Electric Ground Support Equipment Operations at Airports

Fleet sizing under operational uncertainty: A KLM case study at Schiphol

T. de Niet

2025.MME.9118

Delft University of Technology

Electric Ground Support Equipment Operations at Airports

Fleet sizing under operational uncertainty: A KLM case study at Schiphol

by

T. de Niet

Graduation Research

in partial fulfilment of the requirements for the degree of

Master of Science
in Mechanical Engineering

at the Department Maritime and Transport Technology
within the Faculty of Mechanical Engineering
of Delft University of Technology

Student number:	4853857
MSc track:	Multi-Machine Engineering
Report number:	2025.MME.9118
Chair of the Committee & Supervisor:	Dr. B. Atasoy
Supervisor:	Dr. A. Napoleone
Independent Assessment Committee Member:	Dr. M. Ribeiro
Company Supervisor:	Ir. K. Timmermans
Date:	July 14, 2026

Abstract

Airports are increasingly electrifying Ground Support Equipment (GSE) fleets to reduce local emissions and support decarbonisation targets. Although electric Ground Support Equipment (eGSE) are well suited to many airside operations, electrification changes the fleet-sizing problem. Vehicle availability is no longer determined only by task duration, location, and travel time, but also by battery state, charging duration, charger access, and operational charging rules. These constraints are especially relevant during aircraft turnaround operations, where multiple time-critical ground-handling tasks must be completed within narrow service windows.

This thesis investigates how operational requirements, uncertainty, and charging affect the required fleet capacity and operational demand of eGSE in airport turnaround processes. A structured review of the GSE and eGSE modelling literature shows that existing studies provide important building blocks for routing, scheduling, energy management, and charging analysis. However, the joint integration of task execution, individual vehicle availability, charging behaviour, infrastructure constraints, and operational uncertainty remains limited. In particular, many models either simplify charging and vehicle-level states, treat fleet size as a fixed input, or do not represent uncertainty. This motivates the development of a simulation-based decision-support approach in which eGSE fleet sizing is evaluated as a dynamic vehicle-availability problem.

A rule-based Discrete-Event Simulation (DES) framework is developed to represent daily eGSE operations. The model includes individual service tasks, vehicle-specific states, airport location groups, travel times, battery state of charge, charging sessions, charger capacity, refilling, dumping, depot-return behaviour, and service-specific operating rules. Vehicles are assigned to tasks based on task urgency, time-window feasibility, travel time, battery state, and operational resource constraints. Task-timing uncertainty and travel-time uncertainty are included to evaluate how stochastic operational variability affects service-window performance and fleet-size requirements.

The framework is applied in a case study at KLM Ground Services (KLM GS) at Amsterdam Airport Schiphol (AAS). The validated case-study scope includes three vehicle groups: water vehicles, toilet vehicles, and loaders. These groups represent both depot-based resource-constrained operations and stand-side service operations. The model is verified using synthetic test cases and validated through operational data checks, expert judgement, service-demand validation, energy and charging behaviour, water-demand consistency, and uncertainty-representation checks. Simulation experiments are then used to assess baseline performance, deterministic fleet sizing, task-timing and travel-time uncertainty, combined uncertainty, demand-case robustness, idle forward staging, feasibility-aware task selection, and sensitivity to operational and charging parameters.

The results show that required fleet capacity depends strongly on the service-level interpretation. Under deterministic operating conditions, the smallest fleet sizes that achieve 100% on-time task completion are 12 water vehicles, 11 toilet vehicles, and 23 loaders. Under combined task-timing and travel-time uncertainty, the strict robust thresholds increase to 16 water vehicles, 15 toilet vehicles, and 25 loaders. These are the smallest tested configurations that achieve 100% on-time task completion across all 30 stochastic replications. However, if a very small number of short internal service-window violations is operationally acceptable, lower fleet sizes may also be defensible. Under this pragmatic interpretation, at least 11 water vehicles, 9 toilet vehicles, and 23 loaders achieve $\geq$ 99.90% on-time task completion. These results should be interpreted carefully, because the performance metric measures completion within internal service windows and does not directly measure aircraft departure delay.

The experiments also show that operational-control assumptions can materially affect fleet-size outcomes. Idle forward staging, where idle vehicles remain near the aircraft stand instead of returning immediately to the depot, reduces unnecessary deadheading and improves service performance for selected vehicle groups. For toilet vehicles, the deterministic strict requirement decreases from 11 to 9 vehicles. Under combined uncertainty, idle forward staging reduces the strict robust thresh-

old from 16 to 14 vehicles for water vehicles and from 15 to 10 vehicles for toilet vehicles. For loaders, the fleet-size threshold remains unchanged, but driven distance and depot-return movements are substantially reduced. A feasibility-aware task-selection variant further shows that dispatching logic can reduce the number of late tasks under scarce-fleet conditions, although it may increase the lateness severity of tasks that are already infeasible.

The sensitivity analysis indicates that charging is not the main driver of late task completion within the tested configurations. Energy-related parameters mainly affect charging sessions, charger occupancy, and charging-infrastructure utilisation. Task punctuality is more strongly constrained by vehicle availability during demand peaks, travel-time assumptions, service-time assumptions, and operational positioning logic. Charging therefore remains important for infrastructure planning and vehicle availability, but it is not the dominant bottleneck in the tested case-study settings.

