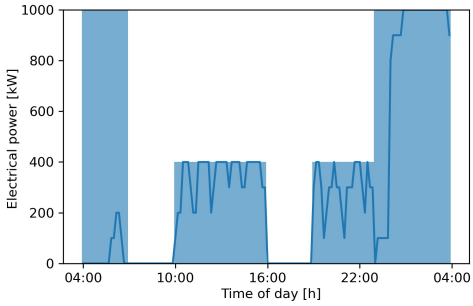


(a) Energy demand for capacity profile A: night capacity.

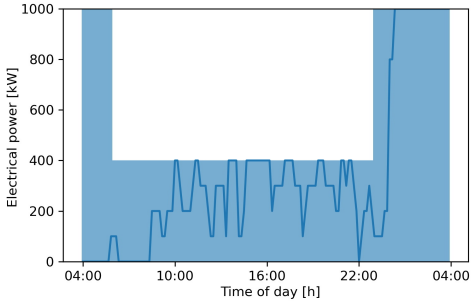


(b) Energy demand for capacity profile B: overall low capacity.

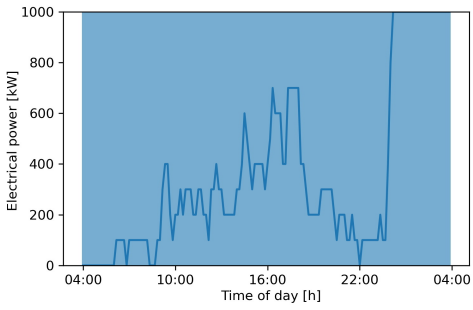




(c) Energy demand for capacity profile C: no capacity during rush hour.



(d) Energy demand for capacity profile D: low capacity during day.



(e) Energy demand for capacity profile E: full capacity.

Figure 3.5: The electrical power available during the day, for the five energy profiles considered. For every profile, the energy demand per time step associated with an ETV schedule generated using the ALNS model is plotted. The fleet size is ten ETVs, and the time period is 04:00-04:00 on 27-12-2021.

with very short towing distance are scheduled during this time. For the larger fleet sizes, the ETV utilization is low during this period, i.e. the schedule is more empty. This is because there are not enough tasks with small towing distance for all ETVs.

- Charging profile D is similar to profile B. The main difference is that night charging (23:00-04:00) allows the ETVs to fully recharge for the next day. In the schedule for profile B most ETVs have one very long charging period during the day, rather than at night. Comparing profiles B and D in Figure 3.6a confirms that allowing night charging leads to a substantial difference in objective value (on average 10% of the maximum objective value per fleet size).

- Last, charging profile E allows charging for every vehicle at any time. The energy demand that appears under these circumstances is as in Figure 3.5e: demand from 06:00-23:00 is roughly triangular shaped, with a peak between 15:00-17:00. Upon examination of the optimized schedules, we find two types of ETVs in the schedule: one with a long charging period near 15:00-17:00, that needs no other charging before the night. The other type is charged for 3 or 4 shorter periods, spread out during the day. Together this forms a triangular shaped demand.

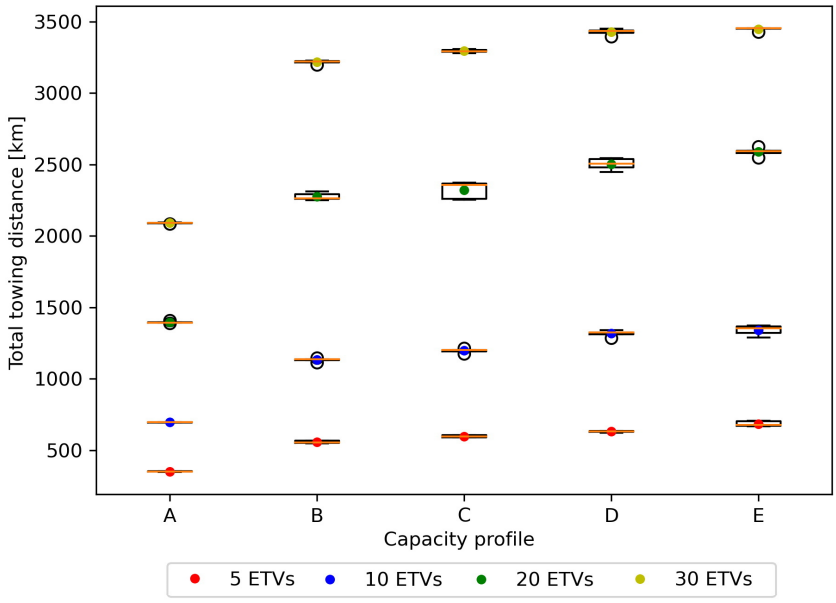
In addition, it is interesting to investigate the *marginal capacity benefit*: the benefit for an airport of providing one more charging position (i.e. increase hourly capacity by P^c kW) for the total towing time in the optimized ETV schedule. This benefit can be weighed against the costs of a new charging position.

Figure 3.6b shows the results of applying the ALNS model with various equidistant values of electricity capacity. The electricity capacity is assumed to be constant over the entire day. Every newly added charging position contributes less to the objective value than the previous one. The figure suggests that after 12 charging positions (in this case 60% of the fleet size), the improvement stagnates. The limit for the ETV fleet with the current characteristics is roughly 2500 towed km (273 ton CO_2). The same analysis can be performed for different capacity profiles and fleet sizes.

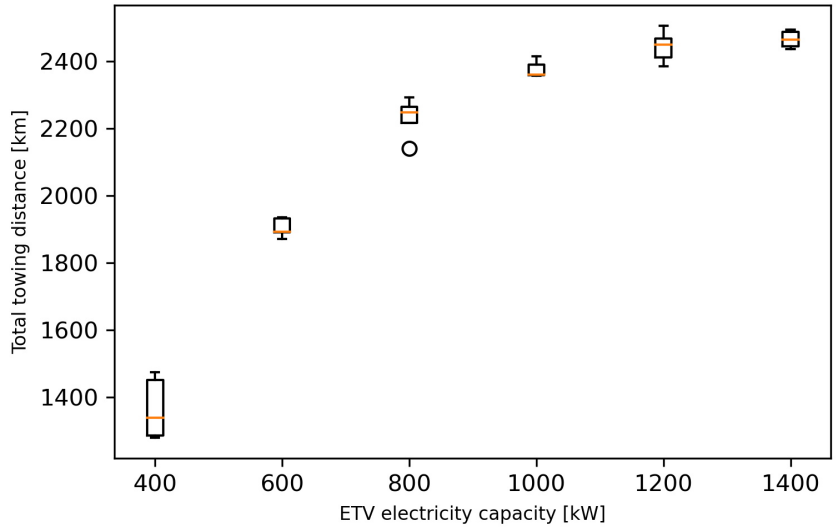
3.6.3. IMPACT OF FAST CHARGING AND BATTERY SIZE ON ETV FLEET UTILIZATION

Throughout the previous sections, the charging rate P^c and battery capacity Q have remained constant. The values chosen are considered realistic for application at the time of writing. However, given their large influence on the ETV schedules resulting from the models, it is instructive to consider (combinations of) other values of these parameters.

Table 3.4 shows the total towed distance obtained when varying the charging rate and battery capacity. The average and standard deviation of 5 runs are shown for every combination. The fleet size is 20 and the electricity capacity is given by profile D. The average runtime is 134 min. Note that we assume constant towing and driving power P^X and P^S. In actuality, increasing the battery capacity will increase these values, due to the ETV becoming heavier.



(a) Total towing distance for aircraft in schedules generated using the ALNS algorithm, for varying fleet size and different ETV charging capacity profiles. Multiple runs for the same combination are grouped in boxplots. Coloured dots indicate the average result.



(b) Total towing distance for aircraft in schedules generated using the ALNS algorithm, for various constant ETV charging capacities. The fleet size is 20 ETVs.

Figure 3.6: Total towing distance in km (Equation(3.28)) of schedules generated with the ALNS algorithm, for various model settings, during 04:00-04:00 on 27-12-2021.

Table 3.4: Impact of fast charging and ETV battery capacity; charging rate $P^c \in \{50, 100, 150, 200\}$ kW and battery capacity $Q \in \{200, 400, 600\}$ kWh. The total towing distance in km (Equation(3.28)) obtained in schedules created with the ALNS algorithm.

		P^c			
		50 kW	100 kW	150 kW	200 kW
Q	200 kWh	1846.8 ± 12.5	2183.0 ± 36.9	2283.6 ± 39.1	2407.4 ± 26.5
	400 kWh	1942.1 ± 20.5	2508.6 ± 18.8	2608.8 ± 21.3	2713.3 ± 8.7
	600 kWh	1940.9 ± 29.6	2634.7 ± 30.8	2731.0 ± 77.6	2797.8 ± 16.6

From the table we deduce that if the charging rate is as small as 50 kW, increasing the battery size from 200 kWh will provide little benefits. At this charging rate, the utilization time of the ETVs is up to 25% smaller than for charging rates of 100 kW and over. There, the increase of objective value with increasing battery size is also considerably larger. The generated schedules show that for $P^c = 50$ kW the ETVs spend up to half the day charging, and the ETVs rarely use more than 300 kWh of their battery.

We observe that at any battery capacity, increasing the charging rate leads to significant increases in objective value. This is mainly because the necessary charging periods during daytime can become shorter, leaving more time to tow more aircraft. The time taken up by charging during the day is roughly inversely proportional to P^c. Therefore, the increases in objective value become smaller with each increase of P^c. Note that the electricity capacity does not grow with the charging rate; if the capacity is 800 kW, then 8 ETVs can charge when $P^c = 100$ kW, and 4 ETVs can charge when $P^c = 200$ kW.

For values in the top right of Table 3.4, an ETV is recharged the fastest; as fast as 1 h. This makes the window for allowed charging periods rather small, since there is also a minimum charging time of 30 min. Given an intermediate solution, the ALNS algorithm may not find a legal move for charging periods when adding or removing aircraft. This results in solutions where many ETVs tow far fewer aircraft than optimal. For a minimum charging time $t^C_{\min}$ of 30 min the results for $Q = 200$ kWh and $P^c = 200$ kW are 1771.6 ± 136.5 km. A solution is to allow a smaller minimum charging time for this combination, which is fitting for a situation with fast charging. The results in Table 3.4 for this combination are obtained using $t^C_{\min} = 20$ min.

From the optimized schedules summarized in the table, and the observations above, we can deduct a relation between the objective value and P^c and Q. We observe:

- In almost all cases, the time available for night charging (00:00-04:00) is fully used, unless a full charge takes less than 4h. Between 04:00 and 06:00 there are very few flights and no charging.
- Any charging during the rest of the day (06:00-00:00) prevents ETVs from towing aircraft. The sum of the time spent charging during the day and the time spent towing aircraft should equal 18 hours.
- It is assumed that at any point in time, the supply of aircraft that can be towed outstrips the towing potential of the ETV fleet size.

An approximation for the total towed distance, in the case where day charging is neces-

sary, is then derived as follows:

$$\begin{aligned} t^d &= t_C^d + t_X^d \\ &= \frac{q_{\text{km}} d_{\text{tot}}^X - q^{\text{night}}}{P^c} + t_{\text{km}}^{\text{ETV}} d_{\text{tot}}^X, \end{aligned} \tag{3.61}$$

so that

$$d_{\text{tot}}^X = \frac{\frac{q^{\text{night}}}{P^c} + t^d}{t_{\text{km}}^{\text{ETV}} + \frac{q_{\text{km}}}{P^c}}. \tag{3.62}$$

Here t^d is the total daytime available to the fleet in hours, i.e. $t^d = 18N^V$, t_C^d and t_X^d are the total daytime spent charging and towing, respectively, q_{km} is the average energy needed per km towing, d_{tot}^X is the total towed distance in the schedule, q^{night} is the total energy that can be charged during the night and $t_{\text{km}}^{\text{ETV}}$ is the average time spent by an ETV towing an aircraft for one km.

These expressions are found as:

$$q^{\text{night}} = N^V P^c \min\{\frac{Q}{P^c}, 4\}, \tag{3.63}$$

$$q_{\text{km}} = \frac{\sum_{a \in A} q^X(a)}{\sum_{a \in A} d_a^X} \tag{3.64}$$

Last, $t_{\text{km}}^{\text{ETV}}$ is found by considering the runs forming Table 3.4, dividing the total time not used for charging by the total towed distance.

In case no day charging is necessary, the expression (3.62) reduces to:

$$d_{\text{tot}}^X = \frac{t^d}{t_{\text{km}}^{\text{ETV}}}. \tag{3.65}$$

By using these formulae all values obtained experimentally in Table 3.4 and the values for profile D in Figure 3.6a are approximated to within 6%. The approximation slightly overestimates the objective values for smaller values of Q, and slightly underestimates them for larger values of Q. A possible explanation is that there is a benefit that comes with a larger Q, that is not factored in with this approximation: an ETV with larger battery capacity will less often reach a SOC of 0, which would leave a time gap without towing until its next charging opportunity.

The approximation given in Equations (3.61)-(3.65) can be further extended to allow for variation in ETV capacity profile and towing power P^X. It can also be altered to adhere to a different charging strategy, for example, if one would relax the restriction of ending the day with full battery capacity. The model can be used by airport planners to gauge the emission reductions associated with acquiring an ETV fleet with any values of Q, P^c or N^V. For example, one can weigh the added costs of using very fast charging technology against the expected environmental benefits. With the current settings, a fast charging system that provides 500 kW per charging point can provide another 7% increase in towed distance for $Q = 200$ kWh. For larger Q this advantage shrinks. For $Q = 750$ kWh night charging suffices from $P^c = 150$ kW onwards, and very fast charging becomes unnecessary. Eq (3.65) suggests a maximum objective value of 2849 km towed distance (311 ton CO_2 saved) for a fleet size of $N^V = 20$ with only night charging.

3.7. Conclusion

In this chapter, two models are proposed to create schedules for electric towing and charging of a fleet of Electric Taxiing Vehicles (ETVs) on an airport. Both models aim to maximize the total towed distance, given a flight schedule, an airport layout, an ETV fleet size, an ETV energy spending and charging module, and an airport electricity capacity profile for charging ETVs. The resulting schedules define which ETV is charging where and for how long, and which ETV is towing an aircraft, or travelling to and from a task.

The first model is a Mixed Integer Linear Programming (MILP) model. An optimal solution or near best bound is found within several hours for instances with small fleet size and a scheduling time period of under ten hours. The goal is to find solutions for a 24 h period and fleet sizes of up to 30 vehicles. Therefore, the second model uses Adaptive Local Neighbourhood Search (ALNS), combined with Simulated Annealing and Tabu Search. We have seen that this model can find solutions with an optimality gap of a few percent for the smaller instances, and can solve the large instances required. Therefore, we conclude that it is possible to pursue optimization of ETV-to-aircraft assignment for a large fleet within a few hours, making the problem tractable. The heuristics introduced to move between different solutions vary in performance: the best heuristics are random removal/insertion, vehicle removal, cluster removal, and greedy insertion.

In order to investigate the effects of limited electricity capacity at the airport, the ALNS model was applied to a full day of operations at Amsterdam Airport Schiphol, with five different ETV electricity capacity profiles. For the battery and ETV properties assumed in this work, it was found that charging at night (when there are no aircraft to tow) is necessary to fulfill the potential of the ETV fleet, but not sufficient to tow all aircraft. Having capacity during the day is therefore crucial to improve ETV utilization time. The results show that even with only small intermittent periods of no charging, the best solution will contain time periods with fewer or shorter tasks than can potentially be scheduled. Last, when charging is allowed for any ETV at any time of day, the ETV daily electricity demand generally forms a triangle shape, placing the largest demand on the electricity network at 15:00-17:00.

The effects of airport investments in ETV battery capacity and fast charging technology on the environmental benefits of the ETV fleet have been explored. A relation between the total towed distance and these parameters was derived by observations of the results from the introduced models. It was found that charging slower than 100 kW will reduce the total towed distance to well below the potential of the ETV fleet. Faster charging will improve the towed distance by freeing time during the day to tow more aircraft, but with diminishing returns. For battery capacities higher than 750 kWh, all necessary charging can be done at night, and very fast charging becomes unnecessary.

The models introduced in this chapter can be applied to other airports and flight schedules, with their own potential ETV electricity capacity profiles. The results can be used to decide whether to implement electric taxiing at the airport, and if so, with how many charging points and ETVs. Or, when such a system is already in place, to consider the costs and benefits of increasing the fleet size, charging capacity or amount of charging locations.

Future work includes optimizing real-time operations of an airport with a fleet of

ETVs, including disruptions such as flight delays and cancellations and ETV unavailability. Such a continuous solution process would aim to maximize emission reduction and robustness. It would constitute the next step in integrating ETV fleet scheduling optimization into actual airport operations. Another upcoming research area is the development of autonomous airport surface movement, including for ETVs.

4

Fleet Scheduling for Electric Towing of Aircraft with Disruption Management

In Chapter 3 models were developed to create assignments of Electric Taxiing Vehicles (ETVs) to aircraft in advance of operation. When disruptions such as flight delays occur during operation, an adapted assignment is required for efficient ETV movement. In this chapter, a strategic assignment model such as those in Chapter 3 is defined. Furthermore, a disrupted assignment model is introduced, with which an adaptive vehicle-to-aircraft assignment is created. The disrupted model maximizes the number of towed aircraft and minimizes the schedule changes for vehicle operators. A case study shows the efficacy of the disrupted model in minimizing schedule changes, which does not come at the expense of emission savings.

This chapter is based on the following research article:
Zoutendijk, M., Van Oosterom, S. J. & Mitici, M. (2023). "Electric Taxiing with Disruption Management: Assignment of Electric Towing Vehicles to Aircraft" in *AIAA AVIATION Forum 2023* [193].

4.1. Introduction

The aviation industry has the ambition of achieving net-zero greenhouse gas emissions before 2050 [49, 75]. An important part of total emissions is the ground-based emissions, created by activities on and around airports. Aircraft taxiing produces 54% of emissions related to the landing/take-off cycle [25], and is therefore a large contributor to ground-based emissions. A promising method to reduce taxiing emissions is to introduce a fleet of electric towing vehicles (ETVs) to airports, which tow the aircraft from gate to runway and vice versa. For example, research has shown that the use of ETVs will reduce the taxiing fuel use by roughly 80% [36, 145].

Implementing a fleet of ETVs on an airport brings about various challenges. In previous work, researchers have investigated the cost-effectiveness of ETVs [70, 170], their effect on on-time performance [134, 88], and they have modelled the expected decrease in fuel use, emissions and noise [63, 88, 65, 103].

In addition, there are operational challenges, such as the assignment of vehicles to aircraft, the charging infrastructure and planning, the shared usage of airport roads, and the robustness of schedules under disruptions.

Vehicle-to-aircraft assignments Several models have been developed to generate vehicle-to-aircraft assignments for ETVs. For example, Baaren and Roling [16] and Soltani et al. [160] created a Linear Programming (LP) model to select aircraft to be towed so that fuel reduction was maximized. The former performed a sensitivity analysis on ETV fleet size, while the latter included collision avoidance in their model. Ahmadi [2] created a complex LP model which minimized not only fuel reduction, but also ground delay costs and operating costs. Oosterom and Mitici [124] created a Mixed Integer Linear Programming (MILP) model minimizing the ETV fleet size. In a different approach, Zaninotto et al. [182] and Salihu, Lloyd, and Akgunduz [137] performed the vehicle-to-aircraft assignment at the same time as the vehicle and aircraft routing, by simulating all ground movement and assigning the ETVs to aircraft that can reach them the first.

Energy module In order to obtain a realistic overview of airport surface movement in the situation of electric taxiing, it is essential to include the energy usage and vehicle recharging in the model. There are few studies which include airport surface movement planning, electric taxiing, and an energy module. One example is Baaren and Roling [16], who prescribe that every ETV is recharged to full capacity upon returning to the depot. Second, Oosterom and Mitici [124] include three different charging strategies: night-charging, fixed-time charging and partial charging. They conclude that partial charging allows for the smallest ETV fleet size.

Disruption management After obtaining a model that takes into account all inputs and requirements and using it to generate a valid schedule, the schedule must be executed. During actual operations, disruptions to the schedule can occur that require airport planners to alter the schedule. We refer to this as *disruption management*. Questions pertaining to disruption management are:

- To what extent can the original schedule be used to obtain the disrupted schedule? What kind of changes to the schedules can be deployed?
- What metric is observed when changing the original schedule? What is the relation of this metric with the metrics observed when creating the original schedule?

An example of an application of disruption management in the aerospace domain is Lee, Marla, and Jacquillat [96]. The authors formulate a Stochastic Mixed Integer Program to obtain a disrupted flight planning schedule for staggered time periods. The tools available to change the schedule include flying altitude and speed, and aircraft swaps and cancellations. Both real-time revealed disruptions and probabilistic forecasts of future disruptions are inputs to the model. The proposed model reduces the disruption recovery costs by 1 to 4% compared to a baseline flight planning model. In later work, the authors show that using probabilistic forecasts reduces expected recovery costs by 1 to 2% more than using only real-time disruptions [97].

To the best of our knowledge, there is no disruption management approach for assignment of electric towing vehicles to aircraft, or for general airport surface movement planning. In this chapter we propose two MILP models specifically designed to facilitate disruption management for ETV-to-aircraft assignment at an operational level. The first is a *strategic model*, which is used to obtain an initial vehicle-to-aircraft assignment for a four hour time window. The second is a *disrupted model*, which is used to obtain a disrupted assignment, taking into account the effects of imminent flight delays, and minimizing the schedule changes from the perspective of ETV operators. This model is run at the start of staggered time periods. Airport layout information and flight schedules are used to obtain the routes for the aircraft and vehicles, after which they are used to generate vehicle-to-aircraft assignments. Realistic parameters and conditions are taken into account, such as the prevention of conflicts in routing. Furthermore, the energy usage and vehicle charging are integrated in the model.

The main contributions of this chapter are:

1. We propose a strategic and a disrupted MILP model for vehicle-to-aircraft assignment which facilitate disruption management throughout a day of operations.
2. We include an energy usage and charging module, conflict and collision avoidance, and realistic parameters.
3. We apply our models to a case study at a large hub airport and investigate the effect of varying model parameters and inputs.

The remainder of this chapter is organized as follows: in Section 4.2 the framework of and inputs to the strategic model for vehicle-to-aircraft assignment are introduced, including the aircraft and vehicle path planning and the ETV energy module. Then, the MILP formulation is stated and explained. In Section 4.3 the disrupted MILP model is formulated and its usage is illustrated. In Section 4.4 the models introduced in Section 4.2 and 4.3 are applied to Amsterdam Airport Schiphol to obtain series of vehicle-to-aircraft assignments. In Section 4.5 sensitivity analysis is performed on the models by varying parameters and inputs. Finally, Section 4.6 summarizes the findings and recommendations for future work.

4.2. Strategic Management of Electric Taxiing Vehicles Towing Aircraft

In this section we propose a model that generates an assignment of departing and arriving aircraft to a fleet of ETVs. These ETVs electrically tow the aircraft assigned to them, and are charged at charging stations when required. The model is an adaptation of the MILP model introduced in Chapter 3. Note that a glossary with terms and notation used in this and other chapters can be found at the end of the dissertation.

4.2.1. Airport layout

An airport is represented by two graphs, with some shared nodes and edges: the taxiway graph $G_X = (N_X, E_X)$ and the service road graph $G_S = (N_S, E_S)$. All edges in E_S are two-directional, and some edges in E_X are one-directional.

In this chapter Amsterdam Airport Schiphol (EHAM) will be used as the reference airport. Figure 4.1a shows a technical map of this airport, and Figure 4.1b shows the graph representation used for the models presented in this chapter. It highlights the taxiways and service roads used to reach the five considered runways. The many gates of the airport are grouped into several gate nodes. The runways are assumed to use only one exit/entry point, and at this point the runway node is located. Furthermore, nodes n_5, n_{108} and n_{110} are designated as charging stations. Formally, the set of charging stations $S_{\text{cs}} = \{n_{\text{cs},i} : i \in \{1, ..., N_{\text{cs}}\}\} = \{n_5, n_{108}, n_{110}\}$ and its total number is $N_{\text{cs}} = 3$. Last, node n_5 also functions as the ETV depot n_{dp}.

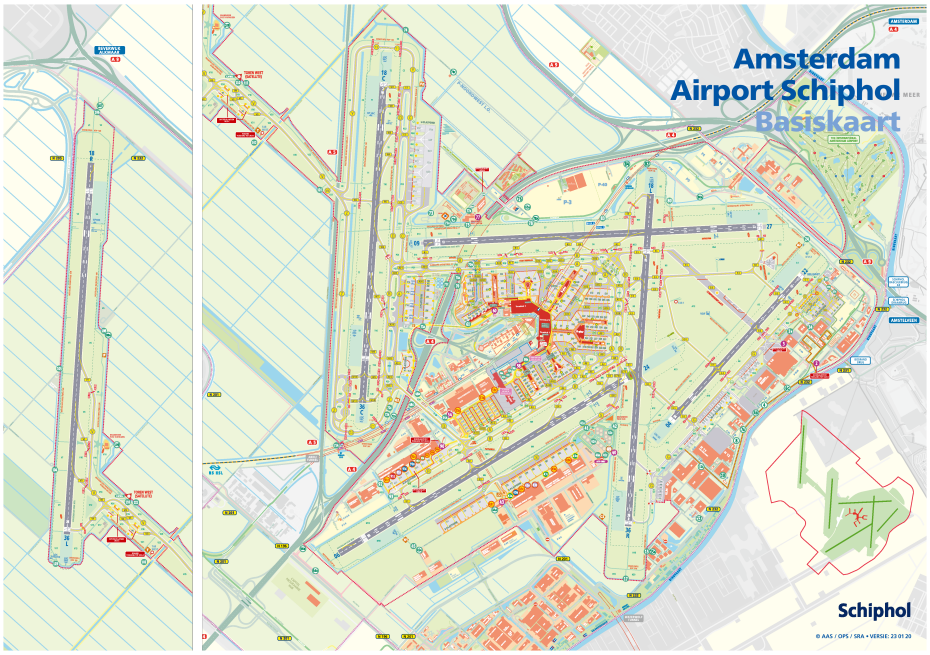
Taxiing aircraft and ETVs towing aircraft are allowed to drive on the taxiways, where a maximum speed of v_x holds. ETVs that are not towing aircraft are allowed to drive on the service roads, where a maximum speed of v_s holds. Although not all airports have a network of service roads available, it is assumed in this work that airports that implement a fleet of ETVs will also implement a service road network. This is because it would be very difficult to route both the aircraft and the travelling ETVs along the taxiways, while retaining the same throughput of flights [146, 180].

We define the distance metrics $d_X(m, n)$ and $d_S(m, n)$ for G_X and G_S, respectively. These distances are the shortest distances within the graphs, which are calculated using Dijkstra's algorithm.

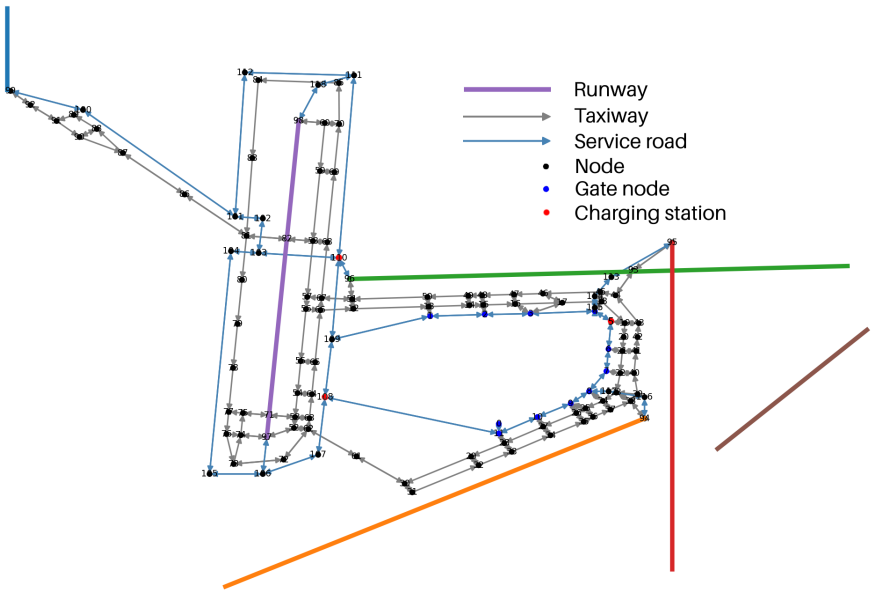
4.2.2. Aircraft path planning

Before we perform vehicle-to-aircraft assignment, we create the routes the aircraft have to traverse. The assignment is made for a certain time period $P = [t^s, t^e]$ on a certain day. The aircraft a that arrive or depart at the reference airport within P form the set A. The number of aircraft within A is denoted as N^F.

Given the flight schedule for that day, we collect the scheduled times of arrival or departure of all aircraft within P. More specifically, the scheduled landing time (SLDT) of an arriving flight or the scheduled off-block time (SOBT) of a departing flight is used, since that is the moment a towing vehicle can start to interact with the aircraft. We will refer to this time as the pick-up time of aircraft a: t_a^p. The time an ETV has finished towing the aircraft is referred to as the drop-off time t_a^d. Furthermore, the pick-up and drop-off locations n_a^p and n_a^d are obtained from the flight schedule and airport layout.



(a) Technical map, displaying all roads, runways, gates and buildings [11].



(b) Graph representation of the airport, $G_X \cup G_S$, including the runways, gates, taxiways, service roads and charging stations.

Figure 4.1: Maps of Amsterdam Airport Schiphol.

For a departing aircraft, the pick-up node is a gate node and the drop-off node is a runway node, and for an arriving aircraft, vice versa.

Using the distance metric d_X defined in Section 4.2.1, we calculate the time needed to traverse the distance between any pair of nodes on G_X as:

$$t_X(m,n) = d_X(m,n)/v_x \quad \forall m,n \in G_X. \tag{4.1}$$

This means that the drop-off time of aircraft a is calculated as:

$$t_a^d = t_a^p + t_X(n_a^p, n_a^d) \quad \forall a \in A. \tag{4.2}$$

It is assumed that no delay is incurred during taxiing due to reasons such as malfunctioning vehicles or aircraft or weather events. However, it is possible that aircraft incur delay because they have to make use of the same nodes or edges at the same time as another aircraft. Furthermore, the minimum separation distance d_{sep} between two aircraft should be respected and is set at 200 m [182, 133]. Since the taxiway speed v_x is roughly 10 m/s, the separation time t_{sep} is set at 20 s.

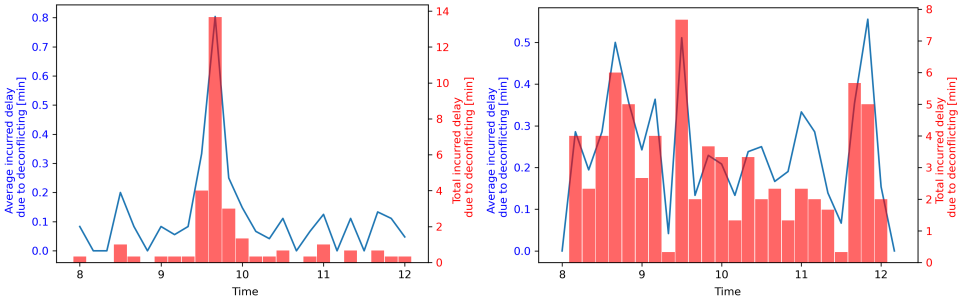
It is necessary to select which aircraft has to wait at which moment in order to resolve potential conflicts or separation violations. This we refer to as *deconflicting routes*. Using the routes, including the time at which each node is reached, the pick-up and drop-off nodes and times for each aircraft, the schedule is deconflicted before performing the vehicle to flight assignment. The route for each aircraft is defined as a series of nodes $n^{\text{route}} = \{n_p^a, ..., n_d^a\}$ and the times the aircraft is to arrive at those nodes as a series of times $t^{\text{route}} = \{t_p^a, ..., t_d^a\}$.

Algorithm 2 summarizes the deconflicting procedure. The procedure runs through a series of time steps, separated from the previous by time step $\Delta = t_{\text{sep}}$, and through all aircraft $a \in A$. If aircraft a will reach a node within this time step, the algorithm checks for two situations. First, it checks whether another aircraft $b \in A$ will reach the same node within the same time step. If there is such an aircraft, the rest of the route of aircraft a is postponed by one time step Δ. Second, when a is about to enter a two-directional taxiway edge $x \in E_X$, the algorithm looks for an aircraft b on the set of edges E_x^{bi} that have to be free from oncoming traffic before a can enter x. For example, let $x = (n_{81}, n_{86})$, see Figure 4.1b. Then $E_x^{\text{bi}} = \{(n_{86}, n_{81}), (n_{87}, n_{86})\}$. If an aircraft b is on either of these edges at the current time step, the rest of the route of aircraft a is postponed by the amount of time needed for b to clear these edges.

Figure 4.2 shows the averaged and total minutes of delay incurred due to the deconflicting procedure, during any ten minute interval during two different periods P. Figure 4.2a shows that very few adjustments are necessary on the morning of March 27th 2022, which is a quiet day, save for the arrival peak from 09:30-10:00. Figure 4.2b shows that on the morning of December 27th 2021, a busy day, many more adjustments are necessary. However, in none of the intervals the average incurred delay exceeds one minute. In addition, one could argue that for actual operation, these delays will be even smaller. This is because part of the delays now attributed to deconflicting during the taxiing period, will in fact be incurred during holding in the air (for arrivals) or waiting at the gate (for departures). This is because the model uses the scheduled off-block and landing times, rather than the actual times.