Overall, the study shows that explicitly simulating operational requirements, uncertainty, and charging changes the interpretation of eGSE fleet sizing from a static vehicle-count problem into a dynamic availability problem. Required fleet capacity depends not only on the number of tasks, but also on when and where vehicles are needed, how uncertainty clusters demand, how quickly vehicles can recover between tasks, and how charging and supporting infrastructure affect vehicle availability. A simulation-based approach therefore provides a useful decision-support method for assessing eGSE fleet capacity, operational robustness, and charging-related resource use in airport ground handling. The reported fleet sizes should be interpreted as operational fleet-capacity requirements under the tested service-level assumptions. Final implementation decisions should add a technical reserve for maintenance, failures, battery degradation, charger unavailability, and other sources of vehicle downtime.

Contents

List of Abbreviations

Companies, Teams and Organisational Units

AFKL	Air France–KLM Group
AHSU KF	KLM Apron Operational Division AHSU KF
ARD KE	KLM Apron Operational Division ARD KE
AS	Apron Services
ATD KN	KLM Apron Operational Division ATD KN
BS	Baggage Services
GSE-Team	Ground Support Equipment Team
K1	KLM Apron Operational Division K1
K4	KLM Apron Operational Division K4
K5	KLM Apron Operational Division K5
KLM	KLM Royal Dutch Airlines
KLM GS	KLM Ground Services
OS	Operational Support
PS	Passenger Services

Airport Operations and Aviation Systems

AAF	Alternative Aviation Fuel
AAS	Amsterdam Airport Schiphol
APU	Auxiliary Power Unit
ASU	Air Starter Unit
COM	Commuter Aircraft
eGSE	electric Ground Support Equipment
ETV	Electric Towing Vehicle
GPU	Ground Power Unit
GSE	Ground Support Equipment
NABO	Narrow-Body Aircraft
PBB	Passenger Boarding Bridge
PCA	Pre-Conditioned Air Unit
PM	Particulate Matter
SAF	Sustainable Aviation Fuel
ULD	Unit Load Device
VOP	Vliegtuigopstelplaats (aircraft stand)

WIBO	Wide-Body Aircraft

Ground Handling Process Metrics

CTP	Critical Path
GH	Ground Handling
IB	In-Blocks
KPI	Key Performance Indicator
OB	Off-Blocks
TRT	Turnaround Time

Vehicles, Batteries and Control Systems

BEV	Battery Electric Vehicle
BMS	Battery Management System
EV	Electric Vehicle
LFP	Lithium Iron Phosphate
LiPo	Lithium-Polymer
NiCd	Nickel–Cadmium
NiMH	Nickel–Metal Hydride
NiZn	Nickel–Zinc
NMC	Lithium Nickel Manganese Cobalt Oxide
SOC	State of Charge
SOH	State of Health

Electrical Power and Charging Concepts

AC	Alternating Current
CC	Constant Current
CP	Constant Power
CV	Constant Voltage
DC	Direct Current
OCV	Open-Circuit Voltage
V1G	Unidirectional Charging
V2G	Bidirectional Charging/Vehicle-to-Grid

Modelling, Optimisation and Research Structure

ABM	Agent-Based Modelling
ALNS	Adaptive Large Neighborhood Search
DES	Discrete-Event Simulation
MILP	Mixed-Integer Linear Programming
MPC	Model Predictive Control
RQ	Research Question
SQ	Sub-question

Glossary

Average-day robustness check

Additional experiment based on a regular-demand operational day. It is used to assess whether conclusions from the baseline day remain reasonable under less congested operating conditions.

Baseline dispatcher

The original rule-based task-assignment logic used in the simulation model before operational-control variants, such as idle forward staging or feasibility-aware triage, are applied.

Combined uncertainty

Scenario in which task-timing uncertainty and travel-time uncertainty are applied simultaneously. It is the most comprehensive uncertainty setting considered in this study, but it does not include all possible operational disturbances.

Depot-return

Repeated movement of idle vehicles back to the depot after short idle periods, followed by new trips from the depot to the apron. This increases deadheading distance and can reduce the number of vehicles already positioned near future demand.

Deterministic baseline

Simulation setting without stochastic uncertainty, using the baseline fleet size and baseline operational-control logic. It provides the reference case against which fleet-size changes, uncertainty effects and operational-control variants are compared.

Deterministic minimum

The smallest fleet size that achieves 100% on-time task completion in the deterministic simulation setting. This threshold does not account for stochastic variation in task timing or travel time.

Feasibility-aware triage

Dispatching variant that prioritises tasks that are still feasible to complete within their service window. Tasks that are already infeasible are deprioritised.

Fleet-size requirement

The number of vehicles required to meet a specified service-level interpretation. In this study, the requirement can differ depending on whether the deterministic minimum, strict robust threshold or pragmatic lower bound is used.

Idle forward staging

Operational-control policy in which idle vehicles remain staged at or near their current apron location when sufficient battery state of charge and nearby future demand make this operationally reasonable, instead of automatically returning to the depot.

Internal service-window violation

A task completion that occurs after the modelled service-window deadline. This is an internal operational performance violation and should not be interpreted directly as an aircraft delay.

Late task

A task that is completed after its assigned service-window deadline. Late tasks are used to quantify service-level violations in the simulation output.

Lateness

The time difference between the actual task completion time and the service-window deadline for a late task, usually expressed in minutes.

On-time completion

The share of tasks completed within their assigned service window. This is the main service-level indicator used for fleet sizing and robustness assessment.

Baseline reference day

The high-demand operational day used as the primary basis for conservative fleet-sizing analysis. It is intended to stress the system under peak workload conditions.

Pragmatic lower bound

The smallest fleet size that achieves a near-perfect service level under stochastic simulation, while allowing a very small number of minor service-window violations. It distinguishes operationally negligible lateness from the stricter requirement of zero late tasks in all replications.

Refill/dump cycle

Operational process in which water vehicles refill potable water and toilet vehicles empty waste or prepare for subsequent service tasks. These processes affect vehicle availability and task sequencing.

Robustness

The ability of a fleet size or operational-control policy to maintain high service-level performance under uncertainty. In this study, robustness is assessed through repeated stochastic simulation runs.

Experiment-based robustness

Robustness assessment based on the uncertainty settings explicitly modelled in this study. It does not imply robustness against all possible real-world disruptions, such as breakdowns, weather effects or unmodelled service-duration variation.

Service-level interpretation

The rule used to translate simulation performance into a fleet-size requirement. Different interpretations, such as deterministic 100% on-time performance, strict stochastic robustness or a pragmatic lower bound, can lead to different fleet-size conclusions.

Service window

The time interval within which a task should be completed according to the modelled operational requirements. Service windows are used to evaluate whether tasks are completed on time.

Strict robust threshold/stochastic strict

The smallest fleet size that achieves 100% on-time completion across all stochastic replications (30) within a given uncertainty setting. This is the most conservative fleet-sizing interpretation used in this study. In this thesis, the term “strict robust” therefore refers to robustness within the sampled stochastic experiment and should not be interpreted as a formal guarantee over all possible operational realisations.

Stochastic fleet sizing

Fleet-sizing analysis based on multiple simulation replications with uncertainty applied to task timing, travel time or both. It is used to assess how fleet-size requirements change when operational variability is included.

Task-timing uncertainty

Uncertainty setting in which task start times or service windows are shifted to represent early or delayed task occurrence, while the task duration itself remains unchanged.

Travel-time uncertainty

Uncertainty setting in which selected travel times are inflated to represent operational variability in apron movements, routing or congestion.

1
Introduction

The chapter is structured as follows. Section 1.1 provides the environmental and operational context and quantifies the role of airports and Ground Support Equipment (GSE) in emissions and local air quality impacts. Section 1.2 formulates the core operational challenges introduced by electrification and motivates the need for modelling and decision support. Based on this context, Section 1.3 formulates the central Research Question (RQ) and associated Sub-questions (SQs), and Section 1.4 outlines the simulation-based methodology used to answer them.

1.1. Introduction

The global urgency to mitigate climate change and reduce greenhouse gas emissions has increased pressure on the transport sector to transition towards more sustainable and energy-efficient operations. In 2024, transport was responsible for approximately 15% of total global carbon emissions [4]. Within the transport sector, aviation accounts for about 13.33% of transport-related CO_2 emissions, corresponding to approximately 2% of total global carbon emissions, making it the second-largest transport emitter after road transport [4]. Airports also contribute to the aviation-sector emission footprint, with airport operations accounting for about 15% of aviation-related CO_2 emissions [29]. Within airports, GSE and passenger transport systems account for up to 24% of airport-specific emissions, including 15% of airport CO_2 emissions and 5–9% of NO_x emissions [16, 84]. A clear overview of the CO_2 emission flow is shown in Figure 1.1, based on sources from 2023 to 2025.

The relevance of GSE becomes particularly clear in apron operations. In these areas, GSE generate approximately 63% of NO_x, 75% of Particulate Matter (PM), and 24% of total fuel consumption [84]. These emissions are relevant not only from a climate perspective, but also because they contribute to local air-quality impacts for airport workers and surrounding communities. Air pollution associated with airport activity has been linked to an estimated 16,000 premature deaths worldwide each year [12].

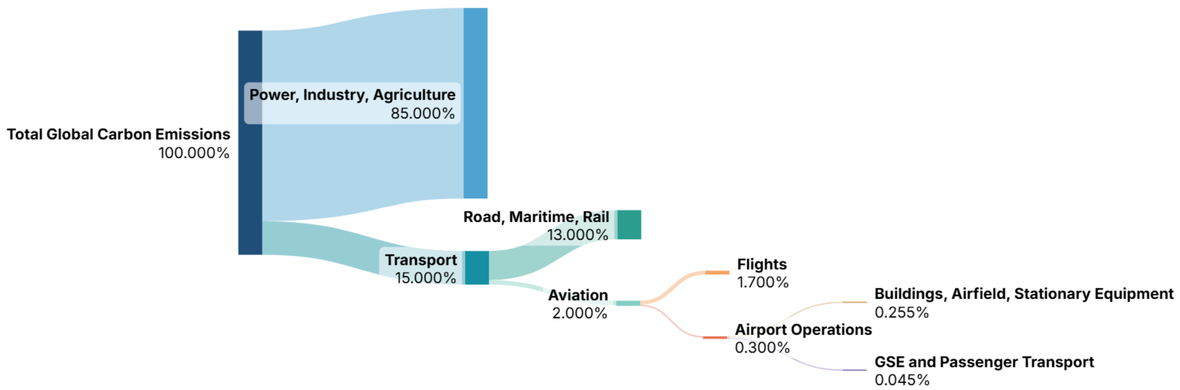


Figure 1.1: *CO_2 emission flow from global sources to airport GSE [4, 9, 16, 84].*

At the same time, global air traffic continues to grow at an average rate of 5% per year [53]. This growth increases the number of flights, turnaround handling operations, taxi movements, operational costs, and potential delays that can propagate through the wider aviation network. Together, these developments place increasing pressure on airport Ground Handling (GH) systems while also increasing the importance of reducing emissions from ground operations [14].