Algorithm 2 Deconflict routes

Require: $t_a^{\text{route}}, n_a^{\text{route}}\ \forall a \in A, \quad E_x^{\text{bi}}\ \forall$edges x in the set of two-directional edges in E_X
  for $i = 0, 1, 2, \ldots$ **do**
    $t \leftarrow t^s + i\Delta$
    for $a = 0, 1, 2, \ldots, N^F$ **do**
      if $\exists j$: a reaches $n_{a,j}^{\text{route}}$ within Δ **then**
        if $\exists$ aircraft b that reaches $n_{a,j}^{\text{route}}$ within Δ **then**
          for $k = j, \ldots, |n_a^{\text{route}}|$ **do**
            $t_{a,k}^{\text{route}} = t_{a,k}^{\text{route}} + \Delta$
            $t_a^d = t_a^d + \Delta$
          end for
        else if $(n_{a,j}^{\text{route}}, n_{a,j+1}^{\text{route}})$ is a two-directional edge **then**
          $x \leftarrow (n_{a,j}^{\text{route}}, n_{a,j+1}^{\text{route}})$
          for $i_x \in E_x^{\text{bi}}$ **do**
            if $\exists$ aircraft b that will be on i_x within Δ **then**
              for $k = j, \ldots, |n_a^{\text{route}}|$ **do**
                $t_{a,k}^{\text{route}} = t_{a,k}^{\text{route}} + \Delta \frac{t_X(i_x)}{v_x}$
                $t_a^d = t_a^d + \Delta \frac{t_X(i_x)}{v_x}$
              end for
            end if
          end for
        else
          Register that aircraft a reaches node $n_{a,j}^{\text{route}}$ within Δ
        end if
      end if
    end for
  end for



(a) The delay incurred due to the deconflicting procedure on March 27th 2022 from 08:00 to 12:00 (148 aircraft).

(b) The delay incurred due to the deconflicting procedure on December 27th 2021 from 08:00 to 12:00 (286 aircraft).

Figure 4.2: The delay incurred to the deconflicting aircraft during a quiet period and during a busy period.

4.2.3. VEHICLE PATH PLANNING

With the aircraft path planning defined, we now turn to ETV path planning. The ETV fleet V with size N^V is assumed to leave the depot n_{dp} at the start of time period P. Each ETV v travels to the pick-up point n_a^p of its assigned aircraft a via the service road network G_S. It arrives at the aircraft at time $t_a^p - t^c$, where it starts the connecting procedure, which takes t^c. At t_a^p the ETV starts towing the aircraft via the taxiway network G_X to its destination, n_a^d. After arriving there at time t_a^d, the disconnecting procedure takes place, lasting until $t_a^d + t^c$. The ETV will then either wait at its location for some time, start to travel to another assigned aircraft, or start to travel to one of the charging stations $n_{\text{cs},i}$.

As in Section 4.2.2, we use the distance metric d_S defined in Section 4.2.1 to calculate the time needed to traverse the distance between any pair of nodes on G_S as:

$$t_S(m,n) = d_S(m,n)/v_s \quad \forall m,n \in G_S. \tag{4.3}$$

The ETVs travel on the service road network G_S when they do not tow aircraft. The service roads are assumed to provide enough opportunities to pass oncoming traffic. Furthermore, no significant separation distance is required between the ETVs. For these reasons, the vehicle path planning is limited to calculating the shortest paths and associated times between destinations of the ETVs, but no deconflicting procedure is required.

The needed jet engine warm-up procedure is assumed to take place during the towing, and the cool-down procedure takes place when the ETV is disconnected after taxi-in. This means that both procedures have no influence on the ETV path planning.

4.2.4. ETV ENERGY CONSUMPTION AND CHARGING

In order to keep track of the state of charge (SOC) of the individual ETVs, it is necessary to determine how fast the ETV batteries charge at a charging station, and discharge while driving or towing an aircraft. For the latter, we are interested in calculating the power $P(v, m_a)$ consumed by an ETV towing a mass m_a at speed v. This can be approximated by considering the rolling resistance and the thrust as the only forces acting on the ETV or ETV + aircraft combination, disregarding e.g. drag or slope of the airport roads. The power consumed is then equal to the rolling force times the velocity, where the rolling force is equal to the rolling resistance coefficient times the normal force, i.e. gravity. Thus we obtain:

$$P(v,m) = \mu^g(v)\, g\, (m_{\text{ETV}} + m_a)\, v, \tag{4.4}$$

with

$$\mu^g(v) = \mu^0 (1 + \frac{v}{v^0}) \tag{4.5}$$

the rolling resistance coefficient at speed v, m_{ETV} the mass of the ETV, m_a the mass of aircraft a, g the gravitational acceleration, μ^0 the rolling resistance base coefficient and v^0 the rolling resistance base velocity [34].

Using Equation 4.4 we calculate the power consumed by ETV during towing aircraft a, $P(v_x, m_a)$, and during driving, $P(v_s, 0)$. We denote these as P^X and P^S, respectively. The charging power, denoted as P^c, is dependent on the capabilities of the ETV battery and the charging infrastructure at the airport.

In this work a constant speed is assumed to be attained by the ETV at all times, without accounting for acceleration and deceleration. With this information and the expressions for power defined above, we can introduce notation for the energy q required to perform certain movements:

$$q^X(a) = P(v_x, m_a)\, t_X(n_a^p, n_a^d) \qquad \forall a \in A, \tag{4.6}$$

$$q^S(n,m) = P(v_s, 0)\, t_X(n,m) \qquad \forall m, n \in N_S, \tag{4.7}$$

$$q^S(a,b) = q^S(n_a^d, n_b^p) \qquad \forall a, b \in A, \tag{4.8}$$

$$q_f^S(v,a) = q^S(n_v^{\text{old}}, n_a^p) \qquad \forall a \in A, v \in V, \tag{4.9}$$

$$q_l^S(a) = q^S(n_a^d, n_{\text{dp}}) \qquad \forall a \in A, \tag{4.10}$$

$$q^C(a,b) = \min_{i \le N_{\text{cs}}} \{q^S(n_a^d, n_{\text{cs},i}) + q^S(n_{\text{cs},i}, n_b^p)\} \qquad \forall a, b \in A, \tag{4.11}$$

$$q_1^C(a) = \min_{i \le N_{\text{cs}}} \{q^S(n_{\text{cs},i}, n_a^p)\} \qquad \forall a \in A. \tag{4.12}$$

Here $q^X(a)$ is the energy required by an ETV to tow aircraft a along the taxiways, $q^S(n,m)$ is the energy required by an ETV to travel from node n to m along the service roads, $q^S(a,b)$ is the energy required by an ETV to travel from the dropoff point of aircraft a to the pickup point of aircraft b along the service roads, $q_f^S(v,a)$ is the energy required by ETV v to travel from its latest location, n_v^{old}, to the pickup point of aircraft a, $q_l^S(a)$ is the energy required by an ETV to travel from the dropoff point of aircraft a to the depot, $q^C(a,b)$ is the energy required by an ETV to travel from the dropoff point of aircraft a to the pickup point of aircraft b along the service roads and via the closest charging station, and $q_1^C(a)$ is the energy required by an ETV to travel from the closest charging station to the pickup point of aircraft a. Last, we denote the maximum energy capacity of an ETV by Q.

4.2.5. MILP FORMULATION FOR VEHICLE-TO-AIRCRAFT ASSIGNMENT

After the deconflicting procedure, the updated values of the pickup and drop-off times t_a^p and t_a^d are used to generate sets of aircraft used in creating the constraints of the MILP model. Before introducing the MILP formulation, we define:

$$t^C(a,b) = t_b^p - t_a^d - t_S(n_a^d, n_b^p) - 2t^c \qquad \forall a, b \in A, \tag{4.13}$$

which is the available time between towing aircraft a and b that can be used for idling or travelling to charging stations and charging. Then:

$$A_{\text{out}}(a) = \{b \in A : t^C(a,b) > 0\} \qquad \forall a \in A, \tag{4.14}$$

$$A_{\text{in}}(a) = \{b \in A : t^C(b,a) > 0\} \qquad \forall a \in A, \tag{4.15}$$

$$A_{\text{PC}}(a) = \{b \in A_{\text{out}}(a) : q^C(a,b) - q^S(a,b) < P^c(t^C(a,b) - t_{\min}^C)\} \qquad \forall a \in A, \tag{4.16}$$

where $A_{\text{out}}(a)$ is the set of aircraft that can be towed by an ETV after it tows aircraft a, $A_{\text{in}}(a)$ is the set of aircraft that can be towed by an ETV before it tows aircraft a, and

$A_{PC}(a)$ is the set of aircraft that can be towed by an ETV after it tows aircraft a and for which there is at least $t^C_{\min}$ time in between for effective charging. Effective charging is the charging that occurs after the energy loss due to the rerouting to the charging station has been replenished. Note that $A_{PC}(a) \subseteq A_{out}(a)$.

We consider the following decision variables:

$$x_{ab} = \begin{cases} 1 & \text{if } a, b \in A \text{ are towed consecutively} \\ 0 & \text{else} \end{cases} \tag{4.17}$$

$$x^f_{av} = \begin{cases} 1 & \text{if } a \in A \text{ is the first aircraft towed by ETV } v \in V \\ 0 & \text{else} \end{cases} \tag{4.18}$$

$$x^l_a = \begin{cases} 1 & \text{if } a \in A \text{ is the last an ETV tows} \\ 0 & \text{else} \end{cases} \tag{4.19}$$

$$q_a \in [q^X(a), Q] \quad \text{ETV state of charge at the start of towing } a \in A \tag{4.20}$$

$$y_a = \begin{cases} 1 & \text{if } a \in A \text{ is towed by an ETV} \\ 0 & \text{if } a \in A \text{ is taxiing by itself} \end{cases} \tag{4.21}$$

The objective function and constraints are given by:

$$\max_{x,q,y} \quad (1-\alpha)\sum_{a\in A} y_a, \tag{4.22}$$

$$\text{s.t.} \quad \sum_{v\in V} x^f_{av} + \sum_{b\in A^{\text{in}}_a} x_{ba} = y_a \qquad \forall a \in A, \tag{4.23}$$

$$\sum_{b\notin A^{\text{in}}_a} x_{ba} = 0 \qquad \forall a \in A, \tag{4.24}$$

$$x^l_a + \sum_{b\in A^{\text{out}}_a} x_{ab} = y_a \qquad \forall a \in A, \tag{4.25}$$

$$\sum_{b\notin A^{\text{out}}_a} x_{ab} = 0 \qquad \forall a \in A, \tag{4.26}$$

$$q_a \le Q(1-x^f_{av}) + x^f_{av}(q^{\text{first}}_v - q^S_f(v,a)) \qquad \forall a \in A, v \in V, \tag{4.27}$$

$$q_a \ge -Q(1-x^f_{av}) + x^f_{av}(q^{\text{first}}_v - q^S_f(v,a)) \qquad \forall a \in A, v \in V, \tag{4.28}$$

$$0 \le q_a - x^l_a(q^X(a) + q^S_l(a)) \qquad \forall a \in A, \tag{4.29}$$

$$q_b \le q_a - x_{ab}(q^X(a) + q^S(a,b)) + Q(1-x_{ab}) \qquad \forall a \in A, b \in A^{\text{out}}_a \setminus A^{\text{PC}}_a, \tag{4.30}$$

$$q_b \le Q - x_{ab} q^C_1(b) \qquad \forall a \in A, b \in A^{\text{PC}}_a, \tag{4.31}$$

$$q_b \le q_a - x_{ab}(q^X(a) + q^C(a,b) - P^c t^C(a,b)) + Q(1-x_{ab}) \qquad \forall a \in A, b \in A^{\text{PC}}_a, \tag{4.32}$$

$$q_b \ge q_a - x_{ab}(q^X(a) + q^S(a,b)) - Q(1-x_{ab}) \qquad \forall a \in A, b \in A^{\text{out}}_a, \tag{4.33}$$

$$N^V \ge \sum_{a\in A} x^l_a, \tag{4.34}$$

$$\sum_{a\in A} x^f_{av} = 1, \qquad \forall v \in V, \tag{4.35}$$

$$\sum_{v\in V} x^f_{av} \le 1, \qquad \forall a \in A, \tag{4.36}$$

The objective (4.22) is to maximize the number of aircraft towed by the fleet of ETVs. The factor $(1-\alpha)$ preceding the expression is there to ensure consistency with the disrupted model, which will be described in the next section. Constraints (4.23) and (4.25) ensure that every aircraft a that is towed by an ETV, is either the first (last) to be towed by an ETV, or has another aircraft b preceding (following) it. Constraints (4.24) and (4.26) ensure that an aircraft b that is not in A^{in}_a (A^{out}_a) cannot precede (follow) aircraft a. Constraints (4.27), (4.28) and (4.29) make sure that ETVs start with a full battery at their first task, and have enough energy at the end of the period P to reach the depot. ETV v starts with a state of charge of q^{first}_v at t^s. In the strategic model the ETVs are assumed to start with a full battery ($q^{\text{first}}_v = Q$). Later, in the disrupted model, the state of charge of ETV v in any assignment will be taken from the previously generated assignment.

Constraint (4.30) subtracts the energy spent on towing aircraft a from the state of charge of aircraft b, if no charging has taken place in between. Constraint (4.31) limits

the state of charge to Q after a period of charging. Constraint (4.32) adds the energy gained from charging to the SOC of an ETV. Constraint (4.33) limits the reduction in SOC to only what was spent on towing aircraft a. Constraint (4.34) limits the number of used ETVs to the maximum of N^V. Constraint (4.35) ensures that for every ETV v, only one aircraft is the first to be towed by that ETV. Constraint (4.36) ensures that for every aircraft a, no more than one ETV is assigned to be the first to tow it.

The advantage of MILP formulation (4.22)-(4.36) is that it is possible to determine all times and durations of all activities of the ETVs and aircraft from the given decision variables, without having to introduce a variable x_{abv}. This would create $N^F \cdot N^F \cdot N^V$ variables, whereas now the number of variables remains of the order $N^F \cdot N^F$. Recall that N^F is the number of flights, and N^V the number of ETVs. The usage of x^f_{av} rather than x^f_a (parallel to x^l_a) ensures that the model assigns a task list of aircraft to a specific vehicle, rather than obtaining N^V task lists that are not tied to any specific vehicle. This will be essential in the disrupted model, which is introduced in the next section.

4.3. Disruption Management of Electric Taxiing Vehicles Towing Aircraft

The disrupted model is used to obtain a vehicle-to-aircraft assignment based on a flight schedule that partly consists of actual departure and arrival times, rather than scheduled times. When creating an assignment for a given day of operations, both models are used. Figure 4.3 shows in what order they are used. An initial assignment of the first four hours of operation is made in advance of the start time t^s, using the strategic model. After this, the disrupted model is run every half hour for the four hours of operation directly after that moment, so that time t^s occurs right after creating the assignment. It is assumed that the flight delays of aircraft departing and arriving in the coming half hour are known at the moment of creating the assignment. That means that for this period, the actual times of arrival and departure are used to compute the assignment, rather than the scheduled times. At the first instance of running the disrupted model, the previously obtained assignment is the result of the strategic model. At the second and later instances, the previous assignment has come from the disrupted model.

The goal of the disrupted model is to obtain an assignment that is close to the previously obtained assignment, i.e. we would like to minimize the number of changes to the towing assignment for ETV operators. In order to quantify this, we introduce an additional decision variable for the disrupted model:

$$s_v = \begin{cases} 1 & \text{if for vehicle } v \in V \text{ the list of aircraft to tow has remained the same as in the previous assignment} \\ 0 & \text{else.} \end{cases} \tag{4.37}$$

When creating an assignment for the next time period P, the aircraft that are towed by vehicle v in the first half hour of the previous time period will have disappeared from the schedule. Likewise, new aircraft will be added to the schedule for the last half hour of this new period. The "list of aircraft to tow" in Equation 4.37 refers to all but these

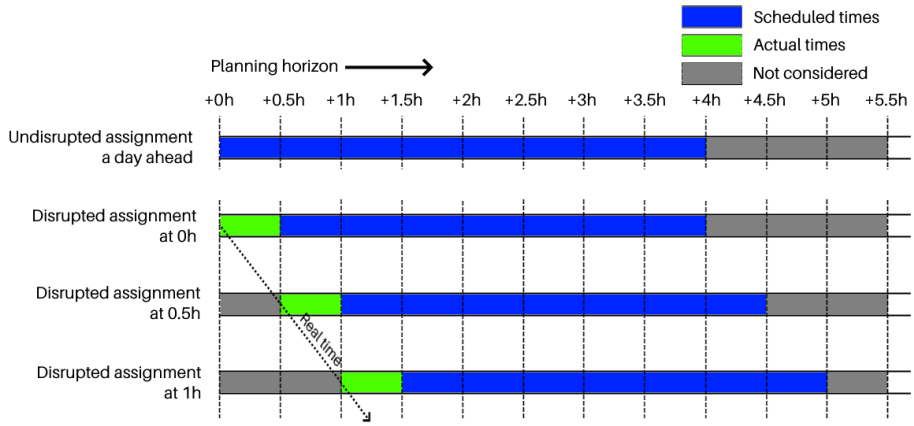


Figure 4.3: Usage of the strategic and disrupted model when creating vehicle-to-aircraft assignments during a day of operations, with an indication of when scheduled or actual arrival and departure times are considered.

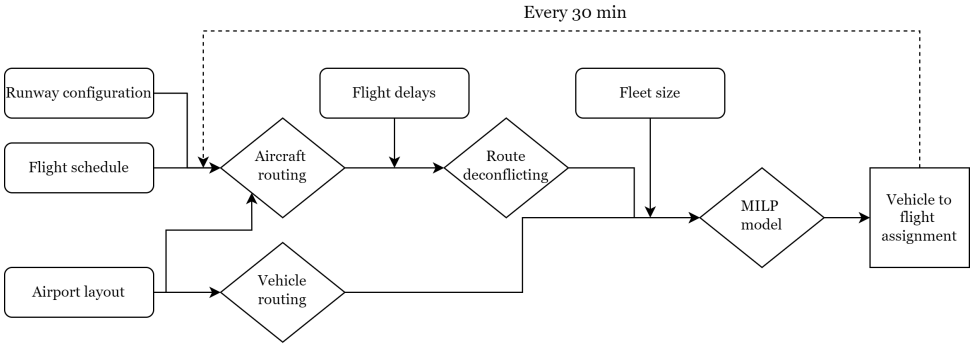


Figure 4.4: Flowchart for the usage of the disrupted model, containing inputs, outputs and processes. Every 30 minutes, a towing assignment is created.

aircraft. The variable has the value 1 if the aircraft are still to be towed by vehicle v in the new assignment, and in the same order as before.

The objective function (4.22) of the strategic model is replaced by Equation (4.38):

$$\max_{x,q,y,s} \quad (1-\alpha)\sum_{a\in A} y_a + \alpha \sum_{v\in V} s_v. \tag{4.38}$$

The objective function (4.38) now contains two terms; the number of aircraft towed, and the number of vehicles with an unchanged schedule. By varying the weight factor α, the trade-off between these terms can be controlled. Since both terms will be used many times throughout this chapter, we introduce shorthand notation, reducing objective function (4.38) to:

$$\max_{x,q,y,s} \quad (1-\alpha)Y + \alpha S. \tag{4.39}$$

4

Constraints (4.23)-(4.36) are the same as in the strategic model. Some additional constraints are added:

$$M(1-s_v) \geq N_v^F + 1 - \sum_{i=1}^{N_v^F} x_{f_{vi}t_{vi}} - x^f_{f_{v0}v} \qquad \forall v \in V, \tag{4.40}$$

$$1 - s_v \leq N_v^F + 1 - \sum_{i=1}^{N_v^F} x_{f_{vi}t_{vi}} - x^f_{f_{v0}v} \qquad \forall v \in V, \tag{4.41}$$

$$x^f_{av} = 1, \qquad \forall a \in A, v \in V : x^{f,\text{old}}_{av} = 1 \tag{4.42}$$

where $N_v^F + 1$ is the number of aircraft that are towed sequentially in the previous solution for ETV $v \in V$ and also appear in the current period P, f_{vi} is the i-th aircraft from which ETV v departed in the previous solution, t_{vi} is the i-th aircraft at which ETV v arrived in the previous solution, M is a large value, and $x^{f,\text{old}}_{av}$ is the value of x^f_{av} from the previous assignment.

Constraints (4.40) and (4.41) define s_v. Constraint (4.42) guarantees continuity between the current and previous assignment, by assigning the aircraft that were being towed when the new assignment was made to the correct ETVs. The path planning, deconflicting procedure and energy module remain the same as for the strategic model.

Figure 4.4 shows a flowchart for the usage of the disrupted model introduced in this section. It shows where the inputs are used, e.g. the flight delays are first applied to the aircraft routes, before the routes are deconflicted. After obtaining the solution, it is used in the next time period to obtain the new aircraft routes.

For completeness, the assumptions applicable for the scheduling models in this chapter are summarized below. For both models:

- The ETV fleet consists of ETVs for narrow-body aircraft.
- Any aircraft can perform electric taxiing or regular taxiing.
- The ETVs have a linear charging profile.
- Routes for aircraft are deconflicted in advance of scheduling.

- ETVs drive on service roads, ETVs tow aircraft on taxiways, and aircraft taxi on taxiways.

In addition, for the disrupted model:

- The flight delays of aircraft departing and arriving in the coming half hour are known at the moment of creating the assignment.

4.4. RESULTS

In this section the strategic and disrupted model are applied to create vehicle-to-aircraft assignments for aircraft arriving and departing at the reference airport Amsterdam Airport Schiphol. Flight schedules are taken from historical data, provided by the Schiphol Developer Center [147]. From these flight schedules the gate numbers and the scheduled and actual off-block time or landing time are obtained. The runway configuration of any day in the past is available from Dutch Air Traffic Control [101, 38].

The choices of parameter values are summarized in Table 4.1.

Table 4.1: Parameter values for the assignment models.

Symbol	Name	Value	Unit	Source
d_{sep}	Separation distance	200	m	[182]
v_s	Speed on service roads	30	km/h	[143, 116]
v_x	Speed on taxiways	42	km/h	[157]
t^c	Connecting/Disconnecting time	3.0	min	[144]
t^C_{min}	Minimum charging time	30	min	
μ^0	Rolling resistance base coefficient	0.010	-	[34]
v_0	Rolling resistance base velocity	41.2	km/h	[34]
m_a	Aircraft mass	$8.0 \cdot 10^4$	kg	
m_{ETV}	ETV mass	$1.45 \cdot 10^4$	kg	[16]
P^c	Charging power	100	kW	[124]
Q	Battery capacity	400	kWh	[124]

The model assumes one weight class of aircraft, with a mass of $8.0 \cdot 10^4$ kg (the maximum take-off weight of a Boeing 737-800), and one type of ETV. The model can be extended to include several weight classes in both the aircraft and the ETVs, see for example Oosterom and Mitici [124]. The mass of the ETV is calculated as

$$m_{\text{ETV}} = m_0 + m_q Q, \tag{4.43}$$

where m_0 is the base mass and m_q is the battery energy density. Given a base mass of $1.2 \cdot 10^4$ kg [16] and an energy density of 6.25 kg/kWh [124], we arrive at the value given in Table 4.1.

The minimum charging time is chosen such that ETVs are not allowed to travel to charging stations to only charge for a few minutes. An ETV will charge either for at least t^C_{min} or until the state of charge reaches Q.

Table 4.2: Objective values for vehicle-to-aircraft assignments of December 27th 2021 with fleet size $N^V = 40$. Shown are the model type, start time t^s, end time t^e, objective value Equation 4.39, the total number of towed aircraft Y, the number of flight movements N^F, the total number of unchanged schedules S and the solution time.

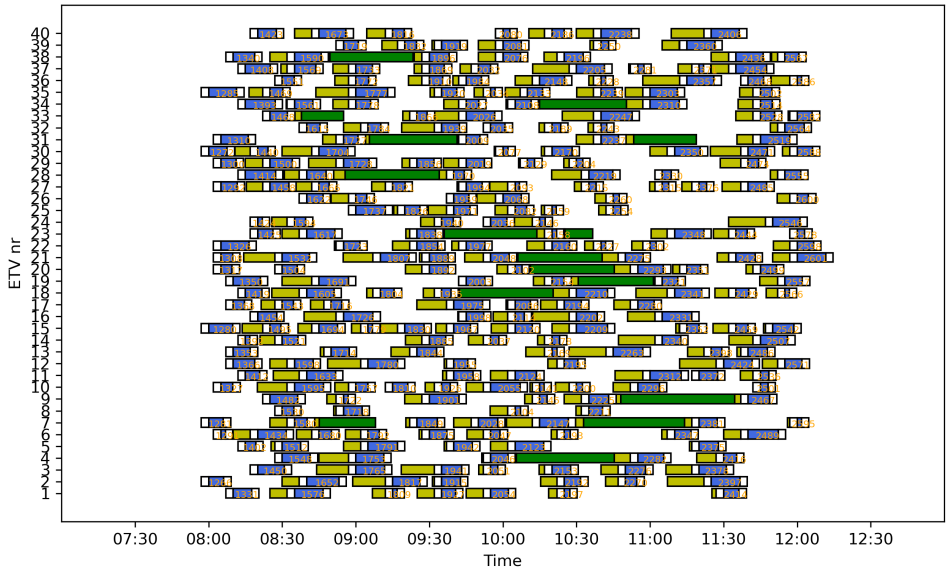
Model	t^s	t^e	Objective	Y	N^F	S	Solution time [s]
Strategic	08:00	12:00	00.286	286	286	N/A	39.5
Disrupted	08:00	12:00	21.259	280	280	21	18.6
Disrupted	08:30	12:30	29.266	295	297	29	15.5
Disrupted	09:00	13:00	26.261	287	289	26	13.6
Disrupted	09:30	13:30	15.269	294	284	15	21.3
Disrupted	10:00	14:00	09.266	275	275	09	14.6
Disrupted	10:30	14:30	22.253	275	275	22	16.1
Disrupted	11:00	15:00	24.231	255	258	24	11.7
Disrupted	11:30	15:30	18.234	252	255	18	12.7
Disrupted	12:00	16:00	22.229	251	251	22	10.8
Disrupted	12:30	16:30	18.225	243	244	18	11.8
Disrupted	13:00	17:00	21.225	246	246	21	12.5
Disrupted	13:30	17:30	22.212	234	236	22	09.8
Disrupted	14:00	18:00	21.205	226	226	21	13.5
Disrupted	14:30	18:30	30.181	211	211	30	06.9
Disrupted	15:00	19:00	32.175	207	207	32	06.6

4.4.1. EXAMPLE DAY OF OPERATIONS

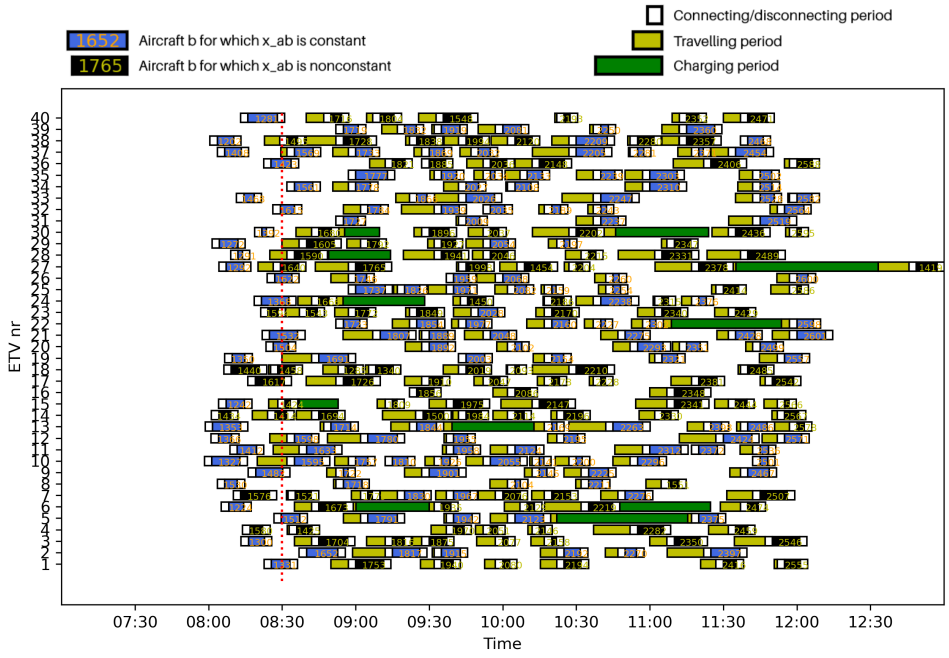
Figures 4.5 and 4.6 show the vehicle-to-aircraft assignments obtained using the strategic and disrupted models for December 27th 2021. First, in Figure 4.5a, the strategic model is run to obtain an initial assignment for the period 08:00-12:00, based on the scheduled off-block and landing times. Then, in Figure 4.5b the disrupted model is run for the same period, but now using the actual times for the first half hour. In Figure 4.6a and Figure 4.6b the period is shifted by half an hour and the disrupted model is applied again. The calculations were performed using Gurobi 9.5 on a Dell Latitude 7490 laptop with an Intel i7-8650U CPU of 1.90 GHz.

In this example, weight factor α is set to 0.999, i.e. the second term of objective function (4.38) from the disrupted model is made the most important; S is prioritized over Y. Furthermore, N^V is set to 40, which is enough to tow all aircraft that arrive and depart during the time period P.

Using black and blue colouring, Figures 4.5 and 4.6 show for individual aircraft whether they are still following the same aircraft as in the previous assignment, i.e. if for aircraft b the value of $x_{ab} = 1$ in the previous assignment, then this is still the case in the current assignment. When a towing assignment has changed for an ETV, we can see that this is due to delays of the aircraft in this schedule. Take for example ETV 24 in Figure 4.5b: aircraft 1666 follows aircraft 1368. However, in Figure 4.6a we see that this cannot be maintained: aircraft 1666, which had a scheduled landing time of 08:45, arrives 11 minutes early. ETV 24 would not have enough time to travel between these tasks, i.e. from

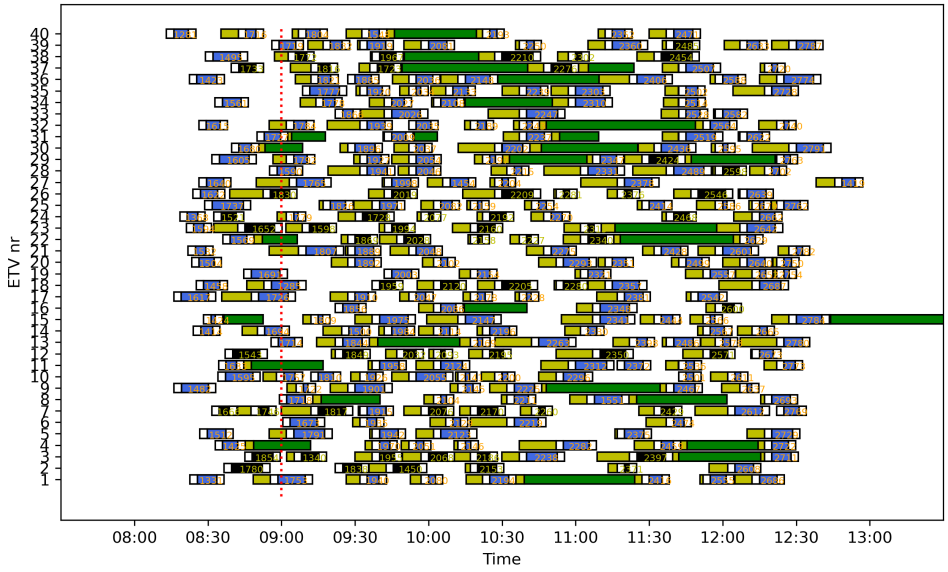


(a) The vehicle-to-aircraft assignment obtained using the strategic model for 08:00-12:00.

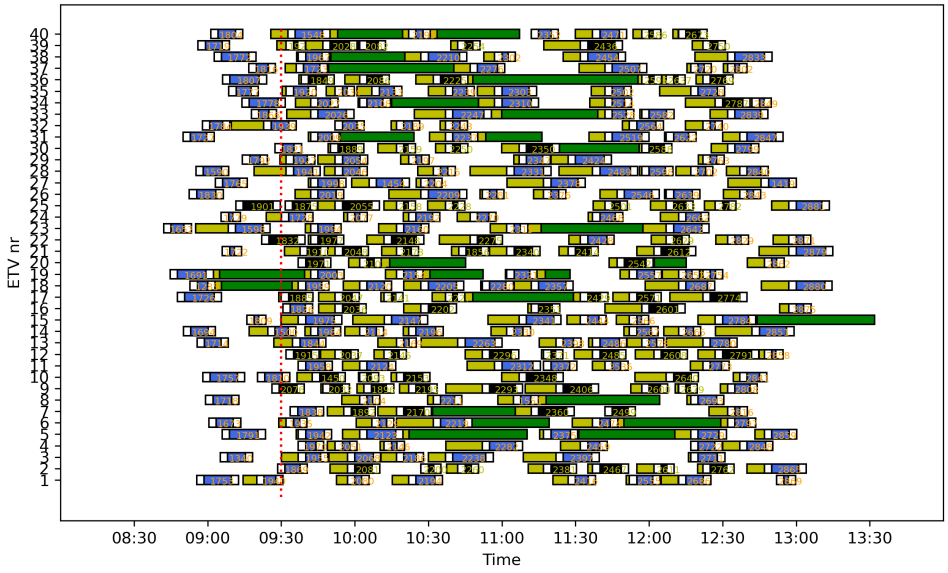


(b) The vehicle-to-aircraft assignment obtained using the disrupted model for 08:00-12:00.

Figure 4.5: Vehicle-to-aircraft assignments for December 27th 2021. A red dashed line indicates the time until which the disruptions are considered known.



(a) The vehicle-to-aircraft assignment obtained using the disrupted model for 08:30-12:30.



(b) The vehicle-to-aircraft assignment obtained using the disrupted model for 09:00-13:00.

Figure 4.6: Continuation of Figure 4.5.

n^d_{1368} to n^p_{1666}. In the new schedule aircraft 1666 is towed by ETV 7.

From these figures we can deduct that the disrupted model is able to keep the list of aircraft to tow the same for a significant amount of the ETVs. Table 4.2 shows the results of applying the disrupted model throughout the day. The number of unchanged schedules, i.e. $S = \sum_{v \in V} s_v$, displayed in the penultimate column, appears to be roughly inversely correlated with the total number of aircraft in the time period. This corresponds with intuition: if the number of aircraft is smaller, but the number of ETVs remains the same, every ETV has fewer aircraft to tow, and its schedule is less prone to the disruptions of flight delay. Second, by averaging the values for S we obtain that on December 27th 2021, roughly 22 out of 40 vehicles, i.e. 55%, receive an unchanged schedule in every time period.

Figure 4.7 shows the state of charge of every ETV during two of the periods appearing in Table 4.2. In Figure 4.7a we see that all ETVs start with a state of charge near the maximum, Q, and are sometimes recharged to full capacity within the time period P. In Figure 4.7b we see that some ETVs are being recharged by a considerable amount, but many of them have run out of charge.

4.4.2. TRADE OFF BETWEEN OBJECTIVE TERMS

The objective function (4.38) consists of two terms. It maximizes both the number of aircraft towed, Y, and the number of ETVs for which the assignment is unchanged, S. In this section we investigate the effect of prioritizing one term over the other.

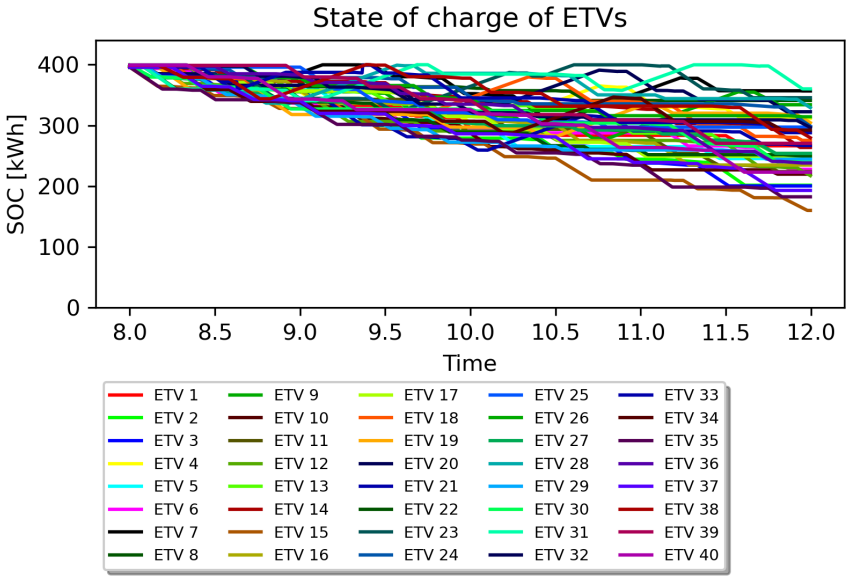
We consider a situation where not all aircraft can be towed by the fleet of ETV: let $N^V = 25$. Then we run the disrupted model, while varying α. The results are shown in Table 4.3, where values for both terms of the objective function (4.38) are displayed for several values of α. We see that when decreasing α, the number of unchanged schedules S decreases, and the number of towed aircraft Y increases slightly. When $\alpha = 0.3$, an extra aircraft towed contributes the same to the objective as two more ETVs with unchanged schedules. When $\alpha = 0.999$, the unchanged schedules contribute a thousand times more to the objective than the towed aircraft. When $\alpha = 0.001$ we obtain the opposite. From Table 4.3 it can be concluded that for December 27th 2021 and with the current model parameters, the results do not vary greatly with varying α values. In order to obtain the maximum number of unchanged schedules S, a maximum of five aircraft (2%) will have to taxi by themselves. This means that it is possible to pursue practicality for ETV operators (number of unchanged schedules), while retaining the minimization of environmental impact (number of towed aircraft).

4.5. SENSITIVITY ANALYSIS

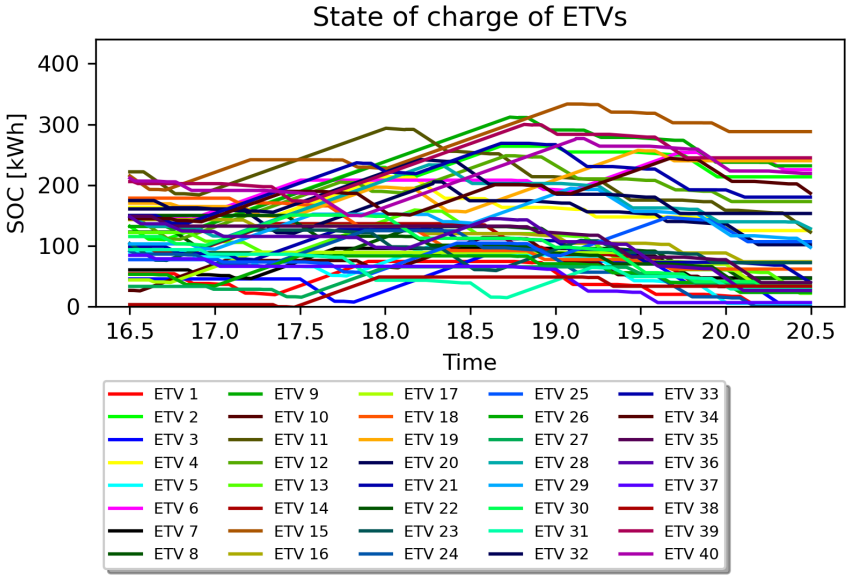
In this section, we analyze the performance of the MILP models introduced in Sections 4.2 and 4.3 when changing certain model parameters or inputs.

4.5.1. THE IMPACT OF FLEET SIZE ON VEHICLE-TO-AIRCRAFT ASSIGNMENTS

An important parameter to investigate is the fleet size N^V. If N^V is large enough, all aircraft can be towed by the fleet, i.e. $Y = N^F$. For lower values of N^V, the model has to make a selection from all $a \in A$. In subsequent time periods, the selection of flights will



(a) State of charge graphs during 08:00-12:00, obtained with the disrupted model.



(b) State of charge graphs during 16:30-20:30, obtained with the disrupted model.

Figure 4.7: State of charge of all ETVs during two different time periods on December 27th 2021 at Schiphol Airport.

Table 4.3: Objective values for vehicle-to-aircraft assignments with varying α of December 27th 2021 with $N^V = 25$.

			$\alpha = 0.001$		$\alpha = 0.3$		$\alpha = 0.7$		$\alpha = 0.999$	
Model	t^s	t^e	Y	S	Y	S	Y	S	Y	S
Strategic	08:00	12:00	266	N/A	266	N/A	266	N/A	266	N/A
Disrupted	08:00	12:00	263	6	263	3	262	8	260	8
Disrupted	08:30	12:30	280	6	280	4	276	10	275	9
Disrupted	09:00	13:00	272	7	271	7	269	8	271	8
Disrupted	09:30	13:30	272	3	272	2	271	1	271	4
Disrupted	10:00	14:00	269	0	269	0	268	1	268	1

Table 4.4: Number of unchanged schedules, S, for vehicle-to-aircraft assignments obtained with the disrupted model, when varying the fleet size N^V, on December 27th 2021 with weight factor $\alpha = 0.999$.

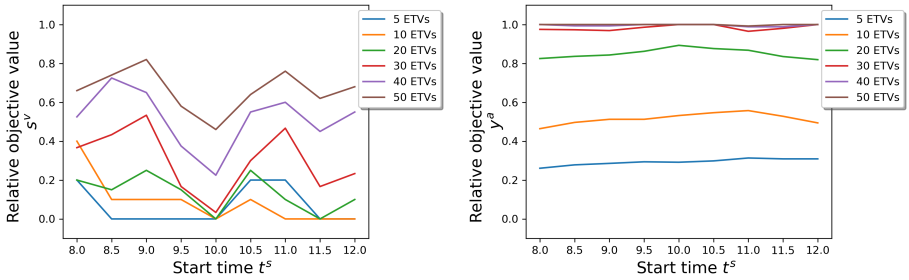
t^s	t^e	$N^V = 5$	$N^V = 10$	$N^V = 20$	$N^V = 30$	$N^V = 40$	$N^V = 50$
08:00	12:00	1	4	4	11	21	33
08:30	12:30	0	1	3	13	29	37
09:00	13:00	0	1	5	16	26	41
09:30	13:30	0	1	3	5	15	29
10:00	14:00	0	0	0	1	9	23

be largely the same, since the goal of the disrupted model is to minimize the number of unchanged schedules. Table 4.4 and Figure 4.8 show the obtained objective values when varying the fleet size. First consider small N^V: observe that for any $\alpha < 1$ the number of towed aircraft is still being maximized. This means any ETV will be assigned many aircraft to tow, with little idle time in between. Therefore, the impact of flight delays on individual towing asignments is large, and Y is small. Now consider N^V large enough so that $Y \approx N^F$. Figure 4.8b shows that this is the case for $N^V \geq 30$. Here S/N^V exceeds 0.3, while it does not for smaller N^V. In such an assignment more time is available between towing tasks, leading to a smaller impact of flight delays, and therefore a larger number of towed aircraft. Figure 4.8a shows that this not only holds in absolute numbers of S, but also relative to the maxima, i.e. S/N^V.

4.5.2. APPLYING THE MODEL TO VARIOUS DAYS OF OPERATION

Figure 4.9 shows the average delay for Schiphol Airport during October 2021 to January 2022. The average of all departure delay averages is 15 minutes, and the average of arrival delay averages is -5 minutes. In addition, the standard deviation is very large. Notable outliers include October 21st 2021 and January 31st 2022, when traffic was disrupted due to weather circumstances [12]. An example of a period with extended large delays is May and June 2022, when Schiphol experienced staff capacity problems. During these months the average departure delay was 31 minutes, and the average arrival delay was 9 minutes.

In order to investigate how the results are affected if the model is faced with more and greater delays, we compare the results of applying the disrupted model with 20 ETVs for



(a) Relative number of unchanged schedules for varying start time t^s and fleet size N^V.

(b) Relative number of aircraft towed for varying start time t^s and fleet size N^V.

Figure 4.8: Relative value of unchanged schedules, S/N^V, and aircraft towed, Y/N^F, in the assignments created with the disrupted model ($\alpha = 0.999$), for varying starting times t^s, on December 27th 2021.

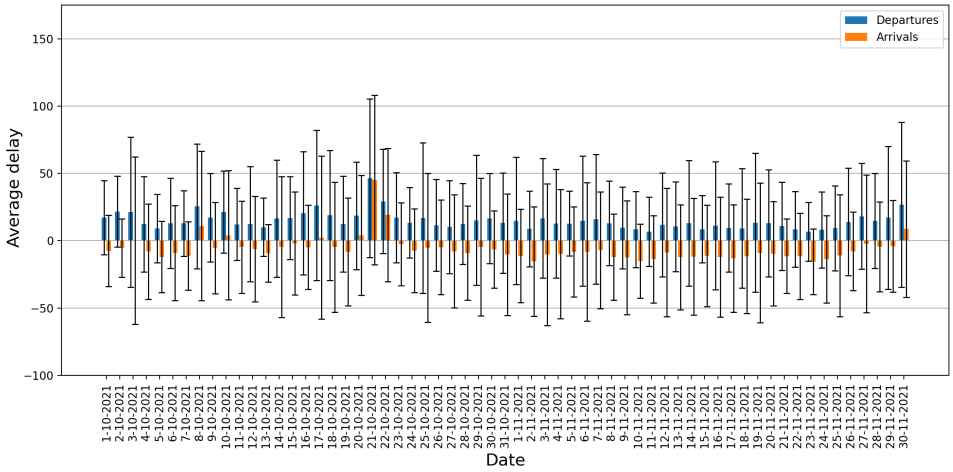
December 27th 2021 (departure delay 15 minutes, arrival delay 2 minutes), to November 30th 2021 (departure delay 27 minutes, arrival delay 9 minutes), and October 21st 2021 (departure delay 47 minutes, arrival delay 45 minutes). In Figure 4.10a the metrics for these days are visualized. The figure includes the values obtained for S and Y, as well as those for N^F, i.e. the maximum value for Y. From Figure 4.10a we note that on November 30th, there were significantly fewer aircraft to tow than on the other days, which means that 20 vehicles is almost enough to tow all aircraft in this case. However, this does not lead to an improved performance in the value of S. Another observation is that a sharp decrease in Y, such as at t^s =14:00 on December 27th, allows for a larger value of S in subsequent time periods. This corresponds to intuition: when the towing assignment becomes relatively empty, there is more space to accommodate flight delays.

Last, an important observation to take from Figure 4.10a is that there is no clear decrease in performance regarding S when the average flight delays of a day is larger. On October 21st there are roughly as many aircraft to tow as on December 27th, but the average delay is much larger. Still, the number of unchanged schedules obtained with the disrupted model is similar throughout the morning and afternoon. Figures 4.10b to 4.11b provide a more detailed illustration of the delays on the tested days. It is clear that especially on October 21st, there are many more large delays than on December 27th.

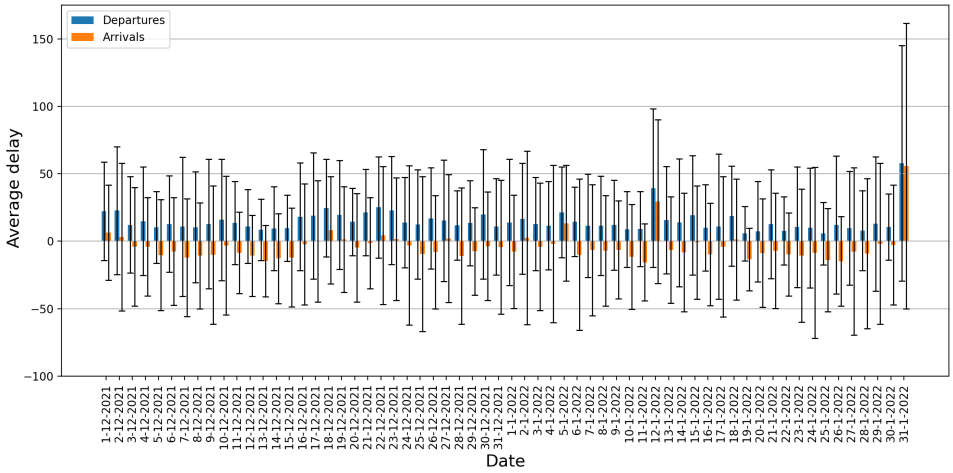
A possible explanation for this difference in delays, but similarity in the value of S, is the following: delays are assumed to become available to the model when the scheduled flight time is less than half an hour away. This holds for large and small delays. This means that both large and small delays impact the vehicle-to-aircraft assignment in a similar way. Only for very small delays the aircraft will not have to be moved out of its current place in the assignment. For this reason, the model performs similar regardless of the severity of the delays.

4.6. Conclusion

This chapter proposes an approach to disruption management for the assignment of electric taxiing vehicles (ETVs) to aircraft. First, a strategic Mixed Integer Linear Pro-

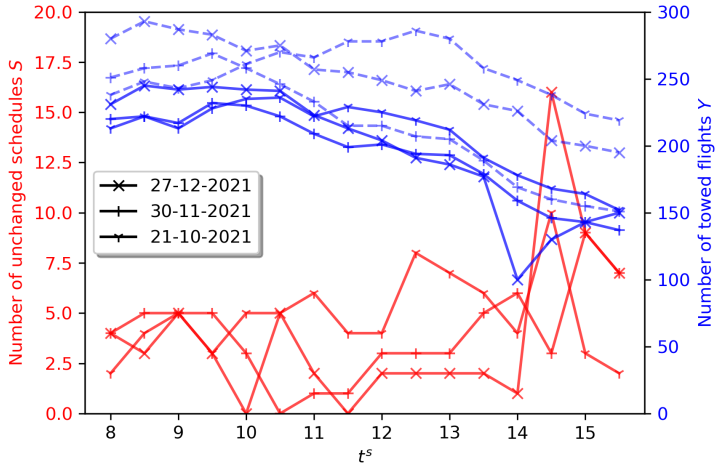


(a) Average daily delay for dates in October and November 2021.

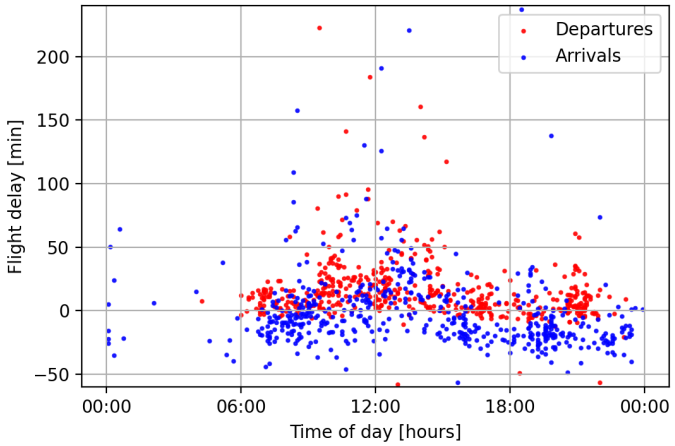


(b) Average daily delay for dates in December 2021 and January 2022.

Figure 4.9: Average daily departure and arrival delay including standard deviation for Amsterdam Airport Schiphol during October 2021 to January 2022.

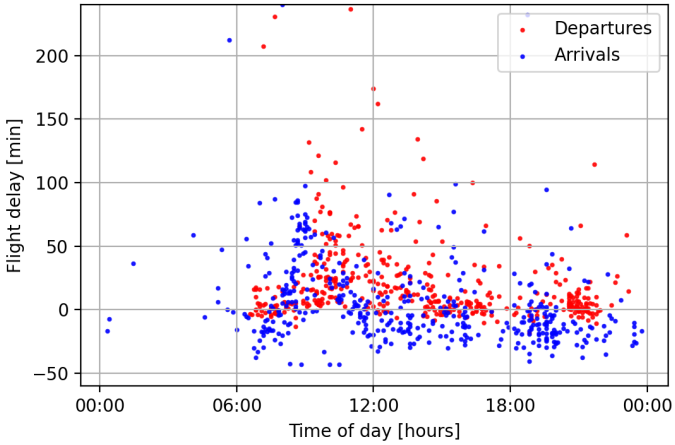


(a) Results of applying the disrupted model to three days. The red solid lines represent the values for S, the blue solid lines the values for Y, and the blue dashed lines the values for the total number of flight movements N^F.

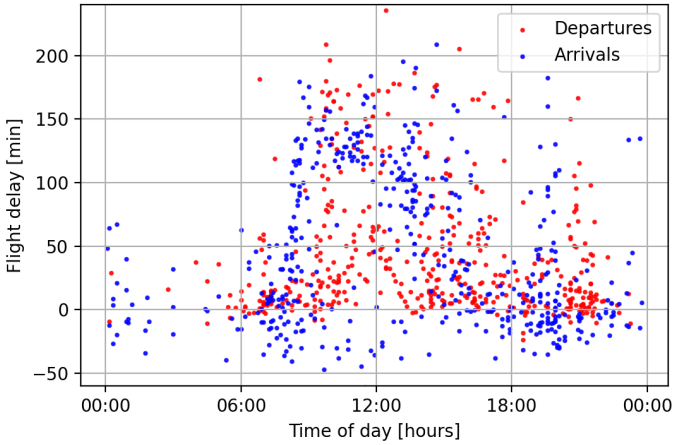


(b) Scatter plot of departure and arrival delays on December 27th 2021.

Figure 4.10: Outcomes of applying the disrupted model for three days with varying degrees of on-time performance using a fleet size of $N^V = 20$ at Amsterdam Airport Schiphol, including scatter plots of delays.



(a) Scatter plot of departure and arrival delays on November 30th 2021.



(b) Scatter plot of departure and arrival delays on October 21st 2021.

Figure 4.11: Continuation of Figure 4.10.

gramming (MILP) model was formulated, which outputs the vehicle-to-aircraft assignment including the state of charge of each vehicle. Second, a disrupted MILP model was introduced, which takes into account the flight delays of the coming half hour when generating its disrupted assignment. By running the disrupted model in staggered time periods, one creates an adaptive vehicle-to-aircraft assignment throughout the day. Both models take the flight schedule and airport layout as inputs, which are used to calculated the vehicle and aircraft routes, for which conflict avoidance is ensured. Both models keep track of the energy spent and gained by the ETVs, and select the times the vehicles are to be charged.

The models are illustrated using a case study of Amsterdam Airport Schiphol. The models are shown to be able to generate vehicle-to-aircraft assignments that correspond to the flight schedules of this hub airport, which include departing and arriving flights that are towed from the gate area to one of the five runways and vice versa. The results show that it is possible to minimize the number of changed schedules, without having to reduce the number of aircraft from the flight schedule that are towed by an ETV, which is beneficial for ETV operators. Furthermore, sensitivity analysis has revealed that the severity of the delays does not impact the model performance, and that if enough ETVs are available to tow all aircraft in the schedule, the number of unchanged schedules increases above 30% of the fleet size.

Future work can include implementing a more detailed routing and energy model within the introduced models by including for example acceleration of aircraft and vehicles, rather than constant velocities. Furthermore, the approach presented in this chapter can be tested in combination with varying charging strategies, for example: a strategy where night charging is not preferred, and the ETVs should sustain their state of charge throughout the day.

5

PROBABILISTIC FLIGHT DELAY PREDICTIONS USING MACHINE LEARNING AND APPLICATIONS TO THE FLIGHT-TO-GATE ASSIGNMENT PROBLEM

In Chapter 4 a model was developed to adapt Electric Taxiing Vehicles to aircraft assignments, when faced with flight delays. A step beyond reacting to disruptions is to predict disruptions. In this chapter, two probabilistic forecasting algorithms, Mixture Density Networks and Random Forest regression, are applied to predict flight delays. The resulting probabilistic predictions are expected to provide more extensive information to airport planners than point predictions. They are evaluated with tailored metrics. To illustrate the utility of the estimated delay distributions, they are integrated into a case study of a probabilistic flight-to-gate assignment problem, which aims to increase the robustness of flight-to-gate assignments. It is shown that our proposed flight-to-gate assignment model reduces the number of gate-conflicted aircraft when compared to a deterministic flight-to-gate assignment model. In general, the results illustrate the utility of considering probabilistic forecasting for robust airport operation optimization.

This chapter is based on the following research article:
Zoutendijk, M. & Mitici, M., & Hoekstra, J. M. (2021). "Probabilistic flight delay predictions using machine learning and applications to the flight-to-gate assignment problem." in *Aerospace*, 8.6, 152 [191].

This paper has been awarded *Honourable Mention* in the *Anna Valicek Competition* of the *Airline Group of the International Federation of Operational Research Societies*, in 2022.

5.1. INTRODUCTION

On-time flight performance is an important measure of the service quality of airports and airlines. During the period 2013–2019, while the number of flights in Europe increased by 16% [44], the average departure delay of European flights increased by 41% [42]. Such an increase has a negative impact on the airports' and airlines' quality of service. As Eurocontrol forecasts the number of flights to be restored to 2019 levels by 2024 [43], large increases in delay can be expected again in the future. Accurate flight delay predictions will therefore remain central to support airports and airlines in offering a high-quality service.

In the past years, several machine learning algorithms have been proposed to predict flight delays. Most studies predict flight delays using (i) binary classifiers (delayed/not delayed flight), (ii) multi-class classifiers (multiple delay classes), or (iii) estimating the delay value.

Binary classifiers are proposed in Kim et al. [90] where recurrent neural networks are used to predict flight delays at airports in the US. The prediction horizon is several hours before the operation. Using this approach, delays are predicted with an accuracy of 0.87. In Lambelho et al. [95], binary classification of flight delays and cancellations is performed for Heathrow airport using three different classification algorithms: LightGBM, Multilayer Perceptron, and Random Forests. The authors predict flight delays and cancellations with an average F1-score of 0.56 using the LightGBM classifier. In Choi et al. [32], the authors propose binary classifiers for flight delays, and applies these with two prediction horizons, one of five days and one of one day. The obtained flight delay predictions have an accuracy of 0.80 using the Random Forest classifier.

Multi-class departure delay predictions are obtained in Alonso and Loureiro [9] for Porto airport for a prediction horizon of several hours before the operation. In Chen and Li [31], flight delay is predicted using multi-label Random Forests classification. Flight delay values from routes flown by an aircraft earlier in a day are used to predict flight delay for the routes flown later in a day.

In Kalliguddi and Leboulluec [85], flight delays are estimated hours ahead of the operation using machine learning algorithms that perform regression. The authors consider delay states of the aviation network as features, in addition to flight schedule-related features. The results obtained using Random Forests have a root mean square error (RMSE) of 12.5 min. It is also shown that the delay states have the largest effect on on-time performance. In Manna et al. [106], the obtained flight delay predictions have an RMSE of 8.2 min and 10.7 min when considering departure delays and arrival delays, respectively. In Yu et al. [178], a deep-belief network is used to predict flight delays several hours before the operation. A reduction of 21% in the RMSE is obtained compared to the best benchmark algorithm, the k-Nearest Neighbours. Thiagarajan et al. [164] propose both classification and regression algorithms to predict flight delay. Here, the regression approach using Random Forests produced an RMSE of 8.7 min. Ayhan, Costas, and Samet [15] and Shao et al. [154] introduce features based on flight trajectory data. Ayhan, Costas, and Samet [15] predict flight delays for domestic flights in Spain within an RMSE of 4 min. A range of prediction algorithms is employed, of which AdaBoost performs best. Shao et al. [154] find that the features based on trajectory data contribute the greatest to the predictive accuracy, and the best result is found using LightGBM.

The classification and regression results obtained in these studies generate an estimate for individual flight delay in the form of a class or a point estimate, respectively. The estimates are often evaluated using metrics based on the confusion matrix and metrics such as RMSE/MAE (Mean Absolute Error), respectively. In order to plan flight operations such as gate allocation or runway allocation in a robust manner, however, it is necessary to also consider the uncertainty of the predicted delays of individual flights. Such measures are not included when obtaining delay classes or point estimates, nor can they be derived directly from the commonly used evaluation metrics. Therefore, in this chapter, we propose to estimate the probability distribution of flight delays on an individual flight basis, using machine learning algorithms. Such probability distributions can support planners to robustly plan flight operations.

Very few studies estimate the probability distribution of flight delays. The common approach is to fit historical delays to *one* probability distribution which is assumed to be representative for *all considered* flights [114, 121, 80, 92, 166]. In Mueller and Chatterji [114] and Novianingsih and Hadianti [121], airport and airline delay distributions are obtained by fitting historical delays to classes of probability distributions. Tu, Ball, and Jank [166] introduce a more complex model, where the national airspace delay distribution is assumed to be the sum of seasonal trends, a daily propagation pattern and random residuals. To the best of our knowledge, however, no studies have been performed that estimate a probability distribution for flight delays on an individual flight basis, i.e., probabilistic flight delay prediction.

To illustrate how probabilistic flight delay prediction on an individual basis can be useful for operation optimisation, we integrate these predictions into a probabilistic flight-to-gate assignment problem (FGAP). Şeker and Noyan [194] were among the first ones to incorporate probabilistic effects in their solution method for the FGAP. The authors evaluate the robustness of FGA's by modelling the departure and arrival flight delays as random variables. A set of scenarios is created, each with random disruptions to flight arrival and departure times. The number of gate conflicts is then minimized for each scenario. The random disruptions utilized in this study model flight delay; however, they are not based on delay predictions. The results of this study provide a general overview of the robustness of the used optimization methods, but it is not possible to directly evaluate the robustness using the actual delay experienced at the airport. Van Schaijk and Visser [171] and L'Ortye, Mitici, and Visser [100] determine the probability that a given arriving/departing aircraft is present at a gate, for a range of time values. This is called the aircraft presence probability, and it is obtained using a regression model based on historic data of aircraft gate presence, using the features 'airline identity' and 'origin/destination region of flight.' The aircraft presence probability of an arriving aircraft is in fact the cumulative distribution function (cdf) of the aircraft's delay. The presence probability of a departing aircraft is the inverted cdf of the aircraft's delay. Using these presence probabilities, robust flight-to-gate assignments are developed. The approach taken by Van Schaijk and Visser [171] makes use of only two features, leading to a limited variation in the constructed presence probabilities. It is possible to use many more features of the flights that need to be assigned to gates, leading to a more accurate prediction of their gate presence.

In this chapter, we obtain probabilistic delay predictions for flights arriving and de-

parting at a regional reference airport. To the best of our knowledge, this is the first time *probabilistic* predictions for flight delays on an individual flight basis are obtained. We employ two machine learning algorithms: Mixture Density Networks and Random Forest regression. We consider features based on flight schedules available at the reference airport, as well as the weather conditions recorded at the origin/destination airport of the flights. Suitable metrics are proposed to evaluate the performance of the considered machine learning algorithms, which estimate delay probability density functions (pdf). Furthermore, the impact of the choice of hyperparameters for these algorithms is analyzed.

The use of the obtained probabilistic predictions is demonstrated in the context of a robust flight-to-gate assignment problem. First, probabilistic predictions for arrival flight delays and departure flight delays are obtained using machine learning algorithms. These predictions are then used to estimate the probability of an aircraft being present at the reference airport. Lastly, these presence probabilities are integrated into a probabilistic FGAP model that aims to robustly assign arriving/departing aircraft to the gates of the reference airport. Here, robustness refers to the assignment model's ability to account for potential flight delays. The results show that, by considering flight delay predictions, flights are allocated to gates more robustly relative to the case when no information about flight delays is considered.

The remainder of this chapter is structured as follows: in Section 5.2, the datasets, machine learning algorithms for probabilistic flight delay predictions, and several performance metrics for these algorithms are introduced. The prediction results are then presented and discussed. In Section 5.3, the obtained probabilistic flight delay predictions are integrated into a flight-to-gate assignment model. Both a deterministic and a probabilistic model for the optimization of the FGAP are formulated. The models are both applied on a short and long term, and the results regarding the robustness of the obtained solutions are presented and discussed. In Section 5.4, conclusions and recommendations for future work are provided.

5.2. Data-Driven Probabilistic Flight Delay Predictions

In this section, we obtain probabilistic flight delay predictions using two machine learning algorithms, Mixture Density Networks and Random Forests Regression.

5.2.1. Data Description

Flight Schedule Dataset

For this analysis, flight schedules available at Rotterdam The Hague Airport (RTM) between 1 January 2017 and 29 February 2020 are considered. In total, 17,365 departing and 17,336 arriving flights are considered. These flights arrive from and depart to 42 airports across Europe and North Africa. The shortest route included is to London City Airport (LCY), and the longest to Tenerife South Airport (TFS), with an average of 1300 km. Figure 5.1 shows a map indicating all airports to or from which flights depart or arrive. The delay distribution of these flights is shown in Figure 5.2. The departing flights have an average absolute delay of 17.8 min with a standard deviation of 25.1 min, and the arriving flights have an average absolute delay of 15.4 min with a standard deviation

of 26.4 min. Here, the delay is considered to be the positive or negative time difference from the scheduled time of arrival/departure.

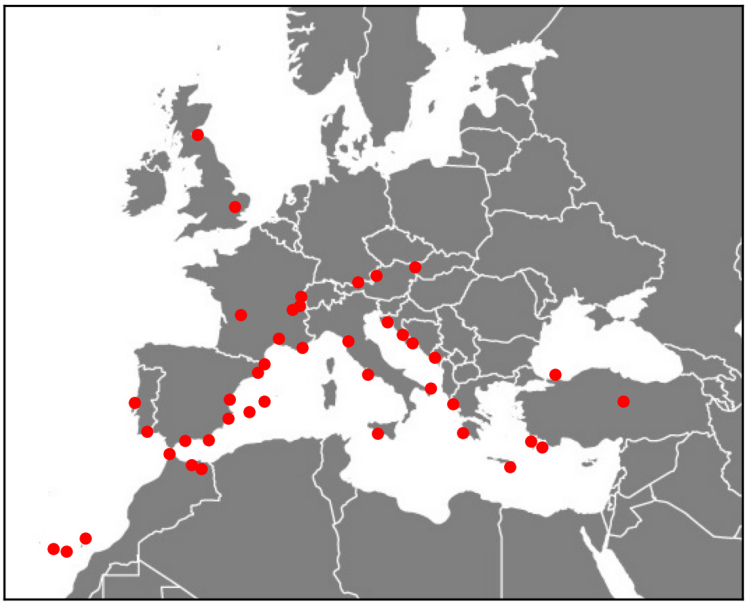


Figure 5.1: Map of origin/destination airports for Rotterdam Airport during the period January 2017–February 2020.

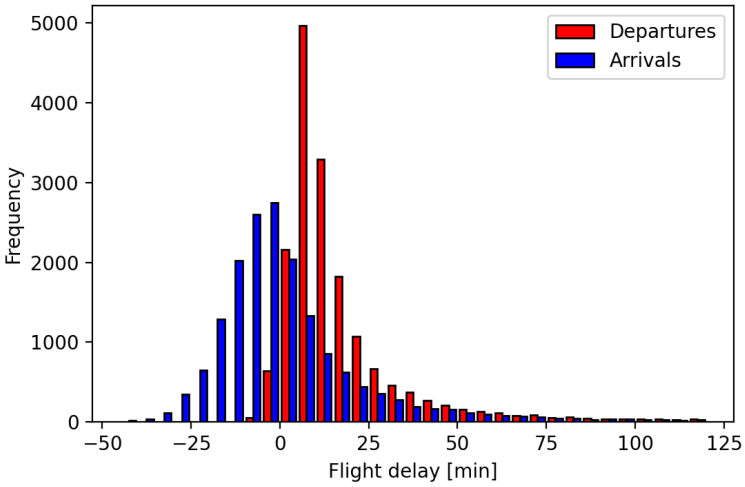


Figure 5.2: Histogram of the flight departure and arrival delays in the period January 2017–February 2020 at Rotterdam Airport.

WEATHER DATASET

We also consider the weather conditions, such as the temperature, pressure, and wind speed, measured at the origin/destination airport of all flights arriving/departing at RTM in the period 2017–2020. Measurements are available every 30 minutes [77].

5.2.2. FEATURE SELECTION

In this section, features are extracted and selected from the datasets described in Section 5.2.1. Feature selection is performed using the Pearson Correlation Coefficient. The correlation between any two features and the correlation between the features and the target (the flight delay) are calculated for a given training set. The features are selected as follows: for any two features that are correlated by more than the threshold value of 0.7, the feature that has the smallest correlation with the target variable is removed. Table 5.1 shows the features that have been selected for flight delay prediction. In Table 5.2, a description is provided for each of the selected features.

The features Airport, Airline, Season, Time of day, Day of week, Day of month, Day of year, Airport latitude and longitude, Distance, Month, Year and Scheduled flights 2h and Scheduled flights day are obtained or calculated from the flight schedule dataset. The feature Seats is derived from the aircraft type assigned to perform a flight. The features Temperature, Dewpoint, Visibility, Pressure, and Wind speed are obtained from the weather dataset.

Table 5.1: Feature encoding and selection for flight delay prediction.