To address these environmental and operational challenges, airports are accelerating the electrification of their GH fleets. In the Netherlands, Amsterdam Airport Schiphol (AAS) has committed to operating emission-free GSE by 2030 [67]. Replacing fossil-fuelled GSE with electric alternatives is therefore an important measure for meeting airport decarbonisation targets. Airports are generally suitable environments for electric Ground Support Equipment (eGSE) deployment because many ground-support operations involve short driving distances, flat terrain, and frequent stop-and-go cycles. These characteristics make electrification comparatively attractive for several GSE types [18]. However, while electric alternatives are already available for smaller GSE, larger variants, such as 45 m^3 refuellers, are still in the development and testing phase [19, 80].

1.2. Research Motivation

The shift towards electrification also introduces new operational challenges. GSE are critical assets for maintaining operational continuity in airport ground handling [17]. However, when these fleets are electrified, vehicle availability becomes dependent not only on task execution and vehicle location, but also on battery state, charging duration, charger access, and vehicle-specific energy performance. Compared with conventional GSE, eGSE have longer recharging times, limited driving ranges, and stronger differences in operational feasibility across vehicle types. These characteristics increase the complexity of dispatching and scheduling decisions. Infrastructure constraints, grid peak loads, and charging-related vehicle downtime can create additional bottlenecks, requiring trade-offs between maintaining vehicle availability, selecting suitable charging moments, and protecting battery health [30, 46, 22].

These challenges are especially relevant in high-traffic hubs such as AAS, where turnaround handling operations are highly time-sensitive and reliability-driven. Multiple GSE types must perform interdependent tasks within narrow service windows and under strict safety protocols [27]. Under such conditions, a fleet that is sufficient under deterministic assumptions may become insufficient when operational variability is included. Delayed task timing, longer travel times, or charger occupancy can reduce the number of vehicles available at the moment and location where demand occurs. Therefore, eGSE fleet sizing should not be treated only as a static vehicle-count problem, but as a dynamic vehicle-availability problem.

As airports move towards full electrification, decision support is needed to evaluate how fleet size, dispatching logic, charging behaviour, infrastructure constraints, and uncertainty interact during daily operations. Models used for this purpose must be able to represent individual vehicle states, time-window feasibility, charging-related unavailability, and stochastic operational disturbances. Simulation-based approaches are particularly suitable for this purpose because they can represent the temporal evolution of service demand, vehicle availability, and infrastructure use under different operational assumptions [49, 44]. This motivates the development of a simulation framework that can assess eGSE fleet-capacity requirements under operational requirements, uncertainty, and charging constraints.

1.3. Research Question (RQ)

What is the impact on the required fleet capacity and operational demand of electric Ground Support Equipment (eGSE) when simulating operational requirements, uncertainty and charging in turnaround processes at airports?

In this research, fleet capacity refers to the number of eGSE vehicles required to execute the relevant ground-handling tasks within the defined service-level requirements. Operational demand refers to the time- and location-specific demand for vehicle availability, task execution, charging access, and supporting infrastructure. This interpretation is used throughout the thesis when evaluating how electrification affects eGSE operations. In order to answer this question, the following SQs are formulated:

SQ 1. Which GSE types exist and what are the functions and duty cycles of these types?

SQ 2. What input requirements and data sources are necessary to model an eGSE fleet while capturing operational requirements, uncertainty, and charging processes?

SQ 3. How can airport eGSE operations be represented in a simulation model using appropriate charging control rules, decision rules and parameters?

SQ 4. How can task-timing and travel-time uncertainty in eGSE operations be incorporated into a simulation model?

SQ 5. Which KPIs best capture the impact of eGSE implementation on operational performance and resource requirements?

SQ 6. How can the required eGSE fleet capacity be determined from simulation outcomes, and how do demand conditions, operational positioning, and task-selection rules affect service performance?

1.4. Methodology

This research follows a simulation-based methodology to analyse how eGSE operations at airports can be represented and evaluated under operational uncertainty and charging constraints. The methodology combines operational analysis, uncertainty modelling, literature review, discrete-event simulation, case-study instantiation, verification, validation, and simulation experimentation. The objective is to develop and apply a decision-support framework that can assess eGSE fleet-capacity requirements under realistic operational constraints.

The research is structured as a connected methodological chain. First, the operational context is analysed to identify the turnaround processes, GSE functions, service windows, aircraft stands, vehicle movements, fleet heterogeneity, and duty-cycle differences that must be represented in the model. This provides the operational basis for the simulation framework and supports SQ1.

Second, the electrification context is analysed to determine how battery-electric operation changes the fleet-sizing problem. Battery capacity, State of Charge (SOC), charging duration, charger availability, charging thresholds, and charging-related vehicle downtime are treated as operational constraints that affect vehicle availability. This defines the electrification-specific model requirements.