Prediction	**Features**
Departure delay	Airport [a], Airline [a], Season [a], Time of day [b], Day of week [b], Day of month [b], Day of year [b], Airport latitude [c], Airport longitude [c], Day of month [c], Seats [c], Year [c], Scheduled flights 2 h [c], Scheduled flights day [c], Dewpoint [c], Visibility [c], Pressure [c], Wind speed [c]
Arrival delay	Airport [a], Airline [a], Aircraft type [a], Season [a], Time of day [b], Day of week [b], Day of month [b], Month [b], Airport longitude [c], Day of month [c], Distance [c], Seats [c], Year [c], Scheduled flights 2h [c], Scheduled flights day [c], Temperature [c], Visibility [c], Pressure [c], Wind speed [c]

[a] This feature is target encoded; [b] This feature is trigonometrically encoded; [c] This feature is numerically encoded.

Table 5.2: Description of features selected for flight delay prediction.

Feature	Description
Airport	the airport of destination (departures) or origin (arrivals)
Airline	the airline operating the flight
Aircraft type	the aircraft type used for the flight
Season	the flight season (summer or winter schedule)
Time of day	scheduled time of day of the flight
Day of week	scheduled day of the week of the flight
Day of month	scheduled day of the month of the flight
Day of year	scheduled day of the year of the flight
Month	scheduled month number of the flight
Airport latitude	the latitude of the destination/origin airport
Airport longitude	the longitude of the destination/origin airport
Distance	the distance between the origin and destination
Seats	the seat capacity of the used aircraft
Year	the year in which the flight was operated
Temperature	the air temperature at the destination/origin airport
Dewpoint	the dewpoint temperature at the destination/origin airport
Visibility	the prevailing visibility at the destination/origin airport
Pressure	pressure altimeter at the destination/origin airport
Wind speed	wind speed at the destination/origin airport
Scheduled flights day	the number of flights scheduled to depart/arrive during the day of the flight
Scheduled flights 2h	the number of flights scheduled to depart/arrive during the period between one hour before and one hour after the scheduled time of the flight

The features are either categorical, time-related, or numerical. The categorical features are target encoded based on a binary delay threshold of 15 min. The encoded value of the sample feature is the delay rate of the category to which the sample belongs. For example: if 8 out of 20 samples flying on Tuesdays are more than 15 min delayed, all Tuesday flights are encoded with value 0.4 for the feature Day of the week. The time features are encoded using trigonometric functions to preserve the periodicity. Two features (sine and cosine) are extracted from every time feature. For example, the features Month sine and cosine are calculated using $\sin(\frac{2\pi\chi}{12})$ and $\cos(\frac{2\pi\chi}{12})$ for a given month χ.

The remaining features are numerically encoded, i.e., the encoded value is the same as the original feature value. Note that the time features are both trigonometrically and numerically encoded. For example, the data field Day of the week yields the features Day of the week sine, Day of the week cosine, and Day of the week. The encoding method of every selected feature is denoted in Table 5.1. After encoding, all feature values are scaled to the interval $[0,1]$ to eliminate undesired feature domination in neural network models.

Table 5.1 shows that most features are selected for at least one of the departure/arrival pair, and that the trigonometrically encoded time features are selected more often than

the non-encoded time features.

5.2.3. MACHINE-LEARNING ALGORITHMS TO ESTIMATE THE PROBABILITY DISTRIBUTION OF FLIGHT DELAYS

Following feature selection, two algorithms are proposed to estimate the distribution of flight delays: Mixture Density Networks (MDN) and Random Forests regression (RFR). These algorithms belong to different classes of machine learning algorithms, neural networks, and decision trees, respectively.

MIXTURE DENSITY NETWORKS (MDNS)

A Mixture Density Network [22] is a combination of a neural network and a Gaussian mixture model. Given feature values $\mathbf{x}_i$ of flight i, an MDN outputs the parameters for each Gaussian in the mixture: the weight α, the mean μ, and the standard deviation σ. With these parameters, the probability density function $p(y_i|\mathbf{x}_i)$ of the target variable y_i, the flight delay, is determined. In general, the MDN is particularly suitable to estimate multimodal probability distributions [151, 183, 176, 27, 172, 50, 185]. It is therefore able to predict a distribution with peaks at, for example, two separate likely delay values.

The flight delay probability distribution is constructed as the weighted sum of Gaussian distributions as follows:

$$p(y_i|\mathbf{x}_i) = \sum_{j=1}^{m} \alpha_j(\mathbf{x}_i)\phi_j(y_i|\mathbf{x}_i), \tag{5.1}$$

$$\phi_j(y_i|\mathbf{x}_i) = \frac{1}{\sqrt{2\pi\sigma_j(\mathbf{x}_i)^2}} \exp\left(-\frac{(y_i-\mu_j(\mathbf{x}_i))^2}{2\sigma_j(\mathbf{x}_i)^2}\right) \tag{5.2}$$

where $p(y_i|\mathbf{x}_i)$ is the probability distribution of delay value y_i given feature values $\mathbf{x}_i$ from flight sample i, while $\alpha_j(\mathbf{x}_i)$, $\mu_j(\mathbf{x}_i)$ and $\sigma_j(\mathbf{x}_i)$ are the weight, mean, and standard deviation of the j^{th} Gaussian component, $1 \leq j \leq m$ with m the total number of Gaussian components considered for the mixture.

For any given flight, the features obtained in Section 5.2.2 are the input to the MDN, while the parameters α_j, μ_j, and σ_j are the output of the MDN. Thus, there are $3m$ outputs of the MDN. The weights use a softmax activation function, and the standard deviations use an exponential activation function, while the means are unrestricted.

The neural network is trained using backpropagation, i.e, the network parameters, the weights and biases of each node are updated using an error function E, which is the negative logarithm of the likelihood that the model derived from the output of the current network gives rise to the training data [22]. This likelihood is the product of the likelihood of every data point, given the current network parameters. Formally [22],

$$E = \sum_{i=1}^{N_f}\left(-\ln\sum_{j=1}^{m}\alpha_j(x_i)\phi_j(y_i|x_i)\right) = \sum_{i=1}^{N_f} -\ln p(y_i|x_i), \tag{5.3}$$

where N_f is the total number of samples in the training set, and we have used Equation (5.2).

For every data point fed to the neural network, the derivatives of the error with respect to all network parameters are used to update the weights and biases of the network. Following training, the MDN is applied to a test set and multimodal probability distributions for the delay of each flight in the test set are estimated. The MDN method is illustrated schematically in Figure 5.3.

RANDOM FORESTS REGRESSION AND KERNEL DENSITY ESTIMATION

Random Forests regression (RFR) is a class of decision tree-based machine learning algorithms [23]. The regular RFR algorithm is an ensemble method that combines the results of a number of decision trees. When building each tree, a random subset of the feature values of each training data point is used to make branches. The algorithm outputs a point estimate for the target variable (flight delay) of every test sample by averaging the output values of all considered decision trees. However, for our analysis, we are interested in estimating the probability distribution for the delay of the given flight, rather than a point estimate.

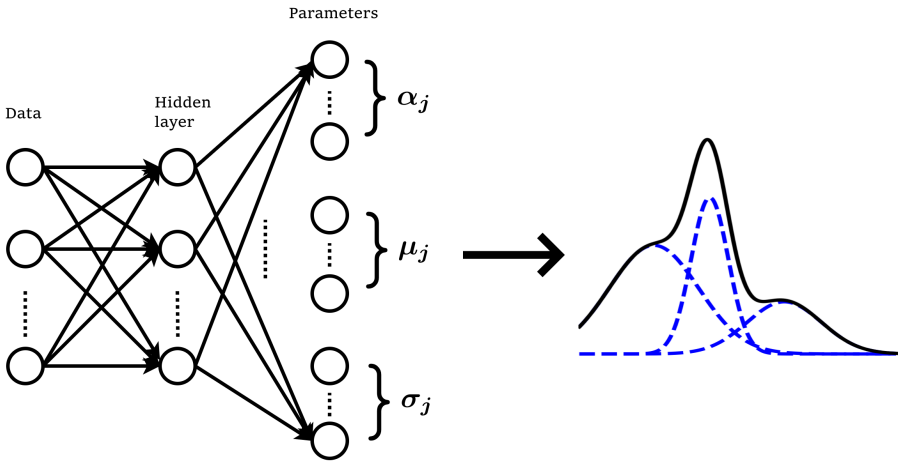


Figure 5.3: Schematic representation of a Mixture Density Network: parameters for a multimodal Gaussian distribution are obtained using a Neural Network.

In order to obtain the flight delay distribution of a flight in the test phase, the output values of the decision trees are not averaged, but collected, and a kernel density estimation (KDE) is performed [186]. A KDE results in a normalized probability density function. Two settings of the KDE are the kernel type and the bandwidth. In our analysis, a bandwidth of 1.5 is used to render the estimated distribution smooth. Gaussian kernels have been selected for their generality.

Random Forests regression is a well-established technique that has been applied in many research areas. However, there are very few examples of studies utilizing the algorithm to obtain probability distributions. Förster, Schultz, and Fricke [53] use quantile values, obtained from Quantile Random Forests, to construct a right-continuous cumulative distribution function of aircraft's time-to-fly from the turn onto the final approach

course to the runway threshold. Schlosser et al. [148] and Rahman et al. [130] use Random Forests algorithms to obtain probability distributions for precipitation forecasts and drug sensitivity, respectively. Both studies make use of feature probability distributions estimated via maximum likelihood to make splitting decisions when constructing the decision trees. Stochastic variables are introduced during or before the growing of the decision trees. In contrast, in this study, the feature values and splitting decisions are kept deterministic throughout the Random Forests algorithm. In this way, the probability density function is estimated from deterministic feature values without the need for stochastic variables. Furthermore, the working of the original Random Forests regression algorithm need not be changed.

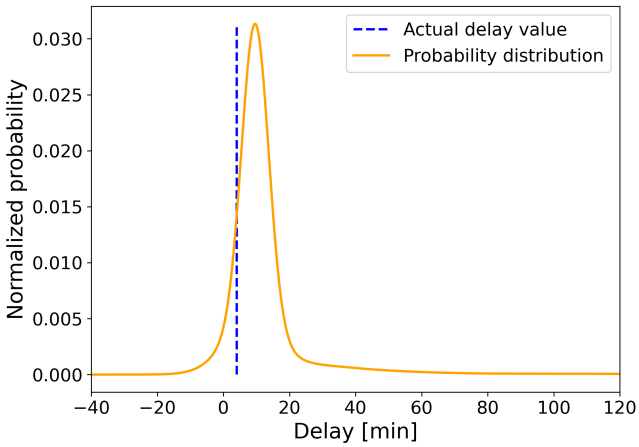
In Figure 5.4, an example of obtained probability distributions is shown for both methods. For both distributions, the actual delay value of the flight example is indicated.

5.2.4. HYPERPARAMETER TUNING

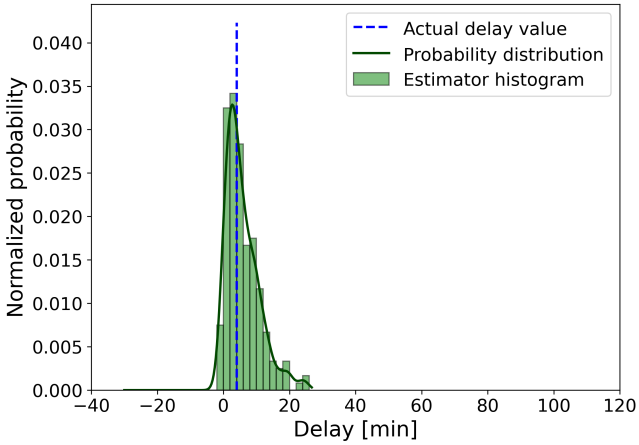
The hyperparameters of the MDN and the RFR prediction algorithms have been optimized using a grid search. The hyperparameters leading to the lowest mean CRPS scores (see Section) have been selected. Table 5.3 shows the selected hyperparameters and their search range. For MDN, a network with three hidden layers of 50 nodes is selected. The output layer of the network consists of 24 nodes, with which an 8-modal Gaussian distribution function is constructed. For RFR, 200 decision trees with a maximum depth of 10 layers are constructed. For every branch split, three out of four features are considered of at least seven training samples.

Table 5.3: Hyperparameters for MDN and RFR.

Mixture Density Network		
Hyperparameter	**Value**	**Range**
Number of modes m	8	[3, 5, 8, 10, 15]
Number of hidden layers	3	[1, 2, 3]
Number of nodes per hidden layer	50	[25, 50, 75, 100]
Number of epochs	1000	[500, 750, 1000, 1250, 1500]
Random Forest Regression		
Hyperparameter	**Value**	**Range**
Number of estimators	200	[100, 150, 200, 300]
Split criterion	Mean-squared error	[MSE, MAE]
Maximum tree depth	20	[4, 6, 8, 10, 12, 15, 20, 30]
Minimum samples per leaf node	7	[0, 3, 5, 7, 9]
Fraction of features considered for split	0.75	[0.25, 0.50, 0.75, 1.00]
KDE Bandwidth h	1.5	[0.5, 1, 1.5, 2]



(**a**) Probabilistic prediction on an example flight using MDN.



(**b**) Probabilistic prediction on an example flight using RFR.

Figure 5.4: An example of probabilistic prediction curves obtained from MDN and RFR for departure flight samples. Blue vertical lines indicate actual sample delay, orange curves depict the probability distribution obtained using MDN, green bars the histogram of RFR estimators, and green curves the probability distribution obtained from this histogram by KDE.

5.2.5. PERFORMANCE METRICS FOR PROBABILISTIC FORECASTING

As discussed before, many studies perform point estimate prediction on flight delays, such as [85, 106, 178, 164]. The most pervasive metrics for point estimate prediction are the root mean square error (RMSE) and mean absolute error (MAE), measured between the actual point and the predicted point. In this study, probabilistic forecasting is performed. Thus, metrics such as the RMSE and MAE cannot be applied, since they cannot be used to compare an entire delay distribution with a point value for actual delay. In

this chapter, the following six metrics are proposed to evaluate the performance of the MDN and RFR algorithms.

CONTINUOUS RANKED PROBABILITY SCORE

Since our aim is to estimate probability distributions for flight delays, a metric is needed that evaluates these distributions. The algorithms aim to obtain a distribution centered on the actual flight delay value, with a small standard deviation. To measure the extent to which the probabilistic prediction algorithms are able to achieve this, the Continuous Ranked Probability Score (CRPS) [108] is proposed. For an estimated flight delay probability distribution $p(y_i)$ and actual delay value $\bar{y}_i$, we define:

$$CRPS(F(y_i), \bar{y}_i) = \int_{-\infty}^{\infty} \left(F(z) - \mathbf{1}_{z \geq \bar{y}_i}\right)^2 dz, \tag{5.4}$$

where $F(y_i)$ is the cumulative distribution function of $p(y_i)$ and $\mathbf{1}$ is the Heaviside step function.

The CRPS is a generalization of the MAE for probabilistic predictions. It measures the deviation of the estimated delay cumulative distribution function from a step function at the actual delay value. This means that the CRPS attains the value 0 in the limit of a correct point prediction with absolute certainty. Since the CRPS is minimized if the model outputs the ideal distribution, the CRPS is a proper scoring rule. Therefore, it is an indication of both the sharpness and the calibration of the probabilistic forecast [57]. Figure 5.5 shows a case where the actual delay is 10 min, and includes examples of cumulative distributions with varying sharpness and calibration. Both a reduced sharpness and a reduced calibration in the distribution will increase the CRPS value. Since the CRPS is calculated for every flight in the test set, we introduce the metrics 'CRPS mean' and 'CRPS std', the mean and standard deviation of all CRPS values, respectively.

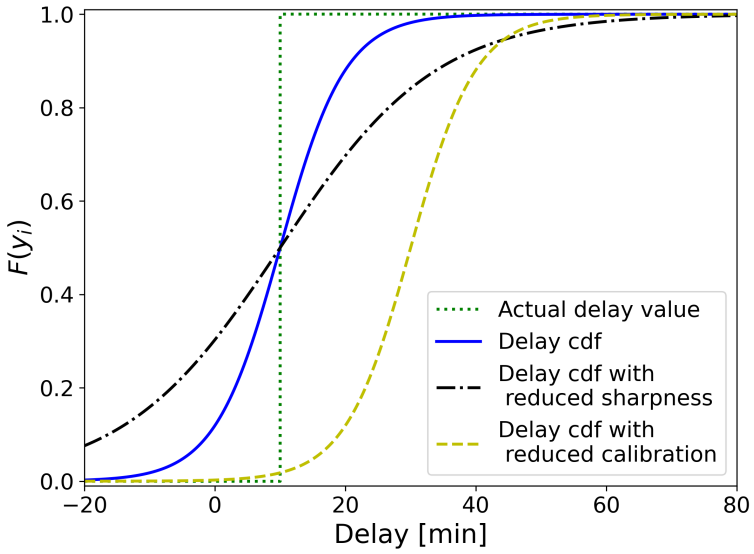


Figure 5.5: Illustration of the relation between the shape of the delay cumulative distribution function and the Continuous Ranked Probability Score (CRPS). The step function at the actual delay value (green dotted) corresponds with a CRPS value of 0. An example of a cdf with nonzero CRPS is plotted in blue. The black dash-dotted and yellow dashed lines show the same cdf with reduced sharpness and calibration, respectively.

RMSE$_M$ AND MAE$_M$

Since the RMSE and MAE are not suitable to assess an estimated flight delay distribution, we propose the variants RMSE$_M$ and MAE$_M$, which are calculated by comparing the mean value of the estimated distribution against the actual delay value. Before introducing the formal notation of these metrics, it is necessary to define the mean value of the estimated distribution. For MDN, the mean is defined as the weighted average of the component means, i.e., $\mu_{\text{MDN}}(\mathbf{x}_i) = \sum_{j=1}^{m} \alpha_j(\mathbf{x}_i)\mu_j(\mathbf{x}_i)$ is the distribution mean of flight sample i, with $\alpha_j(\mathbf{x}_i)$ and $\mu_j(\mathbf{x}_i)$ the weight and mean of component j. When using RFR, the mean $\mu_{\text{RFR}}(\mathbf{x}_i)$ is defined as the mean of the point estimates obtained from each decision tree. The distribution means are referred to as $\mu_{\text{M}}(\mathbf{x}_i)$ with $M \in \{\text{MDN, RFR}\}$.

The RMSE$_M$ and MAE$_M$ are then defined as:

$$\text{RMSE}_M = \sqrt{\frac{1}{N_f}\sum_{i=1}^{N_f}(\bar{y}_i - \mu_{\text{M}}(\mathbf{x}_i))^2}, \tag{5.5}$$

$$\text{MAE}_M = \frac{1}{N_f}\sum_{i=1}^{N_f}|\bar{y}_i - \mu_{\text{M}}(\mathbf{x}_i)| \tag{5.6}$$

The RMSE$_M$ and MAE$_M$ are used to characterize the average deviation of the mean of the estimated distribution from the actual delay $\bar{y}_i$ and thus measure only the calibration of the distribution and not the sharpness.

Metrics Based on the Standard Deviation

For MDN, the standard deviation of a multimodal probability density function for a flight sample $\mathbf{x}_i$ is calculated as follows [22]:

$$\sigma_{\text{MDN}}(\mathbf{x}_i) = \sqrt{\sum_{j=1}^{m} \alpha_j(\mathbf{x}_i)\left(\sigma_j(\mathbf{x}_i)^2 + \left(\mu_j(\mathbf{x}_i) - \mu_{\text{MDN}}(\mathbf{x}_i)\right)^2\right)}, \tag{5.7}$$

with $\alpha_j(\mathbf{x}_i)$ the weight, $\mu_j(\mathbf{x}_i)$ the mean and $\sigma_j(\mathbf{x}_i)$ the standard deviation of component j.

For the RFR algorithm, the standard deviation of the delay distribution is calculated in a similar fashion: a Kernel Density Estimation can be considered a multimodal Gaussian as well. This Gaussian has equal weights $\frac{1}{N_f}$, the RF regression point estimates as means and $\sqrt{h}$ as the standard deviation. This leads to the following expression for $\sigma_{\text{RFR}}(\mathbf{x}_i)$:

$$\sigma_{\text{RFR}}(\mathbf{x}_i) = \sqrt{\frac{1}{n_e}\sum_{j=1}^{n_e}\left(h + \left(\hat{y}_{i,j} - \mu_{\text{RFR}}(\mathbf{x}_i)\right)^2\right)}, \tag{5.8}$$

with n_e the number of estimators used in the algorithm, and $\hat{y}_{i,j}$ the jth point estimate for the delay of flight sample i. The distribution standard deviations are referred to as $\sigma_{\text{M}}(\mathbf{x}_i)$ with $M \in \{\text{MDN,RFR}\}$. Having obtained the distribution standard deviations in Equations (5.7) and (5.8), we can introduce the two metrics based on these. The first metric is the sample average of the standard deviation:

$$\bar{\sigma} = \frac{1}{N_f}\sum_{i=1}^{N_f}\sigma_{\text{M}}(\mathbf{x}_i), \tag{5.9}$$

where N_f is the number of flights in the test set. In order to define the second metric, we first introduce $f_{1\sigma}(\mathbf{x}_i)$, which indicates whether the actual delay $\bar{y}_i$ of a sample i lies within one standard deviation σ_{M} from the distributional mean μ_{M}. The second metric $\bar{f}_{1\sigma}$ is then defined as the average of this quantity over all N_f samples. It measures the ability of the probabilistic algorithm to predict a narrow delay distribution on or near the correct delay value. Together with the $\sigma_{\text{M}}(\mathbf{x}_i)$, it characterizes the spread of the estimated distribution and thus measures only the sharpness of the distribution and not the calibration. Formally,

$$\bar{f}_{1\sigma} = \frac{1}{N_f}\sum_{i=1}^{N_f} f_{1\sigma}(\mathbf{x}_i) \tag{5.10}$$

with

$$f_{1\sigma}(\mathbf{x}_i) = \begin{cases} 1 \text{ if } |\mu_{\text{M}}(\mathbf{x}_i) - \bar{y}_i| < \sigma_{\text{M}}(\mathbf{x}_i) \\ 0 \text{ if } |\mu_{\text{M}}(\mathbf{x}_i) - \bar{y}_i| \geq \sigma_{\text{M}}(\mathbf{x}_i) \end{cases}. \tag{5.11}$$

The six metrics defined in Equations (5.4)–(5.11) are used to assess the estimated flight delay distributions obtained using MDN and RFR. The metrics CRPS mean, CRPS std, RMSE_M, MAE_M and $\bar{\sigma}$ have the same unit as the target variable, i.e., minutes of delay, whereas $\bar{f}_{1\sigma}$ is expressed as a percentage.

5.2.6. RESULTS—PROBABILISTIC FLIGHT DELAY PREDICTIONS

We analyze both departing and arriving flights. For both, train and test sets are constructed using a 5-fold Cross Validation. The MDN and RFR algorithms have been used to estimate the distribution of the arrival and departure flight delays. The use of weather measurements implies a prediction horizon associated with these flight delay predictions of several days long. Table 5.4 shows the performance obtained using these algorithms.

Table 5.4: Performance metrics for probabilistic flight delay prediction.

Flights	**Algorithm**	**CRPS Mean** [min]	**CRPS Std** [min]	**MAE$_M$** [min]	**RMSE$_M$** [min]	$\bar{\sigma}$ [min]	$\bar{f}_{1\sigma}$ [%]
Departures	MDN	9.12	19.15	13.23	24.23	23.85	0.92
	RFR	8.86	18.15	12.51	23.32	12.08	0.69
Arrivals	MDN	10.95	17.59	15.62	24.98	24.60	0.87
	RFR	10.85	17.49	14.99	24.39	14.02	0.61

Table 5.4 shows that both MDN and RFR are able to predict departure and arrival delays within an average CRPS of 11 min. The RFR algorithm results in a smaller prediction error than the MDN algorithm. In addition, the delays of the arriving flights are predicted with larger error than those of the departing flights. This is explained by the fact that the bulk of the arriving flights has a considerably smaller delay than the bulk of the departing flights, as seen in Figure 5.2. Because the algorithms are trained mostly using arrival samples having a small delay, they have a decreased prediction performance for test samples with large delays. This decreased performance contributes greatly to the larger CRPS values.

Furthermore, Table 5.4 shows that the MDN algorithm predicts flight delays with a larger standard deviation than the RFR algorithm, and in turn the actual delay falls within this standard deviation more often. This is explained by the fact that the RFR algorithm produces a more narrow prediction curve than the MDN algorithm, on average.

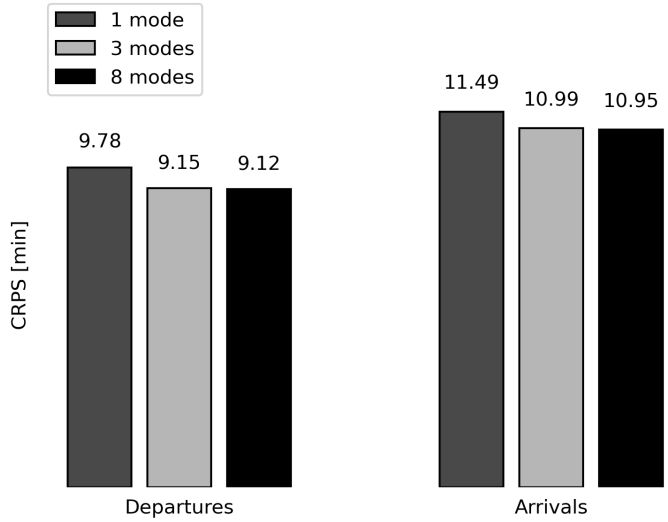
5.2.7. IMPACT OF THE CHOICE OF THE HYPERPARAMETERS

In this section, the influence of the values of important hyperparameters on the probabilistic flight delay prediction performance is assessed. The focus lies on the ability of the algorithms to construct a representative delay distribution; therefore, the mean CRPS is used to quantify the performance.

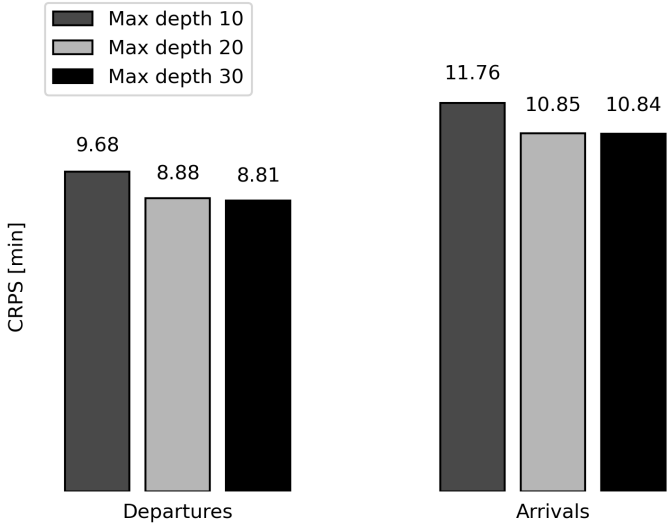
An important hyperparameter of the MDN algorithm is the number of modes. A distribution with more modes allows for more complex shapes, while a distribution with only one mode corresponds to a regular Gaussian distribution. In Figure 5.6a, the performance of the MDN algorithm for a varying number of modes is shown. Using multiple modes leads to a better performance than using a regular Gaussian function. When adding more than three modes, this improvement stagnates.

An important hyperparameter of the RFR algorithm is the maximum tree depth. A greater tree depth leads to a better distinction between different flights in the training set, but a tree depth that is too large can lead to overfitting. In that case, the er-

ror on the test set is not further reduced, while the computational time still increases. In Figure 5.6b, the performance of the RFR algorithm for varying values of the tree depth is shown. By analyzing a range of values between 10 and 30, it is found that a consistent performance is obtained from a max depth value of roughly 20.



(**a**) CRPS values for delay prediction using varying number of modes for MDN.



(**b**) CRPS values for delay prediction using varying values of the tree depth for RFR.

Figure 5.6: CRPS values obtained when varying hyperparameters in the MDN and RFR algorithms.

5.3. INTEGRATING PROBABILISTIC DELAY PREDICTIONS INTO THE FLIGHT-TO-GATE ASSIGNMENT PROBLEM

At an airport, a daily recurring operation is to assign arriving/departing flights to a gate. This is known as the flight-to-gate assignment problem (FGAP). The FGAP has been addressed extensively in literature [35]. An important quality of a given flight-to-gate assignment is its robustness. A greater robustness implies that, when faced with a disturbance (for example a flight is delayed), the model is able to handle this situation without introducing more disturbances. The aim of this section is to use the flight delay predictions obtained in Section 5.2 to obtain a robust flight-to-gate assignment. First, the FGAP is introduced, after which the flight delay predictions are integrated in this problem.

5.3.1. MATHEMATICAL FORMULATION OF THE DETERMINISTIC FGAP MODEL

In the past few decades, the FGAP has been modelled as a linear programming problem, having objectives such as the minimization of the number of towing procedures [14], the minimization of passenger walking distance [105], or obtaining robust flight-to-gate assignments by minimizing the number of gate conflicts [171, 179, 100, 194, 89].

These optimization models use a set of scheduled flights with deterministic flight arrival and departure times as input. Let us first introduce the following deterministic FGAP model [171].

Let N denote the set of n scheduled aircraft at the airport during a planning horizon, let G denote the set of g gates available at the airport, and let K denote the set of k time slots in the planning horizon. Let c_{ij} denote the cost of assigning aircraft $i \in N$ to gate $j \in G$. Let s_{it} denote the following binary presence indicator:

$$s_{it} = \begin{cases} 1, \text{ if aircraft } i \text{ is scheduled to be at the airport at time slot } t \in K \\ 0, \text{ otherwise} \end{cases}$$

The decision variables in this model are denoted as:

$$x_{ijt} = \begin{cases} 1, \text{ if aircraft } i \text{ is assigned to gate } j \text{ at time slot } t \\ 0, \text{ otherwise} \end{cases}$$

Then, the deterministic FGAP model is:

$$\min \sum_{i=1}^{n} \sum_{j=1}^{g} \sum_{t=1}^{k} c_{ij} x_{ijt} \tag{5.12}$$

$$\text{s.t.} \sum_{j=1}^{g} s_{it} x_{ijt} = s_{it} \quad \forall i \in N \text{ and } \forall t \in K. \tag{5.13}$$

$$\sum_{i=1}^{n} s_{it} x_{ijt} \leq 1 \quad \forall j \in G \text{ and } \forall t \in K. \tag{5.14}$$

$$s_{it} x_{ijt+1} - s_{it+1} \cdot x_{ijt} = 0 \quad \forall i \in N \text{ and } \forall j \in G \text{ and } \forall t \in K \setminus \{k\} \tag{5.15}$$

In this problem, the assignment costs over the total assignment are minimized (see Equation (5.12)) under the following conditions: Constraint 5.13 enforces that any aircraft i scheduled to be at the airport at time slot t, is assigned to exactly one gate j. Constraint 5.14 ensures that at most one aircraft i is assigned to a gate j at any time slot t. Lastly, Constraint 5.15 makes sure that an aircraft i cannot switch gates during its presence at the airport.

In this chapter, we consider a planning horizon of 24 h, separated in time slots. Every time slot consists of 5 min, therefore $k = 288$. The cost c_{ij} is assumed to be equal to 1 for any combination of gate and aircraft.

5.3.2. MATHEMATICAL FORMULATION OF THE PROBABILISTIC FGAP

In Section 5.3.1, the *deterministic FGAP model* was introduced. In this model, the aircraft presence is modelled by the binary, deterministic variable s_{it}. In this section, we introduce a *probabilistic FGAP model*, where the variable s_{it} is replaced by a presence probability function p_{it} of an aircraft, i.e., p_{it} is the probability that aircraft i is present at the airport at time slot t. In Van Schaijk and Visser [171], this aircraft presence probability is estimated based on a statistical analysis of a set of historical flights. In contrast, in this study, the presence probability is obtained using the machine learning prediction algorithms in Section 5.2, which provide an estimate of the delay for each individual flight.

Constraint (5.14), which refers to a deterministic aircraft presence in the deterministic FGAP model, is replaced by [171]:

$$\sum_{i=1}^{n} f(p_{it}, r)\, p_{it}\, x_{ijt} \leq 1 \quad \forall\, j \in G \text{ and } \forall\, t \in K, \tag{5.16}$$

where

$$f(p_{it}, r) = \frac{p_{it}}{r + p_{it}^2}, \tag{5.17}$$

with r a maximum overlap probability threshold between any two aircraft assigned to the same gate j at any time slot t. In other words, instead of Constraint 5.14, which ensures in the deterministic FGAP model that at most one aircraft is assigned to a gate in a time slot, Constraint 5.16 in the probabilistic FGAP model ensures that the probability that two aircraft are assigned to the same gate does not exceed a maximum threshold r at any time slot. Constraint 5.16 considers the overlap probability between two aircraft. In the case that the probability that three or more aircraft that are assigned to the same gate at the same time slot exceeds r, the probabilistic FGAP model is solved iteratively: for any instance where the value of r is exceeded, the number of aircraft assigned to the respective gate and time slot is iteratively decremented [100].

5.3.3. AIRCRAFT PRESENCE PROBABILITY FUNCTION

The aircraft presence probability function p_{it} is an input to the probabilistic FGAP model given in Equations (5.12), (5.13), (5.15) and (5.16).

Let y_i^{arr} and y_i^{dep} denote the arrival delay and departure delay of aircraft i. We determine the probability distributions of y_i^{arr} and y_i^{dep} using the machine learning al-

gorithms introduced in Section 5.2. These distributions are further used to obtain the aircraft presence probability p_{it} as follows.

Let STA_i and STD_i be the scheduled times of arrival and departure of aircraft i. Then, the predicted arrival and departure times are $X_i^{\text{arr}} = STA_i + y_i^{\text{arr}}$ and $X_i^{\text{dep}} = STD_i + y_i^{\text{dep}}$, respectively. Let $f_{X_i^{\text{arr}}}(t)$ and $f_{X_i^{\text{dep}}}(t)$ denote the pdf of X_i^{arr} and X_i^{dep}, respectively. Let $F_{X_i^{\text{arr}}}(t)$ and $F_{X_i^{\text{dep}}}(t)$ denote the cdf of X_i^{arr} and X_i^{dep}, respectively. Figure 5.7 shows the pdf of the arrival and departure times of an aircraft with STA = 12:20 and STD = 13:10. The cdf of these arrival and departure times is given in Figure 5.8.