Third, uncertainty modelling is developed as a methodological foundation. The research distinguishes between aleatory uncertainty, representing inherent operational variability, and epistemic uncertainty, representing imperfect knowledge about data, assumptions, or model structure. Task-timing uncertainty and travel-time uncertainty are selected as active stochastic inputs because they directly affect service-demand timing and vehicle availability. Other uncertainties, such as service-duration variation, energy consumption, charging behaviour, infrastructure availability, battery degradation, weather disruption, and operator behaviour, are addressed through assumptions, validation, sensitivity analysis, scope limitations, or future research. This part supports SQ4.

Fourth, a targeted literature review is conducted to assess how existing airport GSE and eGSE studies model fleet sizing, routing, task scheduling, uncertainty, charging integration, and energy management. The review compares studies in terms of modelling objective, model type, time representation, uncertainty treatment, fleet representation, operational detail, and charging integration. The identified research gap motivates the development of an integrated simulation framework that combines task execution, individual vehicle states, charging behaviour, infrastructure constraints, and operational uncertainty.

Fifth, a rule-based Discrete-Event Simulation (DES) model is developed. The model represents service tasks, individual eGSE vehicles, and supporting infrastructure such as chargers, refill points, dump points, depots, and service locations. Tasks are defined by service type, aircraft stand, service duration, and service time window. Vehicles are modelled individually and are characterised by location, SOC, onboard resource level, and operational status. The simulation includes task generation, dispatching, task selection, vehicle selection, task execution, charging, depot return, refilling, dumping, and resource replenishment. This part answers SQ2 and SQ3.

Sixth, the generic framework is instantiated for the KLM Schiphol case study. Case-study data define the flight schedule, aircraft stands, service rules, fleet characteristics, infrastructure locations, charging assumptions, resource constraints, operational policies, and uncertainty profiles. The validated case-study scope includes water vehicles, toilet vehicles, and loaders, because these vehicle groups cover both depot-based resource-constrained operations and stand-side service operations.

Seventh, the model is verified and validated before it is used for experimentation. Verification tests whether the internal simulation logic behaves as intended, including task generation, dispatching, vehicle assignment, task execution, charging, refilling, dumping, depot return, queueing, stop-cutoff handling, and uncertainty mechanisms. Validation assesses whether the model outputs are plausible for the KLM Schiphol case using operational data, expert judgement, service-demand checks, energy and charging behaviour, water-demand consistency, and uncertainty-representation checks. The validation does not make the model a fully calibrated digital twin, but establishes whether it is credible enough for system-level fleet-capacity analysis.

Eighth, simulation experiments are conducted with the verified and validated model. These experiments evaluate baseline performance, deterministic fleet sizing, task-timing uncertainty, travel-time uncertainty, combined uncertainty, demand-case robustness, idle forward staging, feasibility-aware task selection, and sensitivity to selected operational and charging parameters. The main performance metric is on-time task completion, supported by vehicle utilisation, driven distance, charging sessions, charger occupancy, and late-task characteristics. Fleet-capacity requirements are then interpreted under both strict and pragmatic service-level definitions, supporting SQ5, SQ6 and the central RQ.

The experimental design follows a robust fleet-sizing perspective within the tested model configurations. The main experiments are based on a high-demand baseline day, which is used as a conservative stress case for fleet-capacity planning. Three complementary analyses are added. First, an average-day water-vehicle experiment is used as a demand-case robustness check by replacing the full operational input day. Second, idle forward staging is tested as an operational positioning policy to assess whether reducing unnecessary depot-return movements lowers mileage, improves service performance, or reduces fleet-size thresholds. Third, feasibility-aware triage is tested as a task-selection variant to examine how dispatching logic affects the number and severity of late tasks under scarce-fleet conditions. Finally, parameter sensitivity tests assess how selected operational and energy-related assumptions influence the results.

The methodological logic is summarised in Table 1.1. The table links each sub-question to its main output, method, and location in the thesis.

Table 1.1: *Overview of sub-questions, main outputs, methods and thesis chapters*

SQ	Main output	Method	Chapter
SQ1	GSE typology and duty-cycle logic	Operational analysis and literature review	Chapter 2
SQ2	Model input requirements and data sources	Literature review, operational analysis, and case-study data	Chapter 2, Chapter 3, Chapter 4, Chapter 5
SQ3	Rule-based DES framework	Literature-based requirements and simulation model development	Chapter 3, Chapter 4
SQ4	Uncertainty representation for task timing and travel time	Literature review, empirical derivation, and stochastic simulation design	Chapter 2, Chapter 3, Chapter 4, Chapter 5, Chapter 6
SQ5	Performance-metric and validation framework	Literature review, operational data checks, expert judgement, and model validation	Chapter 5, Chapter 6
SQ6	Fleet-sizing outcomes and operational-control insights	Deterministic and stochastic simulation experiments	Chapter 6, Chapter 7, Chapter 8

1.5. Scope and Contribution

This thesis develops and applies a simulation-based approach to assess the operational fleet-capacity requirements of eGSE under time-window, vehicle-availability, charging, and uncertainty constraints. The scope is deliberately operational. The model evaluates whether ground-handling service tasks can be completed within internal service windows, given a demand schedule, a spatial airport representation, vehicle-specific service rules, and capacity-constrained supporting infrastructure. The resulting fleet sizes therefore represent operational fleet-capacity requirements under the tested assumptions, not final procurement numbers. Implementation decisions would still require an additional technical reserve for maintenance, failures, battery degradation, charger downtime, and other sources of vehicle unavailability.