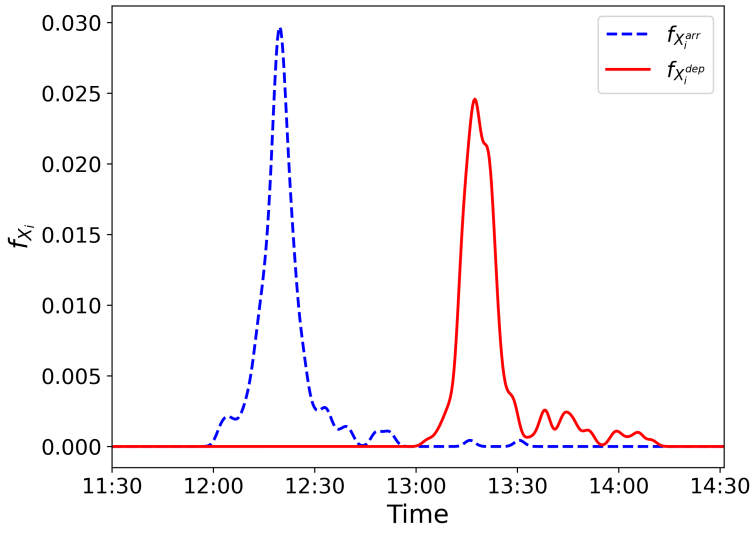


Figure 5.7: Pdf of the arrival and departure time of an aircraft with STA =12:20 and STD =13:10.

Using the cdf of X_i^{arr} and X_i^{dep}, we determine p_{it} as follows:

$$p_{it} = F_{X_i^{\text{arr}}}(t) \cdot (1 - F_{X_i^{\text{dep}}}(t)), \tag{5.18}$$

i.e., the aircraft presence probability p_{it} is calculated as the product of the probability that the aircraft has arrived and the probability that the aircraft has not yet departed, at time t. Figure 5.9 shows the aircraft presence probability p_{it} of an aircraft having STA = 12:20 and STD = 13:10, calculated using Equation (5.18).

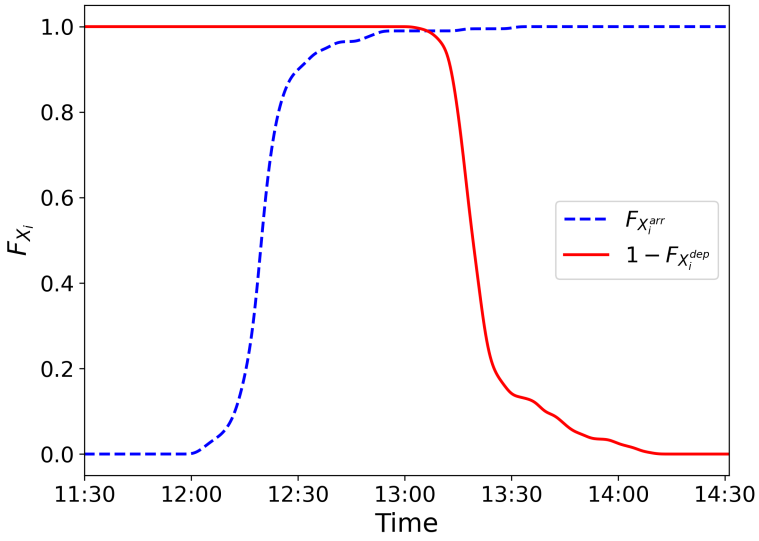


Figure 5.8: Cdf of the arrival and departure time of an aircraft with with STA =12:20 and STD =13:10

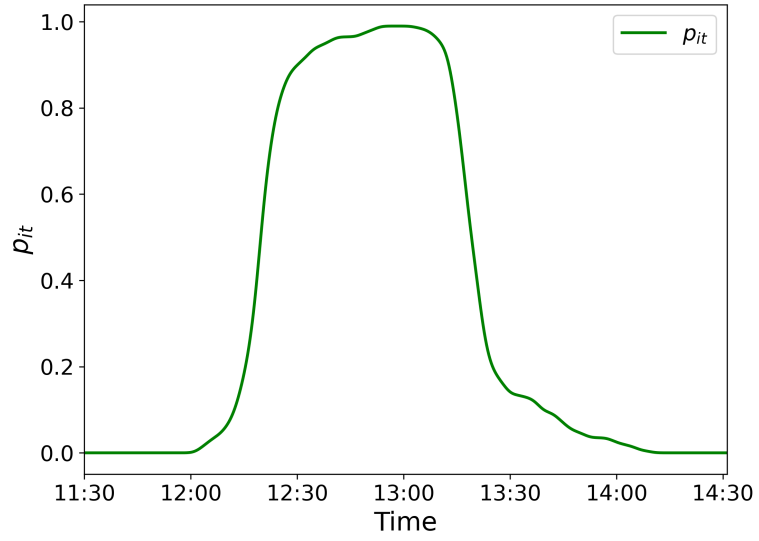


Figure 5.9: Aircraft presence probability (p_{it}) of an aircraft with STA =12:20 and STD =13:10.

If the aircraft has an overnight stay, i.e., does not arrive at and depart from the reference airport on the same day, then the aircraft presence probability p_{it} is calculated as follows:

Case 1: The aircraft has stayed at the airport during the night before the day of interest, and departs at the beginning of the day. The aircraft presence probability is formed

using only the cdf of departure time:

$$p_{it} = 1 - F_{X_i^{\text{dep}}}(t), \tag{5.19}$$

Case 2: The aircraft arrives at the airport in the evening and stays during the night after the day of interest. The aircraft presence probability is formed using only the cdf of arrival time:

$$p_{it} = F_{X_i^{\text{arr}}}(t), \tag{5.20}$$

Having obtained the aircraft presence probability p_{it} for an aircraft i and p_{jt} for an aircraft j, we determine the overlap probability between aircraft i and j at timestep t as $p_{it} \cdot p_{jt}$. Figure 5.10 shows an example of an overlap probability for aircraft 1 and 2.

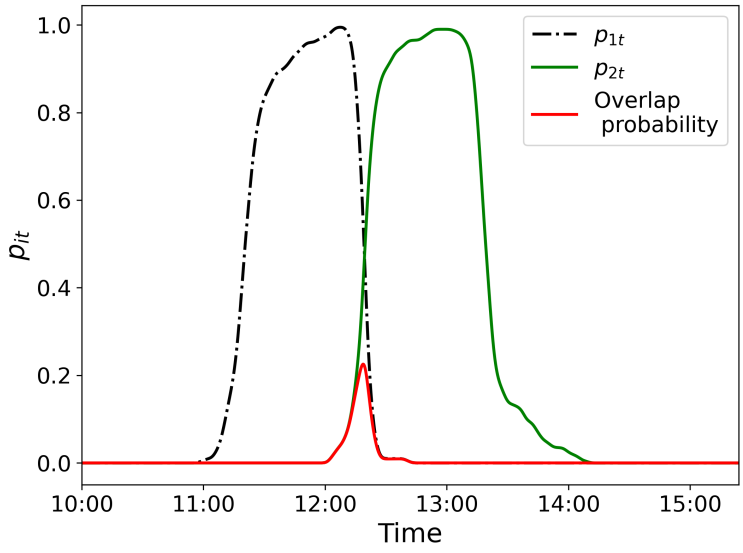
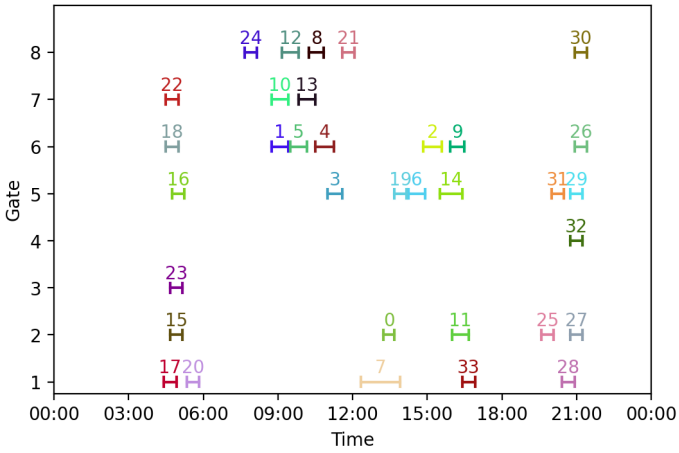


Figure 5.10: Two aircraft presence probability functions with STA_1 = 11:20, STD_1 = 12:00, STA_2 = 12:20, STD_2 = 13:10, and their overlap probability.

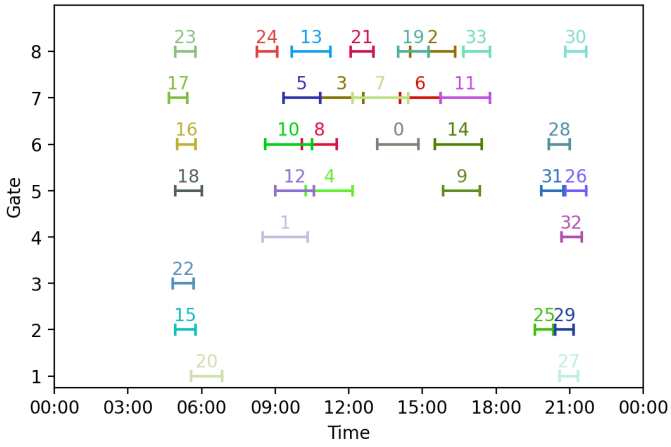
5.3.4. RESULTS—FLIGHT-TO-GATE-ASSIGNMENT INTEGRATING PROBABILISTIC FLIGHT DELAY PREDICTIONS

In this section, the results obtained from the deterministic and probabilistic FGAP models introduced in Sections 5.3.1 and 5.3.2 are outlined. The flight-to-gate assignment is generated at RTM airport for one day of operations: 14 July 2019. On this day, a total of 25 departures and 24 arrivals were scheduled. This day is referred to as the date of interest. The collection of all flights scheduled on the date of interest forms the test set for the delay predictions obtained using the machine learning algorithms in Section 5.2. All flights scheduled in the period from 1 January 2017–13 July 2019 form the training set of the machine learning algorithms. For our FGAP model, we assume eight gates, to which aircraft can be assigned.

Figure 5.11 shows the assignment of aircraft to gates for the date of interest obtained using the deterministic and probabilistic FGAP models. For the assignment obtained with the deterministic model, the presence as indicated by a solid line shows the period of time an aircraft occupies a gate, based on its scheduled arrival and departure time, i.e., based on the aircraft presence s_{it}. For the assignment obtained with the probabilistic FGAP model, the presence indicated by a solid line shows the period of time for which the aircraft presence probability p_{it} is larger than 0.1.



(**a**) FGA obtained with the deterministic model.



(**b**) FGA obtained with the probabilistic model using $r = 0.15$.

Figure 5.11: Flight-to-gate assignments for 14 July 2019 obtained using the deterministic and probabilistic models.

Figure 5.11a shows that, in the flight-to-gate assignment obtained using the deterministic FGAP model, aircraft 1 and 5 are assigned to the same gate (6). However, Fig-

ure 5.11b shows that, in the assignment obtained using the probabilistic FGAP model, aircraft 1 and 5 are assigned to different gates (4 and 7). This is because the overlap probability between aircraft 1 and 5 exceeds the maximum overlap probability threshold r. The same situation occurs for aircraft 10 and 13.

The deterministic and probabilistic flight-to-gate assignments obtained are evaluated using the actual aircraft presence of the aircraft that flew on the date of interest. The actual aircraft presence is the time between the actual arrival and ATA_i and actual departure time and ATD_i.

Figure 5.12 shows the aircraft presence probability p_{it} for all the aircraft assigned to gates 7 and 8 in the solution obtained using the probabilistic FGAP model, as shown in Figure 5.11b, versus the actual times the aircraft were present at gates 7 and 8. The overlap probability between any two aircraft is plotted in red. We consider a maximum permissible overlap probability $r = 0.15$.

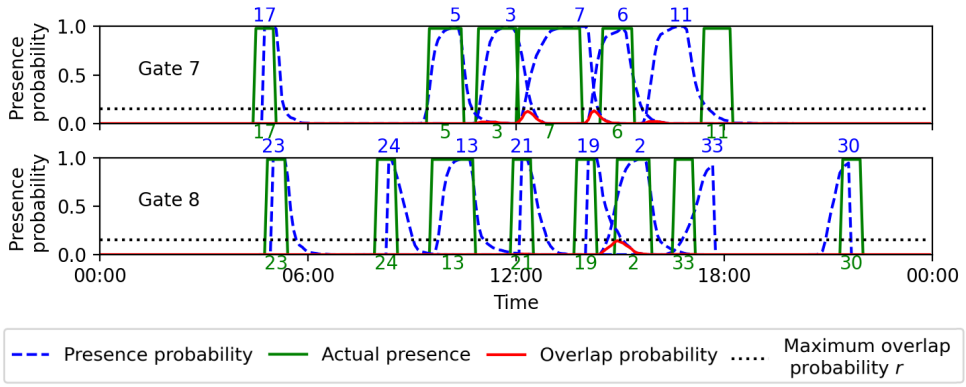


Figure 5.12: Flight-to-gate assignment for gates 7 and 8 on 14 July 2019 obtained using the probabilistic model with maximum overlap probability $r = 0.15$ combined with actual aircraft presence.

Following Constraint 5.16 in the probabilistic FGAP model, Figure 5.12 shows that the overlap probability of any two aircraft assigned to either gate 7 or 8 does not exceed the threshold r. The overlap probabilities between aircraft 19 and 2 (at gate 8) and aircraft 7 and 6 (at gate 7) are near the maximum overlap probability threshold r. The actual presence periods of these pairs of aircraft do not overlap, showing that a threshold of $r = 0.15$ was sufficient to prevent aircraft conflicts for these aircraft pairs. The actual presence periods of aircraft 3 and 7 (at gate 7) do overlap, leading to an aircraft conflict. In this case, the conflict is caused by the fact that the predicted presence of aircraft 7 is later than the actual presence.

5.3.5. RESULTS—LONG RUN PERFORMANCE

In order to evaluate the long run performance of the deterministic and probabilistic FGAP model, they are applied to test data comprising a period of 30 days: from 14 July 2019 up to and including 12 August 2019. Two metrics are used for evaluation:

- An aircraft is defined as a *conflicted aircraft* if there is at least one time slot at which

this aircraft and any other aircraft are both present at the same gate.

- For the probabilistic FGAP model, a gate time slot is defined as a *used gate time slot* if there is an aircraft present at the gate at this time with a probability of more than 0.5—for the deterministic FGAP model, if there is an aircraft present at the gate at this time. Note that the maximum amount of used gate time slots is equal to $g \cdot k$.

The number of conflicted aircraft (CA) is a metric that measures the robustness of the FGA against delay when in operation. The number of used gate time slots (UGT) is a metric that measures to which extent an increase in robustness induces the need for a larger utilization of the available gate capacity.

To evaluate the flight-to-gate assignments, the means and standard deviations of these metrics over all testing days are used. The probabilistic FGAP model has been run with a range of possible conflict probabilities r, namely $r \in [0.05, 0.10, 0.15]$. Since the RFR algorithm has proven to yield the most accurate results in Section 5.2, RFR is used to obtain the presence probabilities. Table 5.5 summarizes the metric values obtained when evaluating the flight-to-gate assignments obtained from the deterministic and probabilistic FGAP model.

Table 5.5: Results for the FGA's at RTM airport, averaged over the days from 14 July until 12 August 2019 (30 days). The mean and standard deviation of the number of conflicted aircraft (CA) and the number of Used Gate Time slots (UGT) are shown for all methods. For reference, the total number of aircraft per day and the total number of available gate time slots are added. The presence probabilities were constructed using RFR.

	CA Mean	**CA σ**	**UGT Mean**	**UGT σ**
Total	**31.6**	**6.7**	**2304**	**N/A**
Deterministic FGAP	5.03	2.87	254	57.9
Probabilistic FGAP, r = 0.15	2.57	2.30	319	76.7
Probabilistic FGAP, r = 0.10	1.73	1.84	319	76.5
Probabilistic FGAP, r = 0.05	1.33	1.49	319	76.6

When considering the probabilistic FGAP model, Table 5.5 shows that the average number of conflicted aircraft is smaller for all values of the maximum overlap probability threshold r, when compared to the deterministic FGAP model. The probabilistic FGAP model results in a more robust assignment than the deterministic FGAP model. The gate usage increases by 25%, while the number of conflicted aircraft is reduced by up to 74%. The number of conflicted aircraft does not decrease further when decreasing r further than 0.05. The maximum overlap probability threshold can thus be used by airport operators to adjust the robustness of the flight-to-gate assignment to the desired level.

5.4. CONCLUSION

In Section 5.2, two probabilistic forecasting algorithms, Mixture Density Networks and Random Forest regression, have been applied to the problem of flight delay prediction. The algorithms were trained using features extracted from a flight schedule dataset and

a weather dataset, which contained data from Rotterdam The Hague Airport. Six performance metrics were defined to evaluate the probabilistic predictions, and the influence of the hyperparameters on the probabilistic prediction performance was investigated.

The results show that it is possible to estimate probability distributions for future flight delays within a Continuous Ranked Probability Score of 11 min, several days in advance. The probabilistic flight delay predictions can provide airport coordinators not only with an estimate for the flight delays of all incoming flights, but also with a measure of the certainty of these estimates. In this way, better informed decisions regarding strategic flight schedules can be made, and on-time performance prediction can be improved.

Subsequently, in Section 5.3, the probabilistic predictions were used as input to a probabilistic linear programming model optimizing the flight-to-gate assignment problem, with the goal of increasing the robustness of this assignment. The results for the flight-to-gate assignment problem show a reduction of up to 74% in the average number of conflicted aircraft per day by incorporating the probabilistic flight delay predictions. The robustness can be adjusted by varying the maximum permissible overlap probability threshold in the probabilistic optimization model. The application of flight delay predictions to the flight-to-gate assignment problem provides a framework for increasing robustness for flight-to-gate operations at airports.

Future work includes the application of the introduced approach to increasing the robustness of flight-to-gate assignments to a larger airport, taking into account e.g., varying assignment costs, airline gate usage and the nearness of changed gates to the original gate, and, second, the integration of probabilistic flight delay predictions into models for other airport operations. Examples are arrival/departure sequencing and scheduling, and electric taxiing operational planning.

6

Considering Imbalanced Datasets and Airport Planners' Preferences when Predicting Flight Delays and Cancellations

In Chapter 5 probabilistic algorithms were used to predict flight delays. In existing literature, a popular approach is to use binary classification for flight delays, flight cancellations and other classification problems. However, such problems often suffer from a large data imbalance, which can reduce the performance of classification algorithms. In this chapter a range of imbalance ratios, classification algorithms and sampling techniques are considered to achieve the best performance for classification problems with imbalanced data. A delay and cancellation case study show the specifications required to achieve the best performance in the desired metric. In general, the results underline the need to investigate the influence of varying data imbalance ratios on the performance of classification algorithms.

This chapter is based on the following research article:
Hendrickx, R., Zoutendijk, M., Mitici, M. & Schaefer, J. (2021). "Considering Airport Planners' Preferences and Imbalanced Datasets when Predicting Flight Delays and Cancellations." in *IEEE/AIAA 40th Digital Avionics Systems Conference (DASC)* [66].

6.1. INTRODUCTION

Flight on-time performance is an important measure for airport and airline service quality. Before the COVID-19 crisis, the continuous growth of air traffic led to challenging scheduling situations and an increase in flight delays and cancellations: In 2018, more than 11 million flights were operated in Europe, with an average delay of 14.7 minutes, an increase of 3.8% and 17% from 2017, respectively [45, 46]. After the crisis, the air traffic volume is expected to restore to its pre-crisis level within 5 years [43]. An increase in the number of flight delays and cancellations has detrimental effects on an airline's and airport's quality of service and revenue [8]. As such, having the ability to anticipate which flights may be cancelled or delayed is of great value for airports and airlines, as it allows for pro-active decision making to mitigate the effects of cancellations/delays. In order to anticipate flight delays and cancellations it is necessary to predict these events ahead of time, preferably with a high certainty, in order to allow efficient managing of the airports resources.

One class of techniques that can be used to predict flight delays and cancellations is that of machine learning classification techniques. In the past years, several studies have developed machine learning algorithms to predict flight delays and cancellations [161], emphasizing the importance of flight on-time performance. One of the challenges of classification problems is the fact that the used datasets can have an imbalanced class distribution, i.e., the amount of samples in the class of interest is only a fraction of the amount of samples in the majority class. This imbalance leads to a low performance of the classification algorithms [188], which usually work best when having a balanced class distribution. Binary flight cancellation and delay prediction is one example of a classification problem where the issue of imbalanced class distribution needs to be addressed. When considering regular operations, a large majority of the flights are not delayed or cancelled, causing the problem to be imbalanced.

In order to address the limitations caused by data imbalance, many studies use oversampling and under-sampling techniques such as Synthetic Minority Oversampling Technique (SMOTE) [30], and Random Undersampling (RUS) [94]. However, when using these techniques, a 50% – 50% sampling ratio is most often used, which is not necessarily the ratio that leads to the best performance of the prediction algorithms with respect to the performance metrics considered. Moreover, the performance metric of choice is usually accuracy, while this may not be the most relevant performance metric for the problem considered.

In this chapter a systematic approach is proposed to analyse and deal with the effects of highly imbalanced datasets when predicting flight delays and cancellations. First, the most relevant performance metric for the prediction problem is selected. Then an adaptive sampling methodology is used to determine which sampling technique and which imbalance ratio yield the best classification performance with regard to this metric. This approach is demonstrated using several sampling techniques and classification algorithms, which are applied to data on flights arriving/departing to and from a large, European hub-airport, in the period 2015 - 2019. In addition to flight operational data, weather data from METAR weather reports [77] are considered. To the best of our knowledge, this paper is the first to propose a systematic approach to deal with the inherent imbalance of the prediction of flight on-time performance, when formulated as a binary

classification problem.

This chapter contributes to the current body of knowledge concerning highly imbalanced datasets and flight on-time performance as follows. From a practical point of view, this proposed approach provides support for air transport stakeholders such as airport coordinators who can use the predictions and the proposed approach to assess flight schedules in advance of the flight execution and take action in order to mitigate the effects of flight delays and cancellations. Second, both flight delays and cancellations are addressed, while existing studies mainly focus on flight delay predictions and not flight cancellations. Predictions for flight cancellations in particular make use of highly imbalanced datasets, which are the focus of this chapter. Third, the approach presented in this chapter can be used to deal with imbalanced data in other fields of research, when considering binary classification problems.

The remainder of this chapter is structured as follows. Section 6.2 presents the systematic approach to deal with the inherent data imbalance, including the binary classification algorithms, feature selection and relevant performance metrics, and addresses the classification results. Section 6.3 concludes the research by discussing the approach, summarizing the most important observations and providing suggestions for future research directions.

RELATED WORK

In recent years, many studies have addressed the flight delay prediction problem using machine learning techniques. Usually, the authors express the problem as a classification task: in Choi et al. [32], the authors predict airline delay on prediction horizons of 5 days, 1 day and 0 days, using Decision Trees, Random Forests, AdaBoost and k-Nearest-Neighbors classifiers. The data is sampled using a combination of SMOTE [30] and RUS [94]. In general, the Random Forest classifier is found to have the best performance, with an accuracy of 0.80. In Horiguchi et al. [69] flight delays are predicted on prediction horizons of 5 months, 1 week and 1 day using Random Forests, XGBoost and Deep Neural Networks. These algorithms make use of airline data, originating from a low cost carrier. The classifiers attain an average Average Under Curve (AUC) score of 0.65 for a horizon of 1 day, with a maximum of 0.75 for certain airports. In Lambelho et al. [95] flight delay and cancellations predictions are used to rank IATA strategic flight schedules at London Heathrow Airport. The predictions are made using three different classification algorithms, of which LightGBM performs best, attaining a maximum F1-score of 0.60 for the cancellation prediction problem. In Kim et al. [90] deep learning algorithms are used to predict flight delays for airports in the US, several hours before the operation. Weather data is also considered in this study. It is found that the Recurrent Neural Networks architecture results in the most reliable delay prediction: an accuracy of 0.87 is obtained. In Chen and Li [31] an air traffic delay prediction model is proposed that combines multi-class Random Forests and an approximated delay propagation model, which results in an accuracy of 0.87. Additionally, it is found that departure delay and late arriving aircraft delay are the most important features for the prediction. The authors use SMOTE to resample the dataset. Finally, Alonso and Loureiro [9] perform multi-class predictions for departing flight delay at Porto Airport, several hours before the flight.

Other studies express the flight delay prediction problem as a regression task. The

authors of Manna et al. [106] investigate the prediction of flight delays several months before the operation for US airports. Using Gradient Boosted Decision Trees, the authors find that the model predicts flight delay patterns with a root mean square error (RMSE) of 8.2 and 10.7 minutes for departure and arrival delay, respectively. Next, Kalliguddi and Leboulluec [85] estimate flight delay several hours ahead of operation using several algorithms, of which Random Forests performs best, with an RMSE of 12.5 minutes. It is concluded that late aircraft delay, carrier delay, weather delay and national airspace delay have the largest effect on on-time performance. Lastly, Thiagarajan et al. [164] perform both classification and regression on the flight delay prediction problem. Classification using the Gradient Boosting Classifier with a combination of SMOTE and Tomek Links [165] yields an accuracy of 0.94 and a recall of 0.91. Regression with Random Forests produced an RMSE of 8.7 minutes.

The topic of flight cancellation has been approached in varying ways in the literature: both Cao and Kanafani [26] and Jarrah et al. [83] are studies utilising on-time performance data to propose an accurate decision-support tool, integrating flight delays and cancellations. They apply network models with minimum cost and maximum profit objectives, respectively. The tool returns an optimal set of flights to either delay or cancel. Furthermore, Seelhorst and Hansen [153] investigate flight cancellation behaviour by using an econometric discrete choice model. The purpose of the research is to identify factors that influence flight cancellations and to predict cancellation probabilities. The results are incorporated in a queuing model, which visualises the effects flight cancellations have on flight delays. Lastly, Alderighi and Gaggero [8] analyze the effect of an airline being part of a global alliance on cancellations. It is concluded that airlines belonging to an alliance are likely to have more flight cancellations compared to non-alliance airlines. Complementary to these studies, in this chapter the cancellation problem is posed as a binary classification problem.

On-time performance datasets are generally imbalanced, and so are flight delays and cancellation datasets. Regarding imbalance, multiple studies have been carried out on different topics. First, Zhao, Wong, and Tsui [188] establish an approach to handle imbalanced healthcare data by incorporating multiple different rebalancing techniques. The proposed framework successfully improves the detection of rare healthcare events due to look-alike sound-alike mix-ups. A 45% increase in recall is observed when combining a logistic regression algorithm with SMOTE. Another study on the effects of data imbalance is Hassanzadeh et al. [64]. Four different rebalancing strategies are presented, combined with a binary classification framework for scientific artifacts in the evidence-based medicine domain. An increase of up to a factor of three in the F1-score of the minority class was found for some of the strategies. Within the field of aircraft on-time performance the most popular approach is to reduce imbalance by sampling with over- or undersampling techniques, such as random oversampling [29], random undersampling [32, 19, 164], SMOTE [18, 32, 60, 164, 31] and Tomek Links [164]. Most studies choose to resample (i.e. bring to an equal amount of samples) the delayed and undelayed classes, without using a systematic approach to choose the sampling ratio.

This chapter aims to elaborate on previous work regarding handling of imbalanced datasets and the prediction of flight delay and cancellation using machine learning, by developing a general approach to handle imbalance in on-time performance datasets.

6.2. DEALING WITH IMBALANCE: A SYSTEMATIC APPROACH

In this section, a systematic approach is presented to select an optimal imbalance ratio for an imbalanced dataset in the context of binary classification for flight cancellation and delay. The approach is demonstrated by predicting cancellation and delays with two different classification algorithms and two different sampling techniques, on a one-day prediction horizon.

6.2.1. DATA DESCRIPTION AND DEFINITIONS

In this study, Amsterdam Airport Schiphol (AAS) is considered as the reference airport where flights are scheduled to depart from/arrive at. Two datasets are considered for the proposed prediction algorithms: i) cancelled arrival/departure flights and, ii) delayed arrival/departure flights.

i) Cancelled flights - Highly imbalanced dataset

A total of 1,956,418 arriving and departing flights to and from AAS in the period 2015-2018 are considered. The dataset is based on the strategic flight schedules available in 2015-2018 and contains information such as scheduled date and time of the flight arrival/departure, origin/destination airport of the scheduled flight and the airline that operates the flight. These flights are operated by 256 airlines that fly to/from 649 airports. Furthermore, 54% of the flights have both the destination and origin airport in the Schengen area. Out of all considered flights 1.6% (30,695) are cancelled. Therefore this dataset is considered to be highly imbalanced.

An arriving/departing flight is considered to be *cancelled* if this flight is scheduled to arrive/depart at the reference airport, but it is not operated on the day of the scheduled arrival/departure.

ii) Delayed flights - Moderately imbalanced dataset

The flight delay dataset contains a total of 479,400 arriving and departing flights to and from AAS during 2019. Similar to the cancelled flights dataset, this dataset is based on the strategic flight schedules available in 2019 and contains information such as date and time of arriving/departing flights, origin/destination airport and the airlines that operate the flights. Specifically, the flights are operated by 99 different airlines, flying from 336 unique origin airports and to 323 unique destination airports. This delay dataset is considered to be moderately imbalanced with 34% (82,350) of all departing flights being delayed, and 24% (57,253) of all arriving flights being delayed.

An arriving/departing flight is considered to be *delayed* if during operation, this flight arrives/departs 16 min or more after the scheduled time of arrival/departure.

With regard to imbalance in datasets, the following definitions are introduced. The *imbalance ratio* of a dataset of flights is defined as the ratio of delayed (cancelled) flights to non-delayed (non-cancelled) flights. The *base imbalance ratio* of a flight dataset is defined as the imbalance ratio the considered dataset initially has. Lastly, the *sampling ratio* applied to a flight dataset is defined as the ratio between the amount of delayed (cancelled) flight samples after sampling and the amount of delayed (cancelled) flight samples before sampling.

As an example, a dataset of 100 flights, of which 20 are delayed, has an imbalance

ratio of 20/80, i.e. 0.25. If the minority class is oversampled to a size of 40, the imbalance ratio increases to 40/100, i.e. 0.40. An imbalance ratio of 100% corresponds with perfect resampling, where the number of delayed (cancelled) and non-delayed (non-cancelled) flights are equal.

Figure 6.1 shows the delay distribution of the arriving/departing flights in 2019 at and from AAS. These histograms show that both the distributions of the arrival and departure flight delays are unimodal with positive skew, i.e., the flights are more likely to arrive/depart later than scheduled compared to earlier than scheduled. Also, as expected, the histograms show that the arriving flights generally experience less delay than the departing flights.

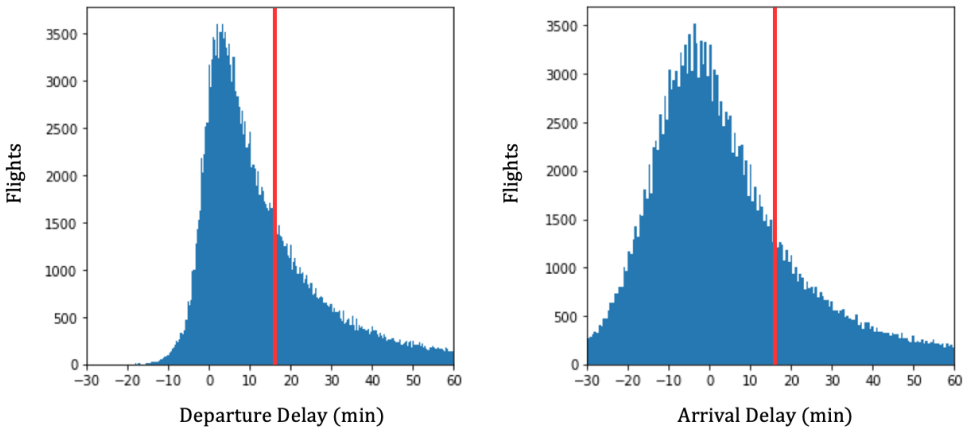


Figure 6.1: Departure and arrival delay distribution of flights arriving and departing at/from AAS in 2019. The vertical red line shows the delay threshold of 16 min.

Apart from the flight schedule specific datasets, the weather conditions at the origin/destination airports such as the air temperature, wind speed, visibility and pressure at sea level are considered. These data are obtained from METAR [77].

6.2.2. A SYSTEMATIC APPROACH TO DEAL WITH IMBALANCED DATA FOR FLIGHT DELAY AND CANCELLATION PREDICTIONS

Given the fact that the flight cancellation and delay datasets are highly and moderately imbalanced, respectively, a systematic approach is proposed to deal with these imbalances when predicting flight delays and cancellations. Figure 6.2 shows a schematic overview of the proposed approach. First, the relevant performance metrics for flight delay and cancellation prediction algorithms are identified. Next, an adaptive sampling procedure is iteratively applied to the flight delay and cancellation prediction algorithms. Finally, an optimal imbalance ratio is determined. The available data is sampled such that this imbalance ratio is attained and several binary classification algorithms are run to predict whether flights are delayed or cancelled.

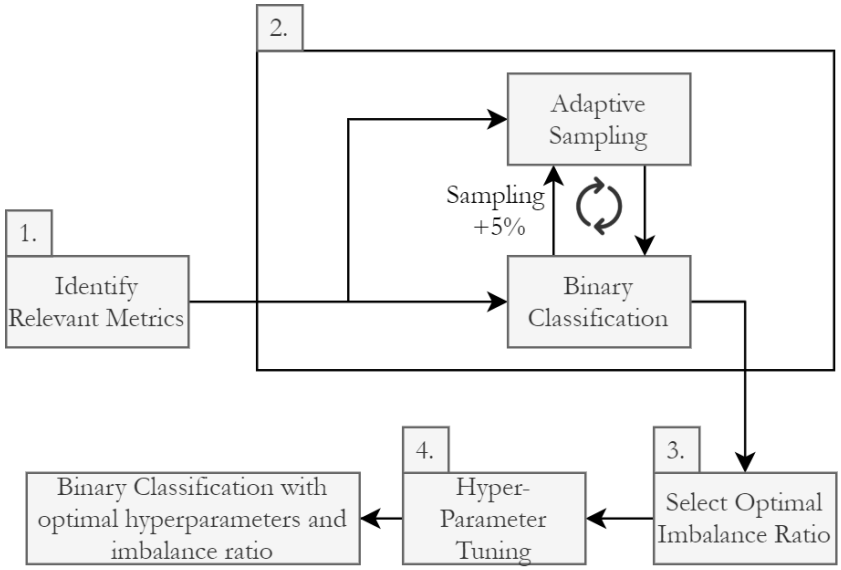


Figure 6.2: A flow diagram of the systematic approach to deal with imbalanced data.

STEP 1: IDENTIFYING RELEVANT PERFORMANCE METRICS

First, performance metrics relevant for the prediction problem are identified. Common metrics for binary classification algorithms are accuracy, precision, recall and F1-score. However, given that the datasets are highly and moderately imbalanced, accuracy is not considered as a relevant performance metric.

Given the specific problem of flight delay/cancellation prediction, in practice it is preferred by airport planners to be able to predict whether flights are delayed/cancelled with a high certainty, even at the cost of mis-classifying many delayed/cancelled flights as not delayed/not cancelled. Otherwise, a low certainty in the flight delay/cancellation prediction may lead to less-informed decisions from an airport planner, which may negatively affect stakeholders such as airlines, passengers, etc. As such, in this study, precision is considered to be the main performance metric (high certainty of predictions), and F1-score as the second most important metric (overall performance of the prediction algorithm).

STEP 2: PREDICTION ALGORITHMS AND ADAPTIVE SAMPLING

In this step, several binary classification algorithms are employed to predict flight delays and cancellations. Below the feature selection and an adaptive sampling approach for these classification algorithms are discussed.

FEATURE ENCODING AND SELECTION

Table 6.1 indicates whether each feature is categorical, numerical or time-related. The categorical features are target-encoded. Here, the target-encoded value of a categorical

feature is the probability of the flight being delayed/cancelled, based on all samples that fall into the same category [109]. For example, if 20 out of all 50 flights from an airline X are delayed, then airline X is encoded with value 0.4. The time features such as hour, day of week and month are encoded using trigonometric functions that preserve periodicity [69]. Lastly, all feature values are scaled to the interval $[0,1]$ to eliminate feature domination or ranking [69, 32].

Table 6.1 also shows which features have been selected for predicting the departure delay, arrival delay and cancellations using binary classification algorithms. The selection is performed based on Pearson's correlation coefficients. These features are the flight number, the airline operating the flight, the apron handler assigned to a flight at the airport, the aircraft type used for the flight, the aircraft registration number, the airport and country of origin/destination, the number of times an origin-destination airport route is operated per day by all aircraft arriving/departing at/from AAS, the service type of the flight (passenger or freight), the month of the year, the time of day, and, for both the destination and origin airport: the wind speed, gust speed, air temperature, air pressure, visibility and snow presence. Table 6.1 shows that the delay classifiers make more use of time features, since busy periods in the flight schedules are causes for flight delay. The cancellation classifiers, however, make more use of weather features such as visibility and snow presence, as they often cause flight cancellations.

Table 6.1: Selected features for the delay and cancellation prediction problems.

Classifier	Features
Departure delay	Flight number[c], Airline[c], Handler[c], Aircraft type[c], Aircraft registration[c], Destination airport[c], Route frequency[n], Month[t], Time[t], Gust speed (origin)[n], Temperature (origin)[n], Temperature (destination)[n]
Arrival delay	Flight number[c], Handler[c], Aircraft type[c], Aircraft registration[c], Origin airport[c], Month[t], Time[t], Gust speed (destination)[n]
Cancellations	Flight number[c], Airline[c], Handler[c], Aircraft registration[c], Origin/destination airport[c], Origin/destination country[c], Service type[c], Wind speed[n], Pressure[n], Visibility[n], Snow[n]

[c] Categorical feature, target encoding
[n] Numerical feature
[t] Time feature, trigonometric encoding

BINARY CLASSIFICATION ALGORITHMS

The flights are classified as delayed or cancelled using two binary classification algorithms: Random Forests (RF) and Multilayer Perceptron (MLP). Random Forests [23] is a collection of many classification trees which are each constructed using a different subset of the training set, and using a different selection of features. Each tree carries out a class vote, after which the RF classifies using the majority vote. This approach reduces overfitting and sensitivity to outliers, and enhances the predictive accuracy. The Multilayer Perceptron [68] is a feed-forward neural network with backpropagation, non-linear activation functions and hidden layers. The MLP has the advantage that it can learn non-linear relations. Both the MLP and RF algorithms are well-established and often used in

the field of machine learning classification and are therefore fitting to be used in the demonstration of our adaptive sampling approach.

For both algorithms the datasets are split into train and test data, with an 80%-20% ratio. Thus, a 5-fold Cross Validation is used for these classifiers.

ADAPTIVE SAMPLING

In this part of the procedure, adaptive sampling is used to investigate the relation between the imbalance ratio of the dataset used for the prediction problem at hand (flight cancellation or delay prediction), and the performance metrics considered relevant for the problem (see Step 1). Adaptive sampling is performed as follows: starting at the base imbalance ratio, the imbalance ratio is iteratively increased by 5%, until it reaches 100%. For each such imbalance ratio, the classification is performed using the two classification algorithms introduced previously and the sampling is performed using two sampling techniques. The resulting values of the performance metrics selected in Step 1, i.e., precision and F1-score, are thus obtained for each imbalance ratio. Lastly, for every combination of algorithm and sampling technique, an optimal imbalance ratio is selected such that precision and F1-score are highest.

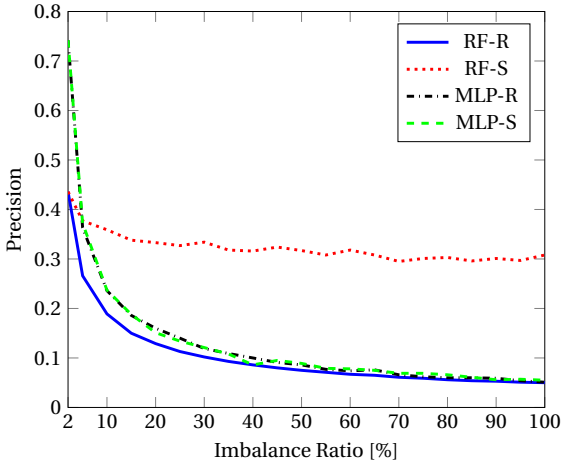
The two sampling techniques, used to sample the considered dataset for every imbalance ratio, remain to be introduced. The first is an oversampling technique and the second is an undersampling technique: Synthetic Minority Oversampling Technique (SMOTE) [30] over-samples the minority class, i.e. the cancelled/delayed flights, by creating synthetic samples between samples and their nearest neighbours. When using SMOTE, the samples are not duplicated. Random Undersampling (RUS) [94] undersamples the majority class by leaving out random samples from this class. Both techniques are well-known in literature, and the approach presented in this chapter can be extended to different sampling techniques. In summary, for every value of the imbalance ratio, the classification is performed with four different settings: RF sampled with SMOTE, RF sampled with RUS, MLP sampled with SMOTE and MLP sampled with RUS.

Figures 6.3 to 6.5 show the precision, recall and F1-score as functions of the imbalance ratio for the cancellations, departure delays and arrival delays, respectively, obtained using the RF and MLP algorithms and the features as described in Section 6.2.1. The sampling techniques SMOTE and RUS are indicated by S and R, respectively. The methods are run with the default hyper-parameter settings as hyper-parameter tuning is performed at a later stage.

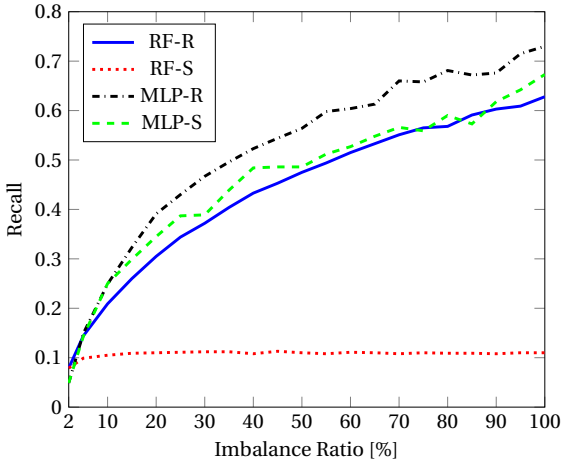
i) Cancellations

Figure 6.3 shows that the precision score is highest at the base imbalance ratio, 1.6%, for all combinations of algorithms and sampling techniques. The precision rapidly decreases with increasing imbalance ratio, until it levels at 0.05. The opposite can be seen for the recall, which starts at a minimum and increases with increasing imbalance ratio. There is a clearly visible trade-off between recall and precision. Finally, the peak of the F1-score is observed near a ratio of 10%. Since the F1-score constitutes the harmonic mean between precision and recall, the peak is observed at an imbalance ratio where neither of the precision and recall attain extreme values. The results also show that RF with SMOTE is insensitive to the imbalance ratios for all metrics.

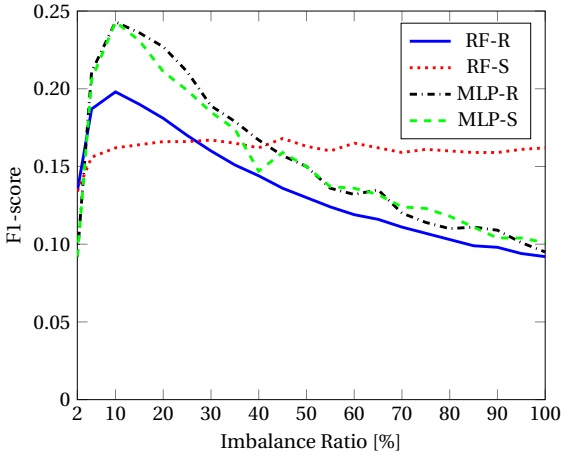
ii) Departure delays



(a)

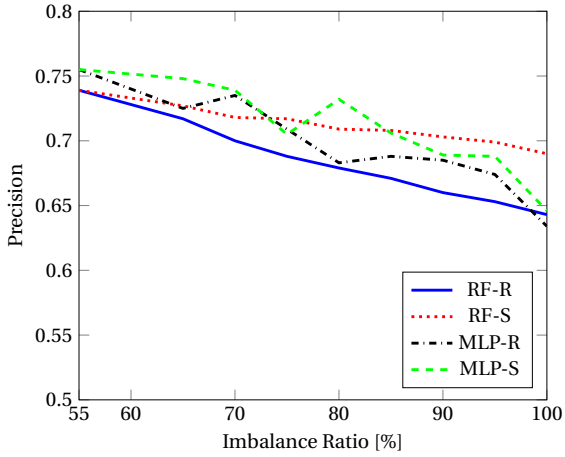


(b)

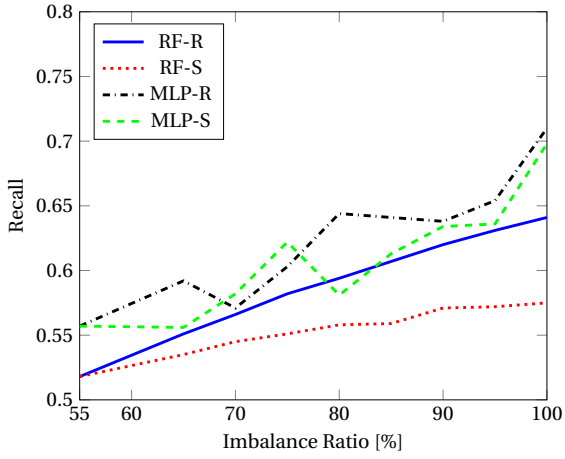


(c)

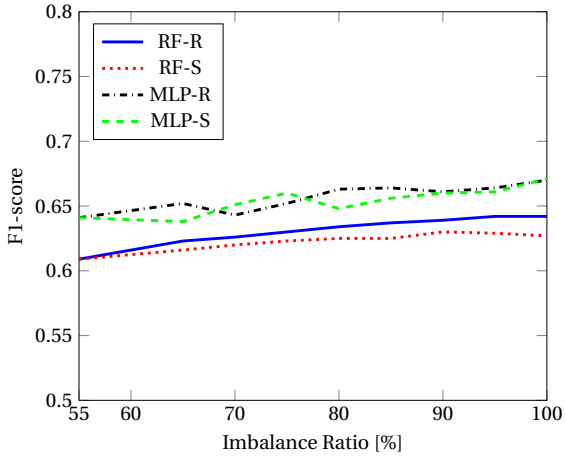
Figure 6.3: Precision (a), recall (b) and F1-score (c) as function of the imbalance ratio, for cancellation prediction (RF = Random Forest, MLP = Multilayer Perceptron, R = RUS, S = SMOTE.)



(a)

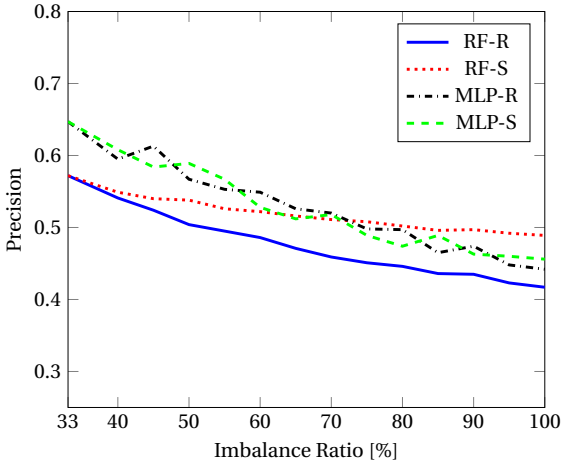


(b)

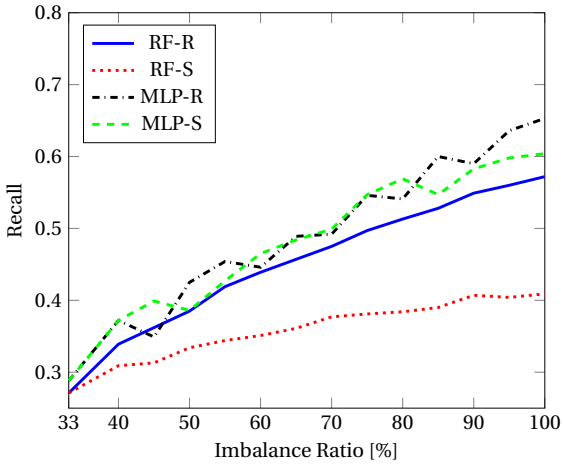


(c)

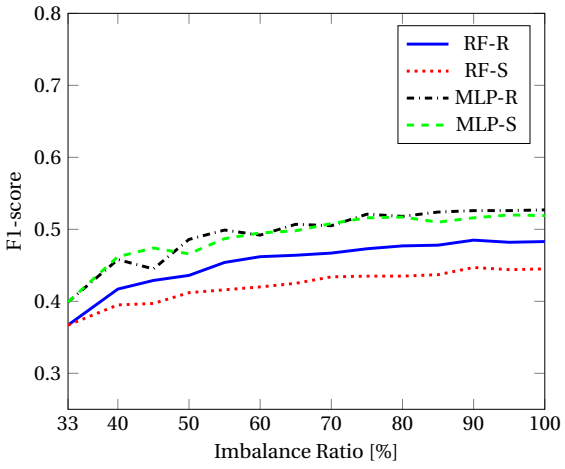
Figure 6.4: Precision (a), recall (b) and F1-score (c) as function of imbalance ratio, for departure delay prediction (RF = Random Forest, MLP = Multilayer Perceptron, R = RUS, S = SMOTE).



(a)



(b)



(c)

Figure 6.5: Precision (a), recall (b) and F1-score (c) as function of the imbalance ratio, for arrival delay prediction (RF = Random Forest, MLP = Multilayer Perceptron, R = RUS, S = SMOTE).

For the departure delays, the imbalance ratio ranges between 55%, the base imbalance ratio, and 100%. The graphs for precision, recall and F1-score are shown in Figure 6.4. The general trends are the same as for the performance of the cancellation classifiers, but the performance differences are smaller. Precision decreases with increasing imbalance ratio, while recall increases with increasing imbalance ratio, for both algorithms and sampling techniques. The F1-score also gradually increases with the imbalance ratio.

iii) Arrival delays

Finally, for the arrival delays, the precision, recall and F1-score graphs are shown in Figure 6.5. The base imbalance ratio for arrival delay lies at 33%. Again, there is a clear decreasing trend for precision and an increasing trend for recall, with the F1-score graph corresponding to their harmonic mean.

STEP 3: SELECTING AN OPTIMAL IMBALANCE RATIO

In this step an optimal imbalance ratio is selected based on the performance achieved in Step 2. As mentioned above, an optimal imbalance ratio is the ratio for which the relevant performance metric value (see Step 1) is highest.

Figures 6.3a, 6.4a and 6.5a show that the highest precision is attained at the base imbalance ratio, i.e. without using sampling, for both classification algorithms. This shows that at the base imbalance ratio the algorithms only classify those samples as positive that have high certainty of being positive. This leads to a small amount of false positives, and consequently to a higher precision than for greater imbalance ratios. As expected, the large amount of positive samples that cannot be classified as such with high certainty by the algorithm lead to a large amount of false negatives, and consequently to a lower recall.

For the F1-score, the highest performance is obtained as follows. For the cancellation results, an optimal F1-score for MLP is obtained when using a 10% imbalance ratio sampled with SMOTE (see Figure 6.3c). An optimal F1-score for RF is located at the 10% RUS imbalance ratio. Considering departure delay results, the MLP achieves the best performance at 100% SMOTE and the RF at 100% RUS, as shown in Figure 6.4c. Finally, for the arrival delay results, the highest F1-score is obtained at an imbalance ratio of 100% RUS for MLP and 90% RUS for RF (see Figure 6.5c). Due to the greater imbalance in the cancellation dataset a larger range of imbalance ratios is considered for the cancellation prediction during the adaptive sampling procedure. This leads to a larger range of precision and recall values for cancellations (Figures 6.3a and 6.3b), as opposed to the values for flight delay (Figures 6.4a, 6.4b, 6.5a and 6.5b). This explains why a clear optimum imbalance ratio appears for the cancellation F1-score near 10% (Figure 6.3c), while for the delay F1-score the values are similar for all considered imbalance ratios, and the optimum is less pronounced compared to that of the cancellation prediction (Figures 6.4c and 6.5c).

A summary of these optimal selected imbalance ratios for each classifier is shown in Table 6.2.

Table 6.2: Optimal imbalance ratios corresponding to the maxima in the performance metric plots, for all classification problems and both the Multilayer Perceptron (MLP) and Random Forest (RF) classifiers.

	Maximize	Cancellations	Departure Delay	Arrival Delay
MLP	Precision	no sampling	no sampling	no sampling
	F1-score	10% SMOTE	100% SMOTE	100% RUS
RF	Precision	no sampling	no sampling	no sampling
	F1-score	10% RUS	100% RUS	90% RUS

Table 6.3: Final hyperparameters for Multilayer Perceptron (MLP).

	Maximize	Sampling	Hidden layer size	Batch size	Acti-vation	Solver	Learning rate
Cancel	Precision	no sampling	100 (1 layer)	1000	ReLu	sgd	constant
-lations	F1-score	10% SMOTE	100 (1 layer)	1000	ReLu	adam	constant
Departure	Precision	no sampling	100 (1 layer)	auto	ReLu	adam	constant
Delay	F1-score	100% SMOTE	100 (1 layer)	auto	ReLu	adam	constant
Arrival	Precision	no sampling	100 (1 layer)	1000	logistic	sgd	adaptive
Delay	F1-score	100% RUS	100 (1 layer)	auto	ReLu	adam	constant

Table 6.4: Final hyperparameters for Random Forest (RF).

	Maximize	Sampling	Number of trees	Criterion	Max depth	Max features
Cancellations	Precision	no sampling	100	Entropy	10	0.2
	F1-score	10% RUS	300	Entropy	6	1.0
Departure	Precision	no sampling	500	Gini	8	0.1
Delay	F1-score	100% RUS	500	Entropy	6	1.0
Arrival	Precision	no sampling	100	Gini	6	0.1
Delay	F1-score	90% RUS	300	Entropy	6	0.7

STEP 4: PERFORMING HYPER-PARAMETER TUNING

Following the selection of an optimal imbalance ratio, hyperparameter tuning is performed for the flight cancellation, departure flight delay and arrival flight delay classifiers. For the RF classifier, the number of trees, selection criterion, maximum tree depth and maximum features per tree are considered for tuning. For the MLP classifier, the hidden layer size, the batch size, activation function, solver and the learning rate are considered. In all cases, a random grid search is performed. Table 6.3 and Table 6.4 show the best hyperparameters for the considered classifiers.

6.2.3. RESULTS - BINARY CLASSIFICATION FOR FLIGHT DELAYS AND CANCELLATIONS WITH OPTIMAL IMBALANCE RATIOS AND HYPER-PARAMETER TUNING

Table 6.5: Final performance metric results for cancellation, departure delay, and arrival delay prediction.

		Cancellations		Departure delays		Arrival delays	
	Indicator	MLP	RF	MLP	RF	MLP	RF
Highest precision	Accuracy	0.986	0.986	0.682	0.681	0.768	0.765
	Precision	**0.809**	**0.853**	**0.614**	**0.660**	**0.692**	**0.713**
	Recall	0.041	0.035	0.303	0.203	0.054	0.028
	F1-score	0.079	0.068	0.406	0.311	0.101	0.054
	AUC	0.772	0.850	0.691	0.691	0.680	0.693
Highest F1-score	Accuracy	0.978	0.981	0.666	0.645	0.710	0.640
	Precision	0.263	0.284	0.524	0.493	0.406	0.362
	Recall	0.237	0.198	0.491	0.601	0.528	0.624
	F1-score	**0.249**	**0.233**	**0.507**	**0.542**	**0.459**	**0.458**
	AUC	0.854	0.839	0.679	0.685	0.712	0.700

Using the obtained optimal imbalance ratios and sampling techniques for each prediction problem and selected metric of interest, the classification algorithms are applied once more to perform the final flight delay and cancellation predictions. The results are summarized in Table 6.5. All results are the mean of a 5-Fold Cross Validation. In this table, "highest precision" and "highest F1-score" indicate that the imbalance ratios have been used that produce optimal results for the respective metric (see Table 6.2). For example, the highest F1-score of 0.507 for departure delays with MLP is obtained using 100% SMOTE.

Table 6.5 can be used to compare the performance of the two used classification algorithms, RF and MLP. For the cancellation problem, the table shows that the precision performance of RF is higher than that of MLP when optimizing for precision (no sampling). The opposite is observed for the value of the F1-score when optimizing for F1-score (10% sampling). For the departure delay problem RF outperforms MLP for both metrics of interest. For the arrival delay problem the difference between the classifier performances is smaller and in the case of F1-score the performance is similar, although the MLP does attain a greater accuracy.

Table 6.5 shows that the general performance, as illustrated by the F1-score, is better

when the base imbalance ratio is larger. When aiming for a high precision, the results show that the departure delay results have the smallest difference between recall and precision, followed by the arrival delay and cancellation results. The trade-off between precision and recall is therefore stronger for smaller base imbalance ratios, as expected.

As shown in Step 3, sampling does not improve the precision in any of the cases. However, for F1-score a clear improvement is observed when choosing an optimal imbalance ratio. For example, when using the MLP classifier, the increase is 243% for cancellation predictions, 74% for the departure delays, and 354% for the arrival delays, compared to the base imbalance ratio.

In general, the fact that large differences in the classification performance are observed when comparing the precision, recall and F1-score between the different imbalance ratios, confirms the need for a systematic approach to deal with imbalanced datasets regarding the flight cancellation and delay classification problem.

6.3. CONCLUSION

In this chapter, a systematic approach to deal with highly imbalanced data for binary classification problems is developed, in order to enhance the performance of machine learning algorithms predicting flight delays and cancellations, while taking into account the preferences of airport planners regarding this performance. The presented approach emphasises the need to identify the performance metrics relevant for the considered problem. In the case of predicting flight delays and cancellations, correct predictions are valuable to airport coordinators. The predictions can be used to propose changes to strategic flight schedules. However, the airlines, which are subject to these change proposals, are expected to accept such change proposals only if the predictions have a high certainty. Hence, in this chapter the performance metric considered to be most relevant has been the precision, as a high precision implies a high certainty in predictions. Additionally, the F1-score has been considered.

The algorithms Random Forests and Multilayer Perceptron are trained and tested with flight operational data from a large European hub airport and weather data. The imbalance of the data is mitigated by applying an adaptive sampling procedure to the prediction problem using the sampling techniques Random Undersampling (RUS) and Synthetic Minority Oversampling Technique (SMOTE), and investigating its effects on the classifier performance.

The imbalance analysis and its results show that optimal performance with respect to the metrics can be obtained by varying the data imbalance ratios. Optimal precision is shown to be found at base imbalance ratio (data without sampling), for all algorithm and sampling technique combinations. In order to find the optimal F1-score, sampling is shown to be necessary. Increasing the imbalance ratio to the optimal amount improves the F1-score by a significant factor for each prediction problem. In the case of cancellation prediction, the optimal imbalance ratio greatly differs from the ratio corresponding to the conventional resampling (100%).

The proposed approach provides support for major hub-airports to perform on-time performance prediction. Furthermore, the approach can be applied within other research areas when considering imbalanced classification problems. Moreover, the presented approach is not dependent on the type of machine learning algorithm, the fea-

tures considered, nor on the type of data. Therefore, it is generic and can be applied to any imbalanced binary classification problem.

Future work includes development of a systematic approach to deal with imbalanced datasets on which multiclass classification or regression is performed, which use different performance metrics than are used for binary classification. Lastly, our approach can be applied to an on-time performance analysis of regional airports.

7

CONCLUSION

In this chapter, the conclusions from the research findings are gathered and the research objectives are reviewed. The limitations of the works are outlined, and recommendations for future work are provided. Last, the scientific and societal contributions are established.

7.1. REVIEWING RESEARCH OBJECTIVES

This section states how the research objectives defined in Section 1.3 are addressed in the dissertation.

Objective 1
Identify the current research status on efficient large-scale application of electric taxiing at airports.

Chapter 2 reviews current literature on electric taxi systems and their implementation at airports. Since reviews on technical feasibility and environmental benefits have already been performed, the chapter focuses on reviewing operational management aspects of electric taxiing implementation at airports, and identifying challenges for future research. One of the difficulties in anticipating the future electric taxiing infrastructure is the current lack of consensus in literature on the values of several key operational management related parameters, such as taxiing speed, battery capacity, charging and discharging rate of the Electric Towing Vehicles (ETVs). This was remedied by extracting information from similar research in other fields, as well as sources from industry. Working to solving the challenges identified in this chapter will contribute to the task of reducing airport ground emissions.

Objective 2
Develop an optimisation model to perform ETV-to-aircraft scheduling that takes into account realistic airport circumstances.

In Chapter 3 two ETV-to-aircraft scheduling models were introduced, a Mixed Integer Linear Programming (MILP) model and an Adaptive Local Neighbourhood Search (ALNS) model. Both models include realistic airport circumstances such as conflict and collision avoidance, aircraft and ETV routing, a mix of certified and non-certified aircraft, and ETV charging properties. Such model requirements and information were gathered from literature and described in Chapter 2. Both models include constraints to reflect the limit electricity capacity expected to be available at airport in the situation of widespread electric taxiing implementation. The models are useful for long-term strategy considerations at airports: whether to implement electric taxiing at the airport, and if so, with how many charging points and ETVs. Or, when such a system is already in place, to consider the costs and benefits of increasing the fleet size, charging capacity or amount of charging locations. For example, for the parameters and airport considered in Chapter 3, it was found that night charging becomes obsolete for ETV battery capacities over 750 kWh, and that increasing the ETV battery capacity has limited use if the ETV charging rate is not increased accordingly.

Objective 3
Develop an optimisation approach to retain efficiency for electric taxiing when confronted with delays.

Chapter 4 introduces two MILP models for ETV fleet scheduling. The first model aims to generate fleet schedules on a strategic level, and the second model is used to create schedules which take into account imminent flight delays. This model allows for reassignment of the ETV fleet throughout a day of operation. The results show that minimizing the effects of disruptions on ETV fleet schedules does not prevent airport planners from maximizing the amount of saved emissions. Although the current approach only considers flight delays, the approach can be extended to include other disruptions, for example ETV unavailability.

Objective 4.1
Perform probabilistic prediction of flight delays using a data-driven approach.

Two probabilistic forecasting algorithms, Mixture Density Networks and Random Forests Regression, have been introduced in Chapter 5. They were applied to probabilistically predict flight delay based on flight schedule and weather data. Performance metrics suitable for probabilistic prediction were developed and used to evaluate the predictions. The results show that such probabilistic algorithms can predict flight delay with accompanying uncertainty information several days in advance, within a Continuous Ranked Probability Score value of 11 minutes.

Objective 4.2
Establish a general approach to flight delay and cancellation classification with imbalanced data.

Chapter 6 introduces a systematic approach to binary classification of flight delays.

The approach consists of first identifying the best performance metrics for the considered prediction problem, and then considering various sampling techniques and rates. In this way, the performance of solving any given imbalanced binary classification problems can be optimized. The approach was applied to predict flight delays given flight schedule data from a large hub airport.

Objective 4.3
Apply the predictions to improve the robustness of flight-to-gate assignments.

In Chapter 5, a deterministic MILP model for creating flight-to-gate assignments was extended to a stochastic MILP model. The probabilistic prediction algorithms defined for ***Objective 4.1*** were used to create the aircraft presence probabilities, an input to the stochastic model. Application of the model to flight schedule data from a small airport showed a potential reduction of up to 74% in the average number of conflicted aircraft per day, which is a measure of the robustness of the assignment. The schedule robustness can be controlled by varying a parameter within the stochastic model. Although the current model only considers uniform gates, the model is extendable to airports with e.g. airline-owned gates.

7.2. GENERAL CONCLUSIONS

In this section, general conclusions derived from the research described in Chapters 2-6 are outlined.

1. Resolving airport space constraints is essential to successful electric taxiing implementation and achieving sustainable airport surface movement.

In addition to the challenges identified regarding Air Traffic Control procedures, electricity capacity and disruption management, the limited availability of space at the airport forms a large challenge to electric taxiing implementation (Chapter 2). Based on available literature, service roads from gate areas to runways are considered a requirement for effective electric taxiing. The service roads also need to be wide enough for two ETVs to pass each other, or have enough passing locations. In addition, runway stands are required for (de)coupling of ETV and aircraft. Such stands take a lot of space, and are currently not present at most airports. Last, the infrastructure associated with providing additional charging capacity, likely required by airports to charge a fleet of ETVs, can also pose a challenge to the available space at airports.

2. The airport electricity capacity profile has a large influence on the efficient use of a fleet of ETVs.

Currently, the ETVs in operations at airports are not electrical, but diesel-powered. Once developed, the charging specifications of the actual ETV, in particular the battery capacity and charging rate, will largely determine the performance of an ETV. However, given these specifications, the electricity capacity profile and layout of the airport will determine the efficiency of the total ETV fleet. The case study of Amsterdam Airport Schiphol,

in combination with the expected values of ETV specifications, shows that night charging is expected to be essential, and limited charging during the day is needed to achieve an efficiency nearing to that associated with 24h unlimited charging availability (Chapter 3).

3. Heuristic approaches must be considered for airside operational scheduling problems.

Airside operational scheduling problem often entail large vehicle fleets, many tasks or locations to be assigned to, and a large time frame (Chapters 3 and 4). This causes the problems to be computationally expensive. Some heuristic approaches can provide solutions to scheduling problems with comparable quality to linear programming approaches, as seen in Chapter 3, and in literature regarding scheduling problems in other fields of study. Such approaches could be helpful for both strategic and tactical/adaptive scheduling.

4. A data-driven approach to operational scheduling will be most beneficial to airport planners.

Reassignment of airside operational schedules is generally performed by manually rearranging tasks. Manual reassignment may not always be the most efficient, and adds to the workload of airport planners. Data-driven techniques can contribute to the development of automated reassignment models. For example, the disrupted model developed in Chapter 4 is used to automate ETV-to-aircraft reassignment, while remaining as sustainable as possible. The stochastic flight-to-gate assignment model developed in Chapter 5 aims to reduce the changes needed during operation, by creating a more robust strategic model in advance. In addition, data-driven delay and cancellation prediction with extensive uncertainty information (Chapter 5) or optimized performance for imbalanced data (Chapter 6) can help airport planners reduce the effects of disruptions to schedules.

5. When developing approaches to solve prediction problems, one should not overlook the importance of selecting suitable metrics.

Binary classification, multi-class classification, point prediction and probabilistic prediction are examples of prediction problem categories. For each category, several types of metrics are used in literature. It is necessary to consider what metric is suitable for the goal of the prediction problem. The choice of metric can have a large influence on the way the prediction method is optimized or tuned. For example, considering binary classification, many studies only consider accuracy as a metric, which is unreliable for imbalanced classification problems (Chapter 6). In addition, for the relatively new approach of probabilistic prediction, a limited amount of metrics are currently available. In Chapter 5 several additional metrics are proposed. Possibilities for development of other metrics remain, each of which might be better suited to another prediction goal.

7.3. LIMITATIONS AND RECOMMENDATIONS

In this section, the limitations of the works in this dissertation are considered and recommendations for future work are outlined.

1. Increasing detail in current models of airside scheduling problems.

In this dissertation models have been introduced to represent the ETV-to-aircraft scheduling problem and the flight-to-gate assignment problem. Future ETV-to-aircraft scheduling models could benefit from more details with respect to routing and energy, by including for example acceleration of aircraft and vehicles where applicable, nonlinear charging profiles, personnel planning and multiple ETV types. Future Flight to Gate Assignment models could benefit from considering varying costs in assignments between aircraft, and considering differentiation in gate usage (e.g. by airlines, by destination or by aircraft size).

2. Considering other airside operational scheduling problems.

In addition to the ETV-to-aircraft scheduling problem and the flight-to-gate assignment problem, other airside operational scheduling problems exist that can benefit from a data-driven optimisation approach. For example, the aircraft-to-runway assignment and the assignment of electric Ground Support Equipment vehicles to charging locations.