The case-study application is limited to three vehicle groups at KLM Ground Services at AAS: water vehicles, toilet vehicles, and loaders. These groups were selected because they represent two relevant operational behaviours within one modelling structure. Water and toilet vehicles represent depot-based, resource-constrained operations with refill or dump requirements, while loaders represent stand-side service equipment without onboard service-resource constraints. Although the broader model structure can be extended to additional GSE types, the numerical fleet-sizing conclusions in this thesis are validated only for these three vehicle groups.

The model boundary is the execution of airside service tasks at aircraft stands. The model includes task generation from operational input data, vehicle dispatching, travel between airport location groups, task execution, charging, depot return, refill and dump processes, and queueing at shared infrastructure. The demand schedule is treated as an input to the model. The model does not optimise flight schedules, stand allocation, turnaround sequencing, or airline planning decisions. It also does not model the full aircraft turnaround as an integrated precedence network. Consequently, a late task in this thesis is an internal service-window violation and should not be interpreted directly as an aircraft departure delay. The model further abstracts from detailed operator behaviour, weather disruption, vehicle breakdowns, battery degradation, charger failures, and grid-level power constraints. Charging is represented as an operational availability constraint within the simulation horizon, not as a grid-optimisation, electricity-market, or battery-health optimisation problem.

Within this scope, the thesis contributes to the modelling and operational planning of eGSE in six ways. First, it develops a rule-based DES framework that represents individual vehicles, task time windows, vehicle locations, SOC, charging behaviour, onboard resource constraints, refill and dump processes, depot-return behaviour, and capacity-constrained supporting infrastructure. This allows fleet sizing to be evaluated as a dynamic vehicle-availability problem rather than as a static vehicle-to-task ratio. The framework also shows how different operational vehicle behaviours, such as depot-based resource-constrained operations and stand-side service operations, can be represented within one simulation structure.

Second, the thesis links fleet-sizing outcomes to explicit service-level interpretations. Deterministic fleet sizing is used to identify nominal sufficiency under fixed operating conditions. Stochastic fleet sizing is used to assess performance under task-timing and travel-time uncertainty. A pragmatic lower-bound interpretation is added to distinguish near-perfect operational performance from the stricter requirement of zero late tasks across all sampled stochastic replications. This distinction is important because different service-level definitions can lead to different fleet-size conclusions.

Third, the thesis applies the framework to a high-demand operational case at KLM Schiphol and quantifies how operational uncertainty changes the interpretation of fleet-capacity requirements. The contribution is therefore not only the calculation of fleet sizes, but also the demonstration of how uncertainty, demand timing, spatial movement, service rules, and charging constraints jointly affect operational vehicle availability. In this way, the thesis translates fleet sizing from a static capacity question into an operational availability problem.

Fourth, the thesis provides operational insight into the use of vehicles and supporting infrastructure. Besides on-time task completion, the model evaluates indicators such as vehicle utilisation, driven distance, charging sessions, charger occupancy, charging-related waiting, refill and dump activity, depot-return movements, and infrastructure use. These outputs help explain why a fleet size performs sufficiently or insufficiently, and whether bottlenecks are primarily caused by vehicle capacity, spatial repositioning, charging behaviour, or supporting infrastructure. The model therefore supports not only fleet-capacity assessment, but also the evaluation of operational resource use and potential infrastructure constraints under the tested configurations.

Fifth, the thesis introduces and evaluates idle forward staging as an operational positioning policy for eGSE. Instead of sending idle vehicles back to a central depot after short idle periods, vehicles can remain staged near their current Vliegtuigopstelplaats (aircraft stand) (VOP) when their battery state and expected future demand make this operationally feasible. This policy is used to assess whether part of the fleet-capacity requirement is caused by avoidable depot-return movements rather than by unavoidable service demand.

Sixth, the thesis uses additional configurations and control-rule experiments to test the sensitivity of fleet-sizing conclusions to alternative operating assumptions. The average-day water-vehicle experiment is interpreted as a demand-case robustness check, because it replaces the complete operational input day rather than varying one isolated parameter. The feasibility-aware triage variant is interpreted as a diagnostic task-selection variant, because it changes the ordering of waiting tasks and reveals a trade-off between reducing the number of late tasks and increasing the lateness severity of tasks that are already infeasible. These experiments show that fleet-sizing conclusions depend not only on demand volume and vehicle count, but also on the operational rules used to position vehicles and select tasks.

2
Operational Foundations for eGSE Fleet Sizing

This chapter establishes the operational foundations required to model eGSE fleet sizing at airports. It first introduces airport ground operations and the aircraft turnaround process, because these determine when and where GSE tasks arise and which operational constraints shape service execution. It then discusses the electrification of GSE, focusing on battery state, charging, and infrastructure constraints as determinants of vehicle availability. Finally, the chapter addresses uncertainty in GSE operations and explains why task timing and travel-time variability are relevant for fleet-sizing analysis.

The chapter therefore defines the modelling requirements that are later evaluated against the literature in Chapter 3 and translated into the simulation framework in Chapter 4. Its purpose is not to provide a complete overview of airport operations or battery technology, but to identify the operational, electrification-related, and uncertainty-related elements that an eGSE fleet-sizing model must represent.

2.1. Airport Operations

This section introduces the airport ground-operation context in which GSE fleets operate. It first describes airports as integrated landside–airside systems and then focuses on the aircraft turnaround process, because turnaround timing, stand configuration, aircraft interfaces, and operational procedures determine the demand for GSE services. The section also introduces the main GSE categories and characterises the operational differences between vehicle types that are relevant for fleet-sizing simulation.

2.1.1. Airports as a System

Airports form the core infrastructure of passenger aviation. They enable aircraft to take off and land and provide the facilities required to process passengers, baggage, cargo, and aircraft services. Figure 2.1 illustrates a generic airport layout. Airport operations are commonly divided into landside and airside processes. Landside processes include passenger arrival, check-in, baggage drop-off, and security screening. Airside processes include passenger boarding and disembarkation, aircraft taxiing, take-off and landing, and the ground-handling activities required while the aircraft is parked.

Within this system, the aircraft turnaround is the key operational interface between airport infrastructure and aircraft operations. It connects aircraft arrival, ground handling, passenger and baggage processes, and departure preparation [74]. This makes the turnaround process central to the operational demand for GSE, because it determines which services are required, where they must be performed, and within which time windows they need to be completed.

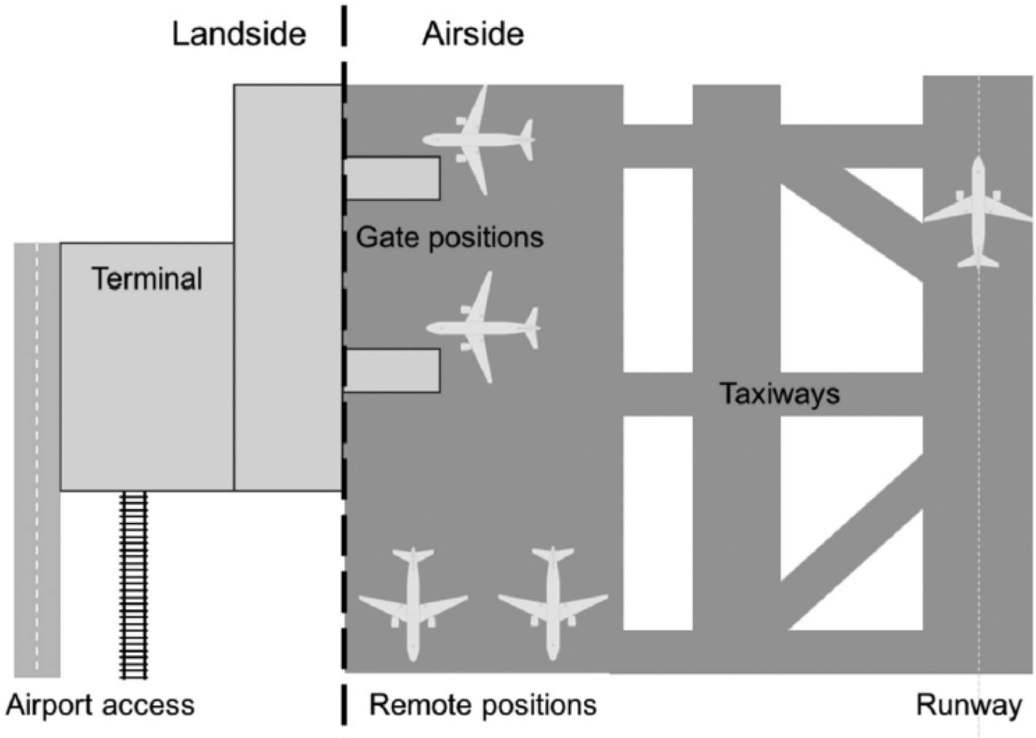


Figure 2.1: *Generic airport with landside and airside [74].*

GH procedures are usually divided into terminal and ramp activities. Terminal activities are performed inside the terminal buildings and mainly concern passenger services. Ramp activities take place at the aircraft parking position between In-Blocks (IB) and Off-Blocks (OB). IB refers to the moment an aircraft is secured with wheel chocks, marking the start of ramp handling activities. OB refers to the moment the wheel chocks are removed, marking the end of the ground-handling process. A turnaround can take place either at a terminal gate or at a remote apron position, as shown in Figure 2.1 [56].

2.1.2. Turnaround Process

Figure 2.2 illustrates the aircraft turnaround process, which begins when the aircraft arrives at the parking position and ends when it departs. The process starts at the IB moment, when the aircraft is parked and secured. The parking position can be located either at a terminal gate or at a remote apron position. Once the aircraft is in position, multiple ground-handling activities can start, often in parallel and under tight time constraints.

During the ground-service phase, different GSE types support different handling activities. Baggage vehicles load and unload baggage, while baggage tractors transport baggage between the aircraft stand and the baggage hall. Passengers disembark and board using passenger bridges, stairs, or buses, depending on the stand configuration. Catering trucks deliver and remove catering supplies, cleaning teams prepare the cabin, refuelling vehicles provide fuel, potable-water trucks refill the aircraft freshwater system, and toilet-service vehicles empty the lavatory system. Some of these activities follow a fixed sequence, while others can be performed in parallel or in interchangeable order [74]. This process structure is important for GSE planning because it creates simultaneous demand for several vehicle types within a limited turnaround window.

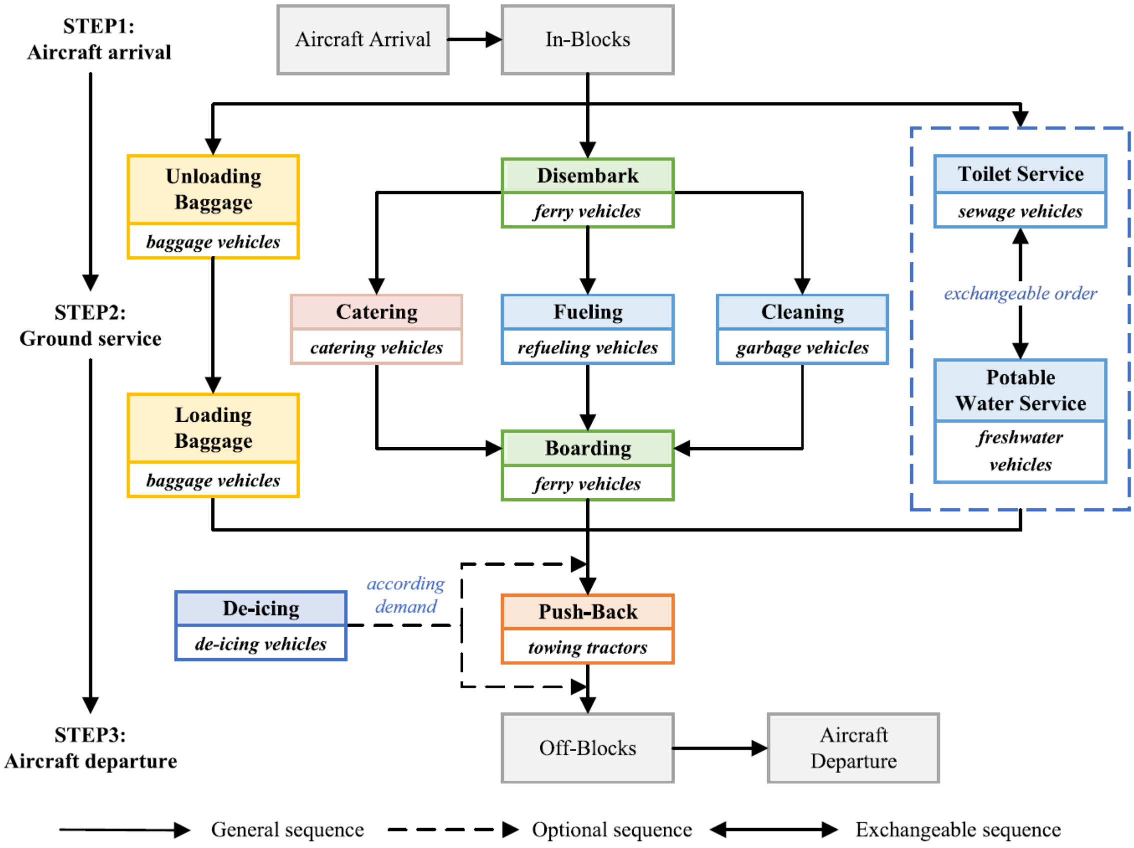


Figure 2.2: *Schematic overview of the turnaround cycle of an aircraft [92].*

Aircraft are equipped with multiple service interfaces, such as doors, hatches, water connections, waste connections, power interfaces, and cargo doors. The position of these interfaces depends on aircraft type and configuration. A cluster analysis of Commuter Aircraft (COM) and Narrow-Body Aircraft (NABO) aircraft shows a trend towards commonality, for example with drinking-water and waste-water connections commonly located near the rear fuselage, electrical power interfaces generally located near the forward fuselage, and cargo doors usually positioned on the right side of the aircraft [74]. Figure 2.3 shows these interface locations.

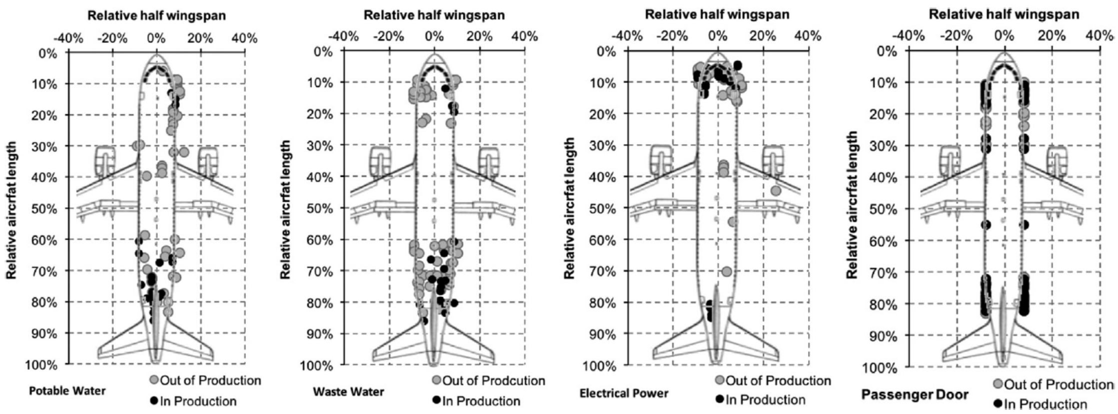