3. Achieving realistic real-time optimization of ETV-to-aircraft schedules.

In Chapter 4 of this dissertation a disrupted model for ETV-to-aircraft assignment was developed, which works with thirty minute time windows. The disruptions, in the form of flight delays, are assumed to be known for the next half hour. In reality, disruptions can occur at any moment. In addition, disruptions can also include e.g. flight cancellations or ETV or personnel unavailability. Ideally, a scheduling model would be able to instantly change the assignment to accommodate these disruptions, append new aircraft to the schedule that are to be towed after the current timeframe, and remain a (near-)optimal schedule with respect to saved emissions. Such a model would require a combination of the optimisation approaches taken in Chapters 3 and 4 of this dissertation and the simulation approaches to electric taxiing assignment as taken by e.g. Zaninotto, Gauci, and Zammit [181]. This combined model would constitute the next step in integrating ETV fleet scheduling optimization into actual airport operations. An additional consideration for such a model could be to consider a buffer fleet of spare vehicles.

4. Applying probabilistic predictions to other airport operation optimisation models.

In this dissertation, probabilistic flight delay predictions were applied to improve the robustness of flight-to-gate assignments. Such predictions can also improve the robustness or efficiency of other operation optimisation models, such as arrival/departure sequencing and scheduling, and electric taxiing operational planning. Another possibility

would be to use probabilistic prediction algorithms to predict other circumstances at the airport, e.g. gate occupation rate. It needs to be noted that the possibilities for prediction rely on the availability and quality of data, and that the possibilities for improving airport operations with predictions rely in turn on the quality of these predictions.

5. Considering heuristic algorithms to solve airport scheduling problems.

Airport scheduling problems often have high complexity, because of the many aircraft in the considered schedule. The advantage of heuristic algorithms over linear programming models is the limited computation time. This is not only advantageous for strategic scheduling, but also for adaptive or real-time scheduling. For this reason, a heuristic approach could be beneficial for the combined model mentioned under item 3 above. Although the computation time of heuristic approaches is often a fraction of that of linear programming approaches, it can still be significant, as the model becomes more realistic and more complex. The scheduling problems typical for airports can limit the application of certain types of heuristic algorithms. For example, ETV-to-aircraft scheduling is hard to combine with a genetic algorithm approach, since aircraft are scheduled at fixed times, and cannot be reordered.

6. Furthering the other aspects of electric-taxiing implementation.

Last, we zoom out from the operational scheduling aspect of electric taxiing implementation. Before an optimized fleet of towing vehicles becomes practice, several obstacles in the development of ETVs need to be cleared. First, the current diesel-powered towing vehicles need to make way for an actual electric towing vehicle, which is still in development. Second, ATC-related organizations managing traffic on airports seeking to implement electric taxiing will need to update airport surface movement and safety protocols to include ETV movement on service roads and taxiways. The workload for ATC will increase as they will monitor the ETV fleet, while keeping in contact with ETV operators and aircraft pilots. Last, the airport itself will need to resolve space constraints mentioned in Section 7.2, and provide sufficient electricity capacity to maintain both an ETV fleet and all other airport processes that require electricity.

7.4. SCIENTIFIC CONTRIBUTIONS

In this section, the specific scientific contributions of this dissertation to the literature are identified.

1. This dissertation is the first study to provide a review on the operational management aspects of electric taxiing implementation. In addition, it identifies future challenges faced by airports during this implementation, from the routing, scheduling and charging aspects. The corresponding contribution led to the following journal publication:
   M. Zoutendijk, M.A. Mitici and J.M. Hoekstra, *Overview of Operational Management Solutions and Challenges for Electric Taxiing of Aircraft*, Research in Transportation Business & Management, **49**, 101019 (2023).

2. This dissertation is the first study to consider limits for airport electricity capacity in an ETV-to-aircraft scheduling model, bringing such scheduling models closer to realistic operation. In addition, it is the first to develop a heuristic approach to ETV fleet scheduling, based on Adaptive Large Neighbourhood Search. This approach allows airport planners to create a 24-hour ETV fleet schedule in a few hours. The corresponding contribution was submitted as the following publication under review:
M. Zoutendijk and M.A. Mitici, *Fleet Scheduling for Electric Towing of Aircraft under Limited Energy Capacity*, Energy, Special Issue "The Role of Smart Technologies in Energy Engineering", 294, p.130924 (2024)

3. This dissertation is the first study to develop an MILP approach to incorporate disruption management into an ETV-to-aircraft scheduling model. The resulting model generates a series of updated assignments throughout the day, both minimizing the number of changes to the schedule, and maximizing the amount of saved emissions. The corresponding contribution led to the following conference proceeding:
M. Zoutendijk, S. van Oosterom and M.A. Mitici, *Electric Taxiing with Disruption Management: Assignment of Electric Towing Vehicles to Aircraft*, AIAA AVIATION 2023 Forum (2023), p.4219

4. This dissertation is the first study to perform probabilistic prediction of flight delays. Using two machine learning algorithms, Mixture Density Networks and Random Forests Regression, flight delays are predicted in advance for an airport case study. Existing and newly developed metrics are used to evaluate the performance. In addition, a stochastic flight-to-gate assignment model was developed, where the robustness of the assignments can be controlled by a parameter. The predicted delays are an input to this model. The corresponding contribution led to the following journal publication, which was awarded *Honourable Mention Anna Valicek Competition, Airline Group of the International Federation of Operational Research Societies*, in 2022:
M. Zoutendijk and M.A. Mitici, *Probabilistic flight delay predictions using machine learning and applications to the flight-to-gate assignment problem*, Aerospace, **8(6)**, 152 (2021).

5. This dissertation is the first study to develop a systematic approach to overcome data imbalance in binary classification problems, which considers metrics, sampling rate and sampling techniques. The model is applied to the case of flight delay and cancellation prediction at a large airport. The corresponding contribution led to the following conference proceeding:
R. Hendrickx, **M. Zoutendijk**, M.A. Mitici and J. Schäfer, *Considering Airport Planners' Preferences and Imbalanced Datasets when Predicting Flight Delays and Cancellations*, IEEE/AIAA 40th Digital Avionics Systems Conference (2021).

7.5. SOCIETAL CONTRIBUTIONS

Aviation is an effective mode of transport that provides many people with the opportunity to travel fast and far. At the same time, the aviation sector creates a large portion of the worldwide emissions in the transport sector. The studies included in this dissertation aim to help increase the sustainability of air transport, to reduce the impact on the climate worldwide, and to help increase the efficiency of air transport, to reduce the resources needed to keep providing service to passengers.

The first chapters focus on electric taxiing, which has been shown to greatly reduce greenhouse gas emissions, noise and air pollution at the airport level. In Chapter 2, the current implementation status of electric taxiing has been investigated, and challenges to be overcome on the road to implementation have been identified. In Chapter 3, the specific challenge of limited electricity capacity at airports has been addressed, showing what charging a fleet of ETVs might look like in the near future, and what is necessary to maximize the distance towed sustainably. Chapter 4 forms a step from theoretical planning towards actual operation, with the inclusion of disruption management.

Chapters 5 and 6 aim to improve flight delay prediction, and show how to use predictions to improve airport operations, specifically the flight-to-gate assignment. Better disruption predictions can be used to retain or approximate the efficiency of undisrupted operations, so that one can reduce the need for extra material resources, personnel costs, and additional emissions.

Overall, the studies included in this dissertation contribute to society by providing airport operational and infrastructural planners with tools to improve the sustainability and efficiency of airport operations, so that passengers and the general public can enjoy more durable and reliable air transportation services.

ABBREVIATIONS

AAS	Amsterdam Airport Schiphol
ALNS	Adaptive Large Neighbourhood Search
APU	Auxiliary Power Unit
ASM	Airport Surface Movement
ATA	Actual Time of Arrival
ATC	Air Traffic Control
ATD	Actual Time of Departure
AUC	Average Under Curve
CA	Conflicted Aircraft
cdf	Cumulative Distribution Function
CRPS	Continuous Ranked Probability Score
DWC	Dynamic Wireless Charging
EASA	European Union Aviation Safety Agency
ESUT	Engine Start-Up Time
ET	Electric Taxiing
ETS	Electric Taxiing Systems
ETV	Electric Towing Vehicle
EV	Electric Vehicle
FAA	Federal Aviation Administration
FGAP	Flight-to-Gate Assignment Problem
FSA	Fleet Scheduling Assignment
GSE	Ground Support Equipment
KDE	Kernel Density Estimation
LP	Linear Programming
LRP	Location Routing Problem
MDN	Mixture Density Networks
MILP	Mixed-Integer Linear Programming
MAE	Mean Absolute Error
ML	Machine Learning
MLG	Main Landing Gear
MLP	Multilayer Perceptron
NB	Narrow-body
NLG	Nose Landing Gear
pdf	Probability Density Function
SMOTE	Synthetic Minority Oversampling Technique
SOC	State of Charge
STA	Scheduled Time of Arrival

STD	Scheduled Time of Departure
RF	Random Forests
RFR	Random Forests Regression
RMSE	Root Mean Square Error
RTM	Rotterdam The Hague Airport
RUS	Random Undersampling
UGT	Used Gate Time Slots
V2G	Vehicle to Grid
VRP	Vehicle Routing Problem
WB	Wide-body

GLOSSARY

Notation used throughout the dissertation is gathered below for reference.

Symbol	Explanation
	Airport representation
$G^S = (N^S, E^S)$	Service road network
$G^X = (N^X, E^X)$	Taxiway network
$n_{\text{cs},i}, N_{\text{cs}}$	Charging station i, number of charging stations
n_{dp}	ETV depot node
v^s, v^x	Speed on service roads / taxiways
	Aircraft routing
$A = A^S \cup A^e$	Set of aircraft (nonartificial + artificial)
d_a^X	Towing distance for aircraft a
$d^X(m,n), d^S(m,n)$	Distance over taxiway network / service road network from m to n
m_a	Mass of an aircraft
$N^{F,S}, N^F$	Number of flight movements during P without / with artificial aircraft
n_a^p, n_a^d	Pick-up and drop-off node of aircraft a
N^T, Δ	Number of time steps, step size
$P = [t^s, t^e]$	Scheduling time period
t_a^p, t_a^d	Pick-up and drop-off time of aircraft a
$t_{\text{sep}}, d_{\text{sep}}$	Separation time, distance
v^0	Rolling resistance base velocity
μ^0	Rolling resistance base coefficient
	Electric Towing Vehicles
m_{ETV}	Mass of the ETV
N^V, V	ETV fleet size, set of ETVs
P^c, P^X, P^S	Charging rate, towing power, driving power
Q	Battery capacity
$q_1^C(a), q_2^C(a)$	Energy needed to travel from the closest $n_{\text{cs},i}$ to n_a^p / from n_a^d to the closest $n_{\text{cs},i}$
$q^S(a,b), q^C(a,b)$	Energy needed to travel from n_a^d to n_b^p / from n_a^d to n_b^p via a $n_{\text{cs},i}$
$q_f^S(a)$	Energy needed to travel from n_{dp} to n_a^p
$q_f^S(v,a)$	Energy needed by ETV v to travel from its latest location to the pickup point of aircraft a
$q_l^S(a)$	Energy needed to travel from n_a^d to n_{dp}
$q^S(n,m)$	Energy needed to travel from node n to m

$q^X(a)$	Energy needed to tow aircraft a
$t^C_{\min}$	Minimum charging time
t^c	Connecting/Disconnecting time
$t^C(a,b)$	Time available between towing consecutive aircraft a and b
	ETV Assignment Model
$A^{\text{out}}_a, A^{\text{in}}_a$	Set of aircraft that can be towed by an ETV after/before it tows aircraft a
A^{PC}_a	Set of aircraft in A^{out}_a for which there is at least $t^C_{\min}$ time for effective charging
c_a	Indicates if ETV is charged after towing aircraft a
c^t_a, c^s_a	Indicates charging time / start time of charging of ETV
C_t	Electricity capacity for ETVs at timestep t
f_{vi}	The i-th aircraft from which ETV v departed in the previous solution
M	Large value
q_a	Indicates state of charge of ETV at the start of towing a
S	Total number of vehicles with unchanged schedules
s_v	Indicates if the aircraft towing list has changed w.r.t. the previous assignment for ETV v
t_{vi}	The i-th aircraft at which ETV v arrived in the previous solution
v_a	Vehicle that tows aircraft a
x_{ab}	Variable indicating if a and b are towed consecutively
x^f_a, x^l_a	Indicates if a is the first/last aircraft an ETV tows
x^f_{av}	Indicates if a is the first aircraft towed by ETV v
Y	Total number of aircraft towed in an ETV-to-aircraft assignment
y_a	Indicates if a is towed by an ETV or taxies by itself
α	Objective function weight factor
α_{at}, β_{at}	Indicates if charging of ETV v_a starts earlier/finishes later than time step t
γ_{at}	Indicates if ETV v_a is charged during time step t
	Adaptive Large Neighbourhood Search
$a^-_{i,j}, a^+_{i,j}$	Selection rate during segment j of removal/insertion heuristic i
$A^{\text{rem}}, A^{\text{ins}}_v$	Set of removable aircraft, set of aircraft insertable in the schedule of ETV v
h^+_i, h^-_j	Insertion heuristic i and removal heuristic j
N^{it}	Number of ALNS iterations
N^{rem}	Number of aircraft removed in heuristic
N^{seg}	Number of iterations in ALNS segment
$s, f(s), s^{\text{all}}$	Solution, its objective value, set of all solutions
s^i, s^c, s^l, s^b	Initial, candidate, last and best solution

$\pi^-_{i,j}, \pi^+_{i,j}$	Removal and insertion selection scores for heuristic i and segment j
$\bar{\pi}^-_{i,j}, \bar{\pi}^+_{i,j}$	Removal and insertion segment selection scores for heuristic i and segment j
	Simulated Annealing
c	Cooling rate
w	Start temperature control parameter
T	Temperature
ρ	Reaction factor
$\sigma_1, \sigma_2, \sigma_3$	Global, previous and accepted solution reward
	Probabilistic Forecasting
E	Error function for MDN
$\bar{f}_{1\sigma}$	Fraction of samples for which the delay lies within one standard deviation from the distributional mean.
$F(y_i)$	Cumulative distribution function of $p(y_i)$
m	Number of Gaussian components considered for MDN
n_e	Number of estimators used in RFR
N_f	Number of samples in the training set
$p(y_i\|\mathbf{x}_i) = p(y_i)$	Probability distribution of delay value y_i given feature values $\mathbf{x}_i$
$\mathbf{x}_i$	Aircraft sample i
y_i	Random variable for flight delay of aircraft sample i
$\bar{y}_i$	Actual flight delay for aircraft sample i
$\hat{y}_{i,j}$	The jth point estimate for the delay of flight sample i
$\alpha_j(\mathbf{x}_i), \mu_j(\mathbf{x}_i)\sigma_j(\mathbf{x}_i)$	Weight, mean and standard deviation of the j^{th} Gaussian component
μ_M	Distribution mean, $\mu_{\text{MDN}}(\mathbf{x}_i)$ or $\mu_{\text{RFR}}(\mathbf{x}_i)$
$\bar{\sigma}$	Sample average of the standard deviation
σ_M	Distribution standard deviation, $\sigma_{\text{MDN}}(\mathbf{x}_i)$ or $\sigma_{\text{RFR}}(\mathbf{x}_i)$
	Gate Assignment Model
c_{ij}	Costs of assigning aircraft i to gate j
$f_{X_i^{\text{arr}}}(t), f_{X_i^{\text{dep}}}(t)$	Probability density function of X_i^{arr} and X_i^{dep}
$F_{X_i^{\text{arr}}}(t), F_{X_i^{\text{dep}}}(t)$	Cumulative density function of X_i^{arr} and X_i^{dep}
G, g	Set of gates at an airport, number of gates
K, k	Set of time slots, number of time slots
N, n	Set of scheduled aircraft, number of aircraft
p_{it}	Probability that aircraft i is at the airport at time slot t
r	Maximum overlap probability threshold
s_{it}	Indicates if aircraft i is at the airport at time slot t
STA_i, STD_i	Scheduled times of arrival and departure of aircraft i
$X_i^{\text{arr}}, X_i^{\text{dep}}$	Random variable of arrival and departure times of aircraft i
x_{ijt}	Indicates if aircraft i is assigned to gate j at time slot t

$y_i^{\text{arr}}, y_i^{\text{dep}}$	Arrival and departure delay of aircraft i

List of Figures

LIST OF TABLES

BIBLIOGRAPHY

[1] Feyijimi Adegbohun, Annette von Jouanne, and Kwang Y. Lee. "Autonomous battery swapping system and methodologies of electric vehicles". In: *Energies* 12.4 (2019), pp. 1–14. ISSN: 19961073. DOI: 10.3390/en12040667.

[2] Sobhan Ahmadi. "Green Airport Operations: Conflict And Collision Free Taxiing Using Electric Powered Towing Alternatives". PhD thesis. 2019.

[3] Sobhan Ahmadi and Ali Akgunduz. "Airport operations with electric-powered towing alternatives under stochastic conditions". In: *Journal of Air Transport Management* 109.January (2023), p. 102392. ISSN: 09696997. DOI: 10.1016/j.jairtraman.2023.102392.

[4] Aircraft Commerce. "Wheeltug self-taxi system prepares for 2021 service entry". In: *Issue No. 131* (2020), pp. 30–31. URL: https://www.wheeltug.com/wp-content/uploads/2020/10/131{_}FLTOPS{_}A.pdf.

[5] Airport Technology. *Frankfurt Airport deploys fuel-saving Taxibot for taxiing aircraft*. 2015. URL: https://www.airport-technology.com/news/newsfrankfurt-airport-deploys-fuel-saving-taxibot-for-taxiing-aircraft-4517031/{\#}:{~}:text=Germany'sFrankfurtAirporthasstarted,AviationSafetyAgency(EASA).

[6] Airports International. *TaxiBot makes Delhi and Bengaluru debut*. 2023. URL: https://www.airportsinternational.com/article/taxibot-makes-delhi-and-bengaluru-debut.

[7] Airside International. *Remote towing: the future?* 2018. URL: https://www.airsideint.com/issue-article/remote-towing-the-future/ (visited on 02/07/2022).

[8] Marco Alderighi and Alberto A Gaggero. "Flight cancellations and airline alliances: Empirical evidence from Europe". In: *Transportation Research Part E: Logistics and Transportation Review* 116 (2018), pp. 90–101. DOI: https://doi.org/10.1016/j.tre.2018.05.008.

[9] Hugo Alonso and António Loureiro. "Predicting flight departure delay at Porto Airport: A preliminary study". In: *2015 7th International Joint Conference on Computational Intelligence (IJCCI)*. Vol. 3. IEEE. 2015, pp. 93–98.

[10] Yaseen Alwesabi et al. "A novel integration of scheduling and dynamic wireless charging planning models of battery electric buses". In: *Energy* 230 (2021), p. 120806. ISSN: 03605442. DOI: 10.1016/j.energy.2021.120806.

[11] Amsterdam Airport Schiphol. *Download a map: aircraft process maps*. 2021. URL: https://www.schiphol.nl/en/operations/page/maps/ (visited on 04/14/2022).

[12] Amsterdam Airport Schiphol. *Possible delays and cancellations due to strong wind*. 2021. URL: https://www.schiphol.nl/en/messages/possible-delays-and-cancellations-due-to-strong-wind (visited on 10/26/2022).

[13] Amsterdam Airport Schiphol. *Reducing energy consumption*. 2023. URL: https://www.schiphol.nl/en/schiphol-group/page/reducing-energy-consumption/.

[14] Amadeo Ascó. "Steady State Evolutionary Algorithm and Operators for the Airport Gate Assignment Problem". In: *Int J Adv Robot Automn* 4.1 (2019), p. 24. DOI: https://doi.org/10.15226/2473-3032%2F4%2F1%2F00139.

[15] Samet Ayhan, Pablo Costas, and Hanan Samet. "Predicting estimated time of arrival for commercial flights". In: *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 2018, pp. 33–42. DOI: https://doi.org/10.1145/3219819.3219874.

[16] Edzard van Baaren and Paul C. Roling. "Design of a zero emission aircraft towing system". In: *AIAA AVIATION Forum*. 2019, pp. 1–11. DOI: 10.2514/6.2019-2932.

[17] Hamsa Balakrishnan, Indira Deonandan, and Ioannis Simaiakis. "6. Opportunities for Reducing Surface Emissions Through Surface Movement Optimization". In: *MIT International Center for Air Transportation (ICAT)* (2008), pp. 1–36.

[18] Nicholas Bambos and Michael Bloem. "Ground delay program analytics with behavioral cloning and inverse reinforcement learning". In: *Journal of Aerospace Information Systems* (2015), pp. 299–313. DOI: https://doi.org/10.2514/1.I010304.

[19] Loris Belcastro et al. "Using scalable data mining for predicting flight delays". In: *ACM Transactions on Intelligent Systems and Technology (TIST)* 8.1 (2016), p. 5. DOI: https://doi.org/10.1145/2888402.

[20] Torben Bernatzky et al. "Development and evaluation of a speed guidance interface for trajectory-based dispatch towing". In: *AIAA/IEEE Digital Avionics Systems Conference*. Vol. September. 2017. ISBN: 9781538603659. DOI: 10.1109/DASC.2017.8102130.

[21] Michal Bischoff, Joschka; Maciejewski. "Electric Taxis in Berlin – Analysis of the Feasibility of a Large-Scale Transition". In: *Communications in Computer and Information Science* 531 (2015), pp. 343–351. ISSN: 18650937. DOI: 10.1007/978-3-319-24577-5.

[22] Christopher M Bishop. *Mixture density networks*. Tech. rep. 1994.

[23] Leo Breiman. "Random forests". In: *Machine Learning* 45 (2001), pp. 5–32.

[24] Luis Cadarso and Raúl de Celis. "Integrated airline planning: Robust update of scheduling and fleet balancing under demand uncertainty". In: *Transportation Research Part C: Emerging Technologies* 81 (2017), pp. 227–245. ISSN: 0968090X. DOI: 10.1016/j.trc.2017.06.003.

[25] Robert Camilleri and Aman Batra. "Assessing the environmental impact of aircraft taxiing technologies". In: *32nd Congress of the International Council of the Aeronautical Sciences, ICAS 2021* September (2021).

[26] Jia-Ming Cao and Adib Kanafani. “Real-time decision support for integration of airline flight cancellations and delays part I: mathematical formulation”. In: *Transportation Planning and Technology* 20:3 (1997), pp. 183–199. DOI: https://doi.org/10.1080/03081069708717588.

[27] Michael Carney, Jim Cunningham, Pádraig Dowling, and Ciaran Lee. “Predicting Probability Distributions for Surf Height Using an Ensemble of Mixture Density Networks”. In: *Proceedings of the 22nd international conference on Machine learning*. 2005. URL: https://doi.org/10.1145/1102351.1102366.

[28] Centraal Bureau voor de Statistiek. *Windenergie op land; productie en capaciteit per provincie*. 2023. URL: https://opendata.cbs.nl/statline/{\#}/CBS/nl/dataset/70960NED/table.

[29] Keshav Ram Chandramouleeswaran et al. “Machine learning prediction of airport delays in the US air transportation network”. In: *Proc. of the 2018 AIAA Aviation Technology, Integration, and Operations Conference* (2018). DOI: https://doi.org/10.2514/6.2018-3672.

[30] Nitesh V Chawla et al. “SMOTE: synthetic minority over-sampling technique”. In: *Journal of artificial intelligence research* 16 (2002), pp. 321–357. DOI: https://doi.org/10.1613/jair.953.

[31] Jun Chen and Meng Li. “Chained predictions of flight delay using machine learning”. In: *AIAA Scitech 2019 Forum* January (2019), pp. 1–25. DOI: 10.2514/6.2019-1661.

[32] Sun Choi et al. “Prediction of weather-induced airline delays based on machine learning algorithms”. In: *2016 IEEE/AIAA 35th Digital Avionics Systems Conference (DASC)*. IEEE. 2016, pp. 1–6. DOI: https://doi.org/10.1109/DASC.2016.7777956.

[33] David Cross, Kristy Kiernan, and Mark Scharf. “Airport Operations Delays and Possible Mitigation Through Electric Taxi Systems : A Qualitative Case Study”. In: *International Journal of Aviation Research* 12.01 (2020), pp. 1–19.

[34] Nihad E. Daidzic. “Determination of taxiing resistances for transport category airplane tractive propulsion”. In: *Advances in Aircraft and Spacecraft Science* 4.6 (2017), pp. 651–677. ISSN: 22875271. DOI: 10.12989/aas.2017.4.6.651.

[35] Gülesin Sena Daş, Fatma Gzara, and Thomas Stützle. “A review on airport gate assignment problems: Single versus multi objective approaches”. In: *Omega* 92 (2020), p. 102146. DOI: http://dx.doi.org/10.1016/j.omega.2019.102146.

[36] Indira Deonandan and Hamsa Balakrishnan. “Evaluation of strategies for reducing taxi-out emissions at airports”. In: *10th AIAA Aviation Technology, Integration and Operations (ATIO) Conference*. Vol. 3. September. 2010, pp. 1–14. ISBN: 9781617825132. DOI: 10.2514/6.2010-9370.

[37] Faiyaz Doctor et al. “Modelling the effect of electric aircraft on airport operations and infrastructure”. In: *Technological Forecasting and Social Change* 177 (2022), p. 121553. ISSN: 00401625. DOI: https://doi.org/10.1016/j.techfore.2022.121553.

[38] Dutch Plane Spotters. *Schiphol Runway Usage*. 2022. URL: https://www.dutchplanespotte nl/runways/.

[39] Niclas Dzikus et al. "Potential for fuel reduction through electric taxiing". In: *11th AIAA Aviation Technology, Integration,and Operations (ATIO) Conference*. 2011, pp. 1–9. ISBN: 9781600869419. DOI: 10.2514/6.2011-6931.

[40] Simon Emde, Hamid Abedinnia, and Christoph H. Glock. "Scheduling electric vehicles making milk-runs for just-in-time delivery". In: *IISE Transactions* 50.11 (2018), pp. 1013–1025. ISSN: 24725862. DOI: 10.1080/24725854.2018.1479899.

[41] EUROCONTROL. *Advanced surface movement guidance and control system*. 2020. URL: https://www.eurocontrol.int/service/advanced-surface-movement-guidance-and-control-system (visited on 04/05/2022).

[42] EUROCONTROL. *Eurocontrol Annual Network Operations Report 2019*.

[43] EUROCONTROL. *Eurocontrol Five-Year Forecast 2020-2024*.

[44] EUROCONTROL. *Eurocontrol Network Manager Annual Report 2019*.

[45] Eurocontrol. "Network Manager Annual Report". In: (). URL: https://www.eurocontrol.int/publication/network-manager-annual-report-2018.

[46] Eurocontrol. "Network Operations Report 2018". In: (). URL: https://www.eurocontrol.int/archive{_}download/all/node/10910.

[47] European Commission. *Reducing emissions from aviation*. 2021. URL: https://ec.europa.eu/clima/eu-action/transport-emissions/reducing-emissions-aviation{_}en (visited on 02/24/2022).

[48] European Commission. *Zero emission vehicles: first 'Fit for 55' deal will end the sale of new CO2 emitting cars in Europe by 2035*. 2022. URL: https://ec.europa.eu/commission/presscorner/detail/en/ip{_}22{_}6462 (visited on 08/07/2023).

[49] Federal Aviation Administration. *Aviation Climate Action Plan*. 2021. URL: https://www.faa.gov/sites/faa.gov/files/2021-11/Aviation{_}Climate{_}Action{_}Plan.pdf (visited on 03/23/2022).

[50] Martin Felder, Anton Kaifel, and Alex Graves. "Wind power prediction using mixture density recurrent neural networks". In: *European Wind Energy Conference and Exhibition 2010, EWEC 2010* 5.April 2010 (2010), pp. 3417–3424.

[51] Ahmed Foda et al. "A generic cost-utility-emission optimization for electric bus transit infrastructure planning and charging scheduling". In: *Energy* 277.April (2023), p. 127592. ISSN: 03605442. DOI: 10.1016/j.energy.2023.127592.

[52] Kate E. Forrest et al. "Charging a renewable future: The impact of electric vehicle charging intelligence on energy storage requirements to meet renewable portfolio standards". In: *Journal of Power Sources* 336 (2016), pp. 63–74. ISSN: 03787753. DOI: 10.1016/j.jpowsour.2016.10.048.

[53] Stanley Förster, Michael Schultz, and Hartmut Fricke. "Probabilistic Prediction of Separation Buffer to Compensate for the Closing Effect on Final Approach". In: *Aerospace* 8.2 (2021), p. 29. DOI: https://doi.org/10.3390/aerospace8020029.

[54] Christian M.M. Frey et al. “The vehicle routing problem with time windows and flexible delivery locations”. In: *European Journal of Operational Research* 308.3 (2023), pp. 1142–1159. ISSN: 03772217. DOI: 10.1016/j.ejor.2022.11.029.

[55] M. Galea et al. “Development of an aircraft wheel actuator for green taxiing”. In: *International Conference on Electrical Machines, ICEM*. 2014, pp. 2492–2498. ISBN: 9781479943890. DOI: 10.1109/ICELMACH.2014.6960537.

[56] Fred Glover. “Future paths for integer programming and links to artificial intelligence”. In: *Computers and Operations Research* 13.5 (1986), pp. 533–549. ISSN: 03050548. DOI: 10.1016/0305-0548(86)90048-1.

[57] Tilmann Gneiting and Matthias Katzfuss. “Probabilistic forecasting”. In: *Annual Review of Statistics and Its Application* 1 (2014), pp. 125–151. DOI: https://doi.org/10.1146/annurev-statistics-062713-085831.

[58] Dominik Goeke and Michael Schneider. “Routing a mixed fleet of electric and conventional vehicles”. In: *European Journal of Operational Research* 245.1 (2015), pp. 81–99. ISSN: 03772217. DOI: 10.1016/j.ejor.2015.01.049.

[59] Goldhofer. *Akasol Supplies High-Performance Battery Systems For Aircraft Pushback Tractor »Phoenix« E From Goldhofer*. 2021. URL: https://www.goldhofer.com/en/press/current-news/akasol-supplies-high-performance-battery-systems-for-aircraft-pushback-tractor-phoenix-e-from-goldhofer (visited on 05/17/2022).

[60] Shon Grabbe, Banavar Sridhar, and Avijit Mukherjee. “Clustering days and hours with similar airport traffic and weather conditions”. In: *Journal of Aerospace Information Systems* 11.11 (2014), pp. 751–763. DOI: https://doi.org/10.2514/1.I010212.

[61] N J F P Guillaume. *Finding the viability of us-ing an automated guided vehicle taxiing system for aircraft*. Tech. rep. 2018.

[62] Kyle Gulan, Eduardo Cotilla-Sanchez, and Yue Cao. “Charging Analysis of Ground Support Vehicles in an Electrified Airport”. In: *IEEE Transportation Electrification Conference (ITEC)*. 2019, pp. 1–6. ISBN: 9781538693100. DOI: 10.1109/ITEC.2019.8790550.

[63] Rui Guo, Yu Zhang, and Qing Wang. “Comparison of emerging ground propulsion systems for electrified aircraft taxi operations”. In: *Transportation Research Part C: Emerging Technologies* 44 (2014), pp. 98–109. ISSN: 0968090X. DOI: 10.1016/j.trc.2014.03.006.

[64] Hamed Hassanzadeh et al. “Load balancing for imbalanced data sets: classifying scientific artefacts for evidence based medicine”. In: *Pricai 2014: Trends in Artificial Intelligence* 8862 (2014), pp. 972–984.

[65] Katja Hein and Sebastian Baumann. “Acoustical comparison of conventional taxiing and dispatch towing - Taxibot’s contribution to ground noise abatement”. In: *30th Congress of the International Council of the Aeronautical Sciences (ICAS)* (2016), pp. 1–7.

[66] Rik Hendrickx et al. “Considering Airport Planners’ Preferences and Imbalanced Datasets when Predicting Flight Delays and Cancellations”. In: *AIAA/IEEE Digital Avionics Systems Conference* 2021-Octob (2021), pp. 1–19. ISSN: 21557209. DOI: 10.1109/DASC52595.2021.9594367.

[67] Gerhard Hiermann et al. “The Electric Fleet Size and Mix Vehicle Routing Problem with Time Windows and Recharging Stations”. In: *European Journal of Operational Research* 252.3 (2016), pp. 995–1018. ISSN: 03772217. DOI: 10.1016/j.ejor.2016.01.038.

[68] Geoffrey E Hinton. “Connectionist learning procedures”. In: *Machine Learning* III (1990), pp. 555–610. DOI: https://doi.org/10.1016/B978-0-08-051055-2.50029-8.

[69] Yuji Horiguchi et al. “Predicting Fuel Consumption and Flight Delays for Low-Cost Airlines”. In: *Innovative Applications of Artificial Intelligence (IAAI) Conference* (2017), pp. 4686–4693. DOI: https://doi.org/10.1609/aaai.v31i2.19095.

[70] Jakub Hospodka. “Cost-benefit analysis of electric taxi systems for aircraft”. In: *Journal of Air Transport Management* 39 (2014), pp. 81–88. ISSN: 09696997. DOI: 10.1016/j.jairtraman.2014.05.002.

[71] Mingyang Huang, Hong Nie, and Ming Zhang. “Dynamic analysis of ground steering response of aircraft with electric taxi system”. In: *Vibroengineering Procedia* 10 (2016), pp. 358–363. ISSN: 23450533.

[72] Mingyang Huang et al. “Main Wheel Prerotation and Ground Taxi Driven by Electric Taxi System”. In: *Journal of Aerospace Engineering* 32.6 (2019), p. 04019088. ISSN: 0893-1321. DOI: 10.1061/(asce)as.1943-5525.0001088.

[73] IATA. *Air Passenger Market Analysis: The year ends on a strong note for the global industry*. Tech. rep. 2022. URL: https://www.iata.org/en/iata-repository/publications/economic-reports/air-passenger-market-analysis---december-2022/.

[74] IATA. *Industry Statistics: Fact Sheet*. Tech. rep. 2023. URL: https://www.iata.org/en/iata-repository/pressroom/fact-sheets/industry-statistics/.

[75] IATA. *Net-Zero Carbon Emissions by 2050*. 2021. URL: https://www.iata.org/en/pressroom/2021-releases/2021-10-04-03/ (visited on 11/01/2022).

[76] International Airport Review. *Sustainable taxiing at Amsterdam Airport Schiphol*. 2022. URL: https://www.internationalairportreview.com/article/176357/sustainable-taxiing-at-amsterdam-airport-schiphol/.

[77] IowaStateUniversity. “ASOS-AWOS-METAR Data Download”. In: (). URL: https://mesonet.agron.iastate.edu/request/download.phtml.

[78] Israel Aerospace Industries. *IAI expands TaxiBot to more airports*. 2020. URL: https://www.iai.co.il/iai-expands-taxibot-to-more-airports (visited on 02/15/2022).

[79] Israel Aerospace Industries. *TaxiBot: Semi-Robotic Aircraft Tractor*. 2022. URL: https://www.iai.co.il/p/taxibot (visited on 03/29/2022).

[80] Eri Itoh and Mihaela Mitici. “Queue-based modeling of the aircraft arrival process at a single airport”. In: *Aerospace* 6.10 (2019), p. 103. DOI: https://doi.org/10.3390/aerospace6100103.

[81] Amin Jamili. “A robust mathematical model and heuristic algorithms for integrated aircraft routing and scheduling, with consideration of fleet assignment problem”. In: *Journal of Air Transport Management* 58 (2017), pp. 21–30. ISSN: 09696997. DOI: 10.1016/j.jairtraman.2016.08.008.

[82] Young Jae Jang. “Survey of the operation and system study on wireless charging electric vehicle systems”. In: *Transportation Research Part C: Emerging Technologies* 95.November 2017 (2018), pp. 844–866. ISSN: 0968090X. DOI: 10.1016/j.trc.2018.04.006.

[83] Ahmad I Z Jarrah et al. “A decision support framework for airline flight cancellations and delays”. In: *Transportation Science* 27(3) (1993), pp. 266–280. DOI: https://doi.org/10.1287/trsc.27.3.266.

[84] Yu Jiang, Zhihua Liao, and Honghai Zhang. “A collaborative optimization model for ground taxi based on aircraft priority”. In: *Mathematical Problems in Engineering* 2013 (2013). ISSN: 1024123X. DOI: 10.1155/2013/854364.

[85] Anish M. Kalliguddi and Aera K. Leboulluec. “Predictive Modeling of Aircraft Flight Delay”. In: *Universal Journal of Management* 5.10 (2017), pp. 485–491. ISSN: 2331-950X. DOI: 10.13189/ujm.2017.051003.

[86] Fabian Kelch et al. “Investigation and design of an axial flux permanent magnet machine for a commercial midsize aircraft electric taxiing system”. In: *IET Electrical Systems in Transportation* 8.1 (2017), pp. 52–60. ISSN: 20429746. DOI: 10.1049/iet-est.2017.0039.

[87] Merve Keskin and Bülent Çatay. “Partial recharge strategies for the electric vehicle routing problem with time windows”. In: *Transportation Research Part C: Emerging Technologies* 65 (2016), pp. 111–127. ISSN: 0968090X. DOI: 10.1016/j.trc.2016.01.013.

[88] Laura Khammash, Luca Mantecchini, and Vasco Reis. “Micro-simulation of airport taxiing procedures to improve operation sustainability: Application of semi-robotic towing tractor”. In: *5th IEEE International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS)*. 2017, pp. 616–621. ISBN: 9781509064847. DOI: 10.1109/MTITS.2017.8005587.

[89] Sang Hyun Kim et al. “Airport gate scheduling for passengers, aircraft, and operations”. In: *Journal of Air Transportation* 25.4 (2017), pp. 109–114. DOI: https://doi.org/10.2514/1.D0079.

[90] Young Jin Kim et al. “A deep learning approach to flight delay prediction”. In: *2016 IEEE/AIAA 35th Digital Avionics Systems Conference (DASC)*. IEEE. 2016, pp. 1–6. DOI: https://doi.org/10.1109/DASC.2016.7778092.

[91] S. Kirkpatrick, C. D. Gelatt Jr, and M. P. Vecchi. “Optimization by simulated annealing”. In: *Science* 220(4598) (1983), pp. 671–680. DOI: https://doi.org/10.1126/science.220.4598.671.

[92] Imke C Kleinbekman, Mihaela Mitici, and Peng Wei. “Rolling-Horizon Electric Vertical Takeoff and Landing Arrival Scheduling for On-Demand Urban Air Mobility”. In: *Journal of Aerospace Information Systems* 17.3 (2020), pp. 150–159. DOI: https://doi.org/10.2514/1.I010776.

[93] Anja Kollmuss. *Carbon offsetting & air travel: Part 1: CO2-emissions calculations*. Tech. rep. Stockholm Environment Institute, 2008. URL: https://mediamanager.sei.org/documents/Publications/Climate/sei{_}air{_}travel{_}emissions{_}paper1{_}27{_}03{_}09.pdf.

[94] Miroslav Kubat and Stan Matwin. “Addressing the curse of imbalanced training sets: one-sided selection”. In: *Proc. of the 14th Int. Conf. on Machine Learning* (1997), pp. 179–186.

[95] Miguel Lambelho et al. “Assessing strategic flight schedules at an airport using machine learning-based flight delay and cancellation predictions”. In: *Journal of Air Transport Management* 82 (2020), p. 101737. ISSN: 09696997. DOI: 10.1016/j.jairtraman.2019.101737.

[96] Jane Lee, Lavanya Marla, and Alexandre Jacquillat. “Dynamic Airline Disruption Management Under Airport Operating Uncertainty”. In: *SSRN Electronic Journal* February (2018). DOI: 10.2139/ssrn.3082518.

[97] Jane Lee, Lavanya Marla, and Alexandre Jacquillat. “Dynamic Disruption Management in Airline Networks Under Airport Operating Uncertainty”. In: *Transportation Science* 54.4 (2020), pp. 973–997. ISSN: 15265447. DOI: 10.1287/trsc.2020.0983.

[98] Xinzhou Li et al. “A cost-benefit analysis of V2G electric vehicles supporting peak shaving in Shanghai”. In: *Electric Power Systems Research* 179.September 2019 (2020), p. 106058. ISSN: 03787796. DOI: 10.1016/j.epsr.2019.106058.

[99] Yuping Lin et al. “Multistage large-scale charging station planning for electric buses considering transportation network and power grid”. In: *Transportation Research Part C: Emerging Technologies* 107.September 2018 (2019), pp. 423–443. ISSN: 0968090X. DOI: 10.1016/j.trc.2019.08.009.

[100] Jules L’Ortye, Mihaela Mitici, and Hendrikus G Visser. “Robust flight-to-gate assignment with landside capacity constraints”. In: *Journal of Transport Planning and Technology* (2021). DOI: 10.1080/03081060.2021.1919347.

[101] Luchtverkeersleiding Nederland. *Runway use*. 2022. URL: https://en.lvnl.nl/environment/runway-use.

[102] M. Lukic et al. “State of the Art of Electric Taxiing Systems”. In: *IEEE International Conference on Electrical Systems for Aircraft, Railway, Ship Propulsion and Road Vehicles and International Transportation Electrification Conference (ESARS-ITEC)*. 2018. ISBN: 9781538641927. DOI: 10.1109/ESARS-ITEC.2018.8607786.

[103] Milos Lukic et al. “Review, Challenges, and Future Developments of Electric Taxiing Systems”. In: *IEEE Transactions on Transportation Electrification* 5.4 (2019), pp. 1441–1457. ISSN: 23327782. DOI: 10.1109/TTE.2019.2956862.

[104] Khizir Mahmud, M. J. Hossain, and Jayashri Ravishankar. “Peak-Load Management in Commercial Systems with Electric Vehicles”. In: *IEEE Systems Journal* 13.2 (2019), pp. 1872–1882. ISSN: 19379234. DOI: 10.1109/JSYST.2018.2850887.

[105] Rami Sabet Mangoubi. “A Linear Programming Solution to the Gate Assignment Problem”. PhD thesis. 1980.

[106] Suvojit Manna et al. “A statistical approach to predict flight delay using gradient boosted decision tree”. In: *ICCIDS 2017 - International Conference on Computational Intelligence in Data Science, Proceedings* 2018-Janua (2018), pp. 1–5. DOI: 10.1109/ICCIDS.2017.8272656.

[107] Dimitris Margaritis et al. “Electric commercial vehicles: Practical perspectives and future research directions”. In: *Research in Transportation Business and Management* 18 (2016), pp. 4–10. ISSN: 22105395. DOI: 10.1016/j.rtbm.2016.01.005.

[108] James E Matheson and Robert L Winkler. “Scoring rules for continuous probability distributions”. In: *Management science* 22.10 (1976), pp. 1087–1096. DOI: https://doi.org/10.1287/mnsc.22.10.1087.

[109] Daniele Micci-Barreca. “A preprocessing scheme for high-cardinality categorical attributes in classification and prediction problems”. In: *ACM SIGKDD Explorations Newsletter* 3 (2001), pp. 27–32. DOI: https://doi.org/10.1145/507533.507538.

[110] John M. Miller et al. “ORNL experience and challenges facing dynamic wireless power charging of EV’s”. In: *IEEE Circuits and Systems Magazine* 15.2 (2015), pp. 40–53. ISSN: 1531636X. DOI: 10.1109/MCAS.2015.2419012.

[111] Mihaela Mitici, Madalena Pereira, and Fabrizio Oliviero. “Electric flight scheduling with battery-charging and battery-swapping opportunities”. In: *EURO Journal on Transportation and Logistics* 11.February (2022), p. 100074. ISSN: 21924376. DOI: 10.1016/j.ejtl.2022.100074.

[112] Hyung Bin Moon et al. “Forecasting electricity demand of electric vehicles by analyzing consumers’ charging patterns”. In: *Transportation Research Part D: Transport and Environment* 62.February (2018), pp. 64–79. ISSN: 13619209. DOI: 10.1016/j.trd.2018.02.009.

[113] Robert Morris et al. “Self-Driving aircraft towing vehicles: A preliminary report”. In: *AAAI Workshop - Technical Report* WS-15-05.December (2015), pp. 41–48.

[114] Eric Mueller and Gano Chatterji. “Analysis of aircraft arrival and departure delay characteristics”. In: *AIAA Aircraft Technology, Integration, and Operations (ATIO) 2002 Technical Forum* (2002). DOI: 10.2514/6.2002-5866.

[115] Munich International Airport. *E-mobility at Munich Airport picks up speed.* 2021. URL: https://www.munich-airport.com/press-all-electric-aircraft-pushback-tug-in-operation-10524189 (visited on 05/17/2022).

[116] Munich International Airport. *Traffic and Safety Rules: for the non-public area at Munich Airport*. Tech. rep. Munich International Airport, 2016, p. 8. URL: https://www.munich-airport.de/{_}b/0000000000000008368639bb5e302e6f/traffic-safety-rules-2016.pdf.

[117] M. Muthukrishnan and Rekibuddin Ahmed. *Green taxiing solution at Delhi Airport*. Tech. rep. 2022. URL: https://www.icao.int/environmental-protection/Documents/EnvironmentalReports/2022/ENVReport2022{_}Art40.pdf.

[118] Yann Nicolas. "eTaxi - Taxiing aircraft with engines stopped". In: *Fast 51* (2013), pp. 2–10. URL: https://www.smartcockpit.com/docs/taxiing-aircraft-with-engines-stopped.pdf.

[119] NIO. *NIO Power: The World's First Smart Power Solution*. URL: https://www.nio.com/nio-power (visited on 03/31/2022).

[120] H. Nojoumi, I. Dincer, and G. F. Naterer. "Greenhouse gas emissions assessment of hydrogen and kerosene-fueled aircraft propulsion". In: *International Journal of Hydrogen Energy* 34.3 (2009), pp. 1363–1369. ISSN: 03603199. DOI: 10.1016/j.ijhydene.2008.11.017.

[121] Khusnul Novianingsih and Rieske Hadianti. "Modeling flight departure delay distributions". In: *2014 International Conference on Computer, Control, Informatics and Its Applications (IC3INA)*. IEEE. 2014, pp. 30–34. DOI: https://doi.org/10.1109/IC3INA.2014.7042596.

[122] Nikolai Okuniek and Dustin Beckmann. "Towards higher level of A-SMGCS: Handshake of electric taxi and trajectory-based taxi operations". In: *AIAA/IEEE Digital Avionics Systems Conference*. Vol. September. 2017. ISBN: 9781538603659. DOI: 10.1109/DASC.2017.8102047.

[123] Luis Oliveira et al. "Wireless charging of electric taxis: Understanding the facilitators and barriers to its introduction". In: *Sustainability* 12.21 (2020), pp. 1–21. ISSN: 20711050. DOI: 10.3390/su12218798.

[124] Simon van Oosterom and Mihaela Mitici. "Analyzing the impact of battery capacity and charging protocols when dispatching electric vehicles for aircraft towing". In: *AIAA AVIATION Forum*. 2022, pp. 1–11. ISBN: 9781624106354. DOI: 10.2514/6.2022-3920.

[125] Simon van Oosterom, Mihaela Mitici, and Jacco Hoekstra. "Dispatching a fleet of electric towing vehicles for aircraft taxiing with conflict avoidance and efficient battery charging". In: *Transportation Research Part C: Emerging Technologies* 147.January (2023). ISSN: 0968090X. DOI: 10.1016/j.trc.2022.103995.

[126] Sergio Ortega Alba and Mario Manana. "Characterization and analysis of energy demand patterns in airports". In: *Energies* 10.1 (2017), pp. 1–35. ISSN: 19961073. DOI: 10.3390/en10010119.

[127] Sergio Ortega Alba and Mario Manana. "Energy research in airports: A review". In: *Energies* 9.5 (2016), pp. 1–19. ISSN: 19961073. DOI: 10.3390/en9050349.

[128] David Pisinger and Stefan Ropke. “A general heuristic for vehicle routing problems”. In: *Computers and Operations Research* 34.8 (2007), pp. 2403–2435. ISSN: 03050548. DOI: 10.1016/j.cor.2005.09.012.

[129] Maria Nadia Postorino, Luca Mantecchini, and Ettore Gualandi. “Integration between aircraft and handling vehicles during taxiing procedures to improve airport sustainability”. In: *International Journal of Transport Development and Integration* 1.1 (2017), pp. 28–42. ISSN: 20588313. DOI: 10.2495/TDI-V1-N1-28-42.

[130] Raziur Rahman et al. “Design of probabilistic random forests with applications to anticancer drug sensitivity prediction”. In: *Cancer informatics* 14 (2015), CIN–S30794. DOI: https://doi.org/10.4137/CIN.S30794.

[131] Tsarafidy Raminosoa et al. “Feasibility and electromagnetic design of direct drive wheel actuator for green taxiing”. In: *IEEE Energy Conversion Congress and Exposition: Energy Conversion Innovation for a Clean Energy Future (ECCE)*. IEEE, 2011, pp. 2798–2804. ISBN: 9781457705427. DOI: 10.1109/ECCE.2011.6064145.

[132] Fabrizio Re and Ricardo De Castroy. “Energetically optimal path following for electric aircraft taxi systems based on convex optimization”. In: *IEEE International Electric Vehicle Conference (IEVC)*. 2014. ISBN: 9781479960750. DOI: 10.1109/IEVC.2014.7056205.

[133] P. C. Roling and H. G. Visser. “Optimal Airport Surface Traffic Planning Using Mixed-Integer Linear Programming”. In: *International Journal of Aerospace Engineering* (2008), pp. 1–11. ISSN: 1687-5966. DOI: 10.1155/2008/732828.

[134] Paul Roling, Pjotr Sillekens, and Richard Curran. “The effects of electric taxi systems on airport surface congestion”. In: *15th AIAA Aviation Technology, Integration, and Operations Conference*. 2015, pp. 1–10. ISBN: 9781624103698. DOI: 10.2514/6.2015-2592.

[135] Stefan Ropke and David Pisinger. “An adaptive large neighborhood search heuristic for the pickup and delivery problem with time windows”. In: *Transportation Science* 40.4 (2006), pp. 455–472. ISSN: 15265447. DOI: 10.1287/trsc.1050.0135.

[136] Daniel Alejandro Rossit, Fernando Tohmé, and Mariano Frutos. “A data-driven scheduling approach to smart manufacturing”. In: *Journal of Industrial Information Integration* 15.April (2019), pp. 69–79. ISSN: 2452414X. DOI: 10.1016/j.jii.2019.04.003.

[137] Abdulrazaq Lemu Salihu, Shannon M. Lloyd, and Ali Akgunduz. “Electrification of airport taxiway operations: A simulation framework for analyzing congestion and cost”. In: *Transportation Research Part D: Transport and Environment* 97.July (2021). ISSN: 13619209. DOI: 10.1016/j.trd.2021.102962.

[138] F. Salucci et al. “An optimization model for airport infrastructures in support to electric aircraft.” In: *IEEE Milan PowerTech* (2019), pp. 1–5. DOI: https://doi.org/10.1109/PTC.2019.8810713.

[139] Nicolae Scarlat, Matteo Prussi, and Monica Padella. “Quantification of the carbon intensity of electricity produced and used in Europe”. In: *Applied Energy* 305 (2022), p. 117901. ISSN: 03062619. DOI: 10.1016/j.apenergy.2021.117901.

[140] Michael Schier et al. “High Integrated Electric Machine for Aircraft Autonomous Taxiing”. In: *International Conference on Electric Vehicles and Renewable Energies (EVER 11)*. 2011.

[141] Maximilian Schiffer and Grit Walther. “The electric location routing problem with time windows and partial recharging”. In: *European Journal of Operational Research* 260.3 (2017), pp. 995–1013. ISSN: 03772217. DOI: 10.1016/j.ejor.2017.01.011.

[142] Schiphol. *Is sustainable taxiing possible at a busy Schiphol?* 2021. URL: https://www.schiphol.nl/en/innovation/blog/is-sustainable-taxiing-possible-at-a-busy-schiphol/ (visited on 04/06/2022).

[143] Schiphol. *Schiphol Regulations*. Tech. rep. Amsterdam Airport Schiphol, 2022. URL: https://www.schiphol.nl/en/download/1609745377/43q9kGoE92CccmEeC6awa4pdf.

[144] Schiphol. *Sustainable taxiing and the Taxibot*. 2020. URL: https://www.schiphol.nl/en/download/b2b/1594319068/40LYaERw5x8BuRuXpVQkiK.pdf (visited on 02/03/2022).

[145] Schiphol. *The benefits of sustainable taxiing*. 2021. URL: https://www.schiphol.nl/en/innovation/page/the-benefits-of-sustainable-taxiing/ (visited on 02/23/2022).

[146] Schiphol. *What is sustainable taxiing? (Part 1)*. 2021. URL: https://www.schiphol.nl/en/innovation/page/what-is-sustainable-taxiing-part-1/ (visited on 02/03/2022).

[147] Schiphol Group. *Schiphol Developer Center*. 2022. URL: https://developer.schiphol.nl/login.

[148] Lisa Schlosser et al. “Distributional regression forests for probabilistic precipitation forecasting in complex terrain”. In: *Institute of Mathematical Statistics* (2019). DOI: 10.1214/19-AOAS1247.

[149] Johannes Schmidt et al. “Using battery-electric AGVs in container terminals - Assessing the potential and optimizing the economic viability”. In: *Research in Transportation Business and Management* 17 (2015), pp. 99–111. ISSN: 22105395. DOI: 10.1016/j.rtbm.2015.09.002.

[150] Michael Schmidt et al. “Challenges for ground operations arising from aircraft concepts using alternative energy”. In: *Journal of Air Transport Management* 56.Part B (2016), pp. 107–117. ISSN: 09696997. DOI: 10.1016/j.jairtraman.2016.04.023. URL: http://dx.doi.org/10.1016/j.jairtraman.2016.04.023.

[151] Mike Schuster. “Better Generative Models for Sequential Data Problems: Bidirectional Recurrent Mixture Density Networks.” In: *Advances in Neural Information Processing Systems*. 1999, pp. 589–595.

[152] Scopus. *Scopus*. 2022. URL: https://www.scopus.com/search/form.uri?display=basic{\#}basic (visited on 03/11/2022).

[153] Michael Seelhorst and Mark Hansen. "Flight cancellation behavior and delay savings". In: *5th International Conference on Research in Air Transportation*. 2012.

[154] Wei Shao et al. "Flight delay prediction using airport situational awareness map". In: *GIS: Proceedings of the ACM International Symposium on Advances in Geographic Information Systems* (2019), pp. 432–435. DOI: 10.1145/3347146.3359079.

[155] Sacha Silvester et al. "Exploring design scenarios for large-scale implementation of electric vehicles; The Amsterdam Airport Schiphol case". In: *Journal of Cleaner Production* 48 (2013), pp. 211–219. ISSN: 09596526. DOI: 10.1016/j.jclepro.2012.07.053.

[156] Giuseppe Sirigu et al. "Autonomous taxi operations: Algorithms for the solution of the routing problem". In: *AIAA Information Systems-AIAA Infotech at Aerospace*. 2018, pp. 1–7. ISBN: 9781624105272. DOI: 10.2514/6.2018-2143.

[157] Smart Airport Systems. *Taxibot International*. 2022. URL: https://www.taxibot-international.com/ (visited on 02/03/2022).

[158] J. W. Smeltink et al. "An Optimisation Model for Airport Taxi Scheduling". In: *INFORMS Annual Meeting*. 2004, pp. 1–25.

[159] Laura Soares and Hao Wang. "Economic feasibility analysis of charging infrastructure for electric ground fleet in airports". In: *Transportation Research Record* 2675.12 (2021), pp. 1–12. ISSN: 21694052. DOI: 10.1177/03611981211033859.

[160] Mojdeh Soltani et al. "An eco-friendly aircraft taxiing approach with collision and conflict avoidance". In: *Transportation Research Part C: Emerging Technologies* 121.December 2019 (2020), p. 102872. ISSN: 0968090X. DOI: 10.1016/j.trc.2020.102872.

[161] Alice Sternberg et al. "A review on flight delay prediction". In: *Transport Reviews* (2017). DOI: https://doi.org/10.48550/arXiv.1703.06118.

[162] Xindi Tang, Xi Lin, and Fang He. "Robust scheduling strategies of electric buses under stochastic traffic conditions". In: *Transportation Research Part C: Emerging Technologies* 105.March (2019), pp. 163–182. ISSN: 0968090X. DOI: 10.1016/j.trc.2019.05.032.

[163] Enrico Testa et al. "Analysis of environmental benefits resulting from use of hydrogen technology in handling operations at airports". In: *Clean Technologies and Environmental Policy* 16.5 (2014), pp. 875–890. ISSN: 16189558. DOI: 10.1007/s10098-013-0678-3.

[164] Balasubramanian Thiagarajan et al. "A machine learning approach for prediction of on-time performance of flights". In: *AIAA/IEEE Digital Avionics Systems Conference* (2017), pp. 5–10. ISSN: 21557209. DOI: 10.1109/DASC.2017.8102138.

[165] Ivan Tomek. "An experiment with the edited Nearest-Neighbour Rule". In: *IEEE Transactions on Systems. Man, and Cybernetics* 6 (1976), pp. 448–452.

[166] Yufeng Tu, Michael O Ball, and Wolfgang S Jank. “Estimating flight departure delay distributions—A statistical approach with long-term trend and short-term pattern”. In: *Journal of the American Statistical Association* 103.481 (2008), pp. 112–125. DOI: 10.1198/016214507000000257.

[167] Enis T. Turgut, Oznur Usanmaz, and Marc A. Rosen. “Empirical model assessment of commercial aircraft emissions according to flight phases”. In: *International Journal of Energy and Environmental Engineering* 4.1 (2013), pp. 1–12. ISSN: 22516832. DOI: 10.1186/2251-6832-4-15.

[168] Moslem Uddin et al. “A review on peak load shaving strategies”. In: *Renewable and Sustainable Energy Reviews* 82 (2018), pp. 3323–3332. ISSN: 18790690. DOI: 10.1016/j.rser.2017.10.056.

[169] Murat Pasa Uysal and M. Ziya Sogut. “An integrated research for architecture-based energy management in sustainable airports”. In: *Energy* 140 (2017), pp. 1387–1397. ISSN: 03605442. DOI: 10.1016/j.energy.2017.05.199.

[170] Parth Vaishnav. “Costs and benefits of reducing fuel burn and emissions from taxiing aircraft”. In: *Transportation Research Record* (2014), pp. 65–77. ISSN: 03611981. DOI: 10.3141/2400-08.

[171] Oscar R P Van Schaijk and Hendrikus G Visser. “Robust flight-to-gate assignment using flight presence probabilities”. In: *Transportation Planning and Technology* 40.8 (2017), pp. 928–945. ISSN: 10290354. DOI: 10.1080/03081060.2017.1355887.

[172] Julian Vossen, Baptiste Feron, and Antonello Monti. “Probabilistic Forecasting of Household Electrical Load Using Artificial Neural Networks”. In: *IEEE International Conference on Probabilistic Methods Applied to Power Systems (PMAPS)*. IEEE, 2018, pp. 1–6. ISBN: 9781538635964. DOI: https://doi.org/10.1109/PMAPS.2018.8440559.

[173] Andreas Westenberger. “Hydrogen and Fuel Cells: Mobile Application in Aviation”. In: *Hydrogen and Fuel Cell: Technologies and Market Perspectives*. Springer, 2016, pp. 107–125. ISBN: 978-3-662-44972-1. DOI: 10.1007/978-3-662-44972-1_5.

[174] Ji Wu et al. “Electric vehicle charging scheduling considering infrastructure constraints”. In: *Energy* 278.February (2023). ISSN: 03605442. DOI: 10.1016/j.energy.2023.127806.

[175] Yue Xiang et al. “Techno-economic design of energy systems for airport electrification: A hydrogen-solar-storage integrated microgrid solution”. In: *Applied Energy* 283 (2021), p. 116374. ISSN: 03062619. DOI: 10.1016/j.apenergy.2020.116374.

[176] Jun Xu et al. “Real-time prediction of taxi demand using recurrent neural networks”. In: *IEEE Transactions on Intelligent Transportation Systems* 19.8 (2017), pp. 2572–2581. DOI: https://doi.org/10.1109/TITS.2017.2755684.

[177] Jun Yang and Hao Sun. "Battery swap station location-routing problem with capacitated electric vehicles". In: *Computers and Operations Research* 55 (2015), pp. 217–232. ISSN: 03050548. DOI: 10.1016/j.cor.2014.07.003.

[178] Bin Yu et al. "Flight delay prediction for commercial air transport: A deep learning approach". In: *Transportation Research Part E: Logistics and Transportation Review* 125.March (2019), pp. 203–221. ISSN: 13665545. DOI: 10.1016/j.tre.2019.03.013.

[179] Chuhang Yu, Dong Zhang, and H. Y.K. Lau. "MIP-based heuristics for solving robust gate assignment problems". In: *Computers and Industrial Engineering* 93 (2016), pp. 171–191. ISSN: 03608352. DOI: 10.1016/j.cie.2015.12.013.

[180] Stefano Zaninotto, Jason Gauci, and Brian Zammit. "A Testbed for Performance Analysis of Algorithms for Engineless Taxiing with Autonomous Tow Trucks". In: *AIAA/IEEE Digital Avionics Systems Conference*. Vol. October. IEEE, 2021. ISBN: 9781665434201. DOI: 10.1109/DASC52595.2021.9594411.

[181] Stefano; Zaninotto, Jason; Gauci, and Brian; Zammit. "Tow Truck Taxi Algorithm: An Engineless Taxi Operations System Using Battery-Operated Autonomous Tow Trucks". In: *33rd Congress of the International Council of the Aeronautical Sciences*. 2022, pp. 1–18.

[182] Stefano Zaninotto et al. "Design of a Human-in-the-Loop Aircraft Taxi Optimisation System Using Autonomous Tow Trucks". In: *AIAA AVIATION Forum*. 2019, pp. 1–13. DOI: 10.2514/6.2019-2931.

[183] Heiga Zen and Andrew Senior. "Deep mixture density networks for acoustic modeling in statistical parametric speech synthesis". In: *IEEE international conference on acoustics, speech and signal processing (ICASSP)*. IEEE. 2014, pp. 3844–3848. DOI: https://doi.org/10.1109/ICASSP.2014.6854321.

[184] Weili Zeng et al. "A data-driven flight schedule optimization model considering the uncertainty of operational displacement". In: *Computers and Operations Research* 133.April (2021), p. 105328. ISSN: 03050548. DOI: 10.1016/j.cor.2021.105328.

[185] Jinhua Zhang et al. "Short-term forecasting and uncertainty analysis of wind turbine power based on long short-term memory network and Gaussian mixture model". In: *Applied Energy* 241 (2019), pp. 229–244. DOI: https://doi.org/10.1016/j.apenergy.2019.03.044.

[186] Lei Zhang et al. "Probability Density Forecasting of Wind Speed Based on Quantile Regression and Kernel Density Estimation". In: *Energies* 13.22 (2020), p. 6125. DOI: https://doi.org/10.3390/en13226125.

[187] Ming Zhang et al. "Assessment method of fuel consumption and emissions of aircraft during taxiing on airport surface under given meteorological conditions". In: *Sustainability (Switzerland)* 11.21 (2019). ISSN: 20711050. DOI: 10.3390/su11216110.

[188] Yang Zhao, Zoie Shui-Yee Wong, and Kwok Leung Tsui. "A framework of rebalancing imbalanced healthcare data for rare events' classification: a case of look-alike sound-alike mix-up incident detection". In: *Journal of Healthcare Engineering* (2018). DOI: https://doi.org/10.1155/2018/6275435.

[189] Guo Zhenfeng et al. "The electric vehicle routing problem with time windows using genetic algorithm". In: *Proceedings of 2017 IEEE 2nd Advanced Information Technology, Electronic and Automation Control Conference, IAEAC 2017* March (2017), pp. 635–639. DOI: 10.1109/IAEAC.2017.8054093.

[190] Mike Zoutendijk and Mihaela Mitici. "Fleet scheduling for electric towing of aircraft under limited airport energy capacity". In: *Energy* 294 (2024), p. 130924. ISSN: 03605442. DOI: 10.1016/j.energy.2024.130924.

[191] Mike Zoutendijk and Mihaela Mitici. "Probabilistic flight delay predictions using machine learning and applications to the flight-to-gate assignment problem". In: *Aerospace* 8.6 (2021). ISSN: 22264310. DOI: 10.3390/aerospace8060152.

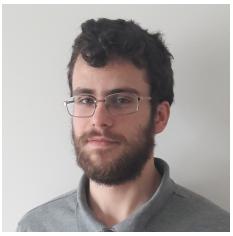
[192] Mike Zoutendijk, Mihaela Mitici, and Jacco Hoekstra. "An investigation of operational management solutions and challenges for electric taxiing of aircraft". In: *Research in Transportation Business and Management* 49.January (2023), p. 101019. ISSN: 22105395. DOI: 10.1016/j.rtbm.2023.101019.

[193] Mike Zoutendijk, Simon van Oosterom, and Mihaela Mitici. "Electric Taxiing with Disruption Management: Assignment of Electric Towing Vehicles to Aircraft". In: *AIAA Aviation Forum 2023*. 2023, pp. 1–20. DOI: https://doi.org/10.2514/6.2023-4219.

[194] Merve Şeker and Nilay Noyan. "Stochastic optimization models for the airport gate assignment problem". In: *Transportation Research Part E: Logistics and Transportation Review* 48.2 (2012), pp. 438–459. ISSN: 13665545. DOI: 10.1016/j.tre.2011.10.008.

CURRICULUM VITÆ

Mike ZOUTENDIJK



Mike Zoutendijk was born on March 16th, 1997 in Dordrecht, The Netherlands. In 2014, he started a double Bachelor of Science degree in Applied Mathematics and Applied Physics at Delft University of Technology (TU Delft). He ended the double Bachelor by conducting research on "A cascade of autoresonances in an accelerating elevator cable system". In 2017, he started the master Applied Mathematics at TU Delft, specializing in Computational Science & Engineering. During the Master program, he investigated how to take temporal aliasing into account when calculating wave current velocity and wave spectrum from marine radar data, during an internship at offshore firm NextOcean and software firm VORtech. In 2019, he conducted his MSc thesis "Applying Deflation Methods in a Topology Optimization Procedure".

In 2020, Mr. Zoutendijk started as a PhD candidate at the Control & Operations (C&O) department of the Aerospace Engineering Faculty of the TU Delft. His research was partly funded by the Airport Technology Lab (ATL) of the European Fund for Regional Development (EFRD). The research focused on developing and applying data-driven optimisation models to increase sustainability, efficiency and robustness of airside airport operations. His contributions have been published in journals and presented at several international conferences, leading to an Honourable Mention for best paper. In addition to the research, he was technical assistant for the course "Stochastic Processes and Simulation". He assisted in supervising four MSc students during their thesis projects and coached a Design Synthesis Exercise named "Air-Guard Bio-Inspired Morphing UAV as a Counter-Drone System".

List of Publications

Journal Publications

3. **M. Zoutendijk** and M.A. Mitici, *Fleet scheduling for electric towing of aircraft under limited energy capacity*, Energy, Special Issue "The Role of Smart Technologies in Energy Engineering", 294, p.130924 (2024).

2. **M. Zoutendijk**, M.A. Mitici and J.M. Hoekstra, *Overview of Operational Management Solutions and Challenges for Electric Taxiing of Aircraft*, Research in Transportation Business & Management, **49**, 101019 (2023).

1. **M. Zoutendijk** and M.A. Mitici, *Probabilistic flight delay predictions using machine learning and applications to the flight-to-gate assignment problem*, Aerospace, **8(6)**, 152 (2021).

Conference Proceedings

2. **M. Zoutendijk**, S. van Oosterom and M.A. Mitici *Electric Taxiing with Disruption Management: Assignment of Electric Towing Vehicles to Aircraft*, AIAA AVIATION 2023 Forum (2023), p.4219.

1. R. Hendrickx, **M. Zoutendijk**, M.A. Mitici and J. Schäfer, *Considering Airport Planners' Preferences and Imbalanced Datasets when Predicting Flight Delays and Cancellations*, IEEE/AIAA 40th Digital Avionics Systems Conference (2021).

Awards

1. *Honourable Mention Anna Valicek Competition, Airline Group of the International Federation of Operational Research Societies*, in 2022.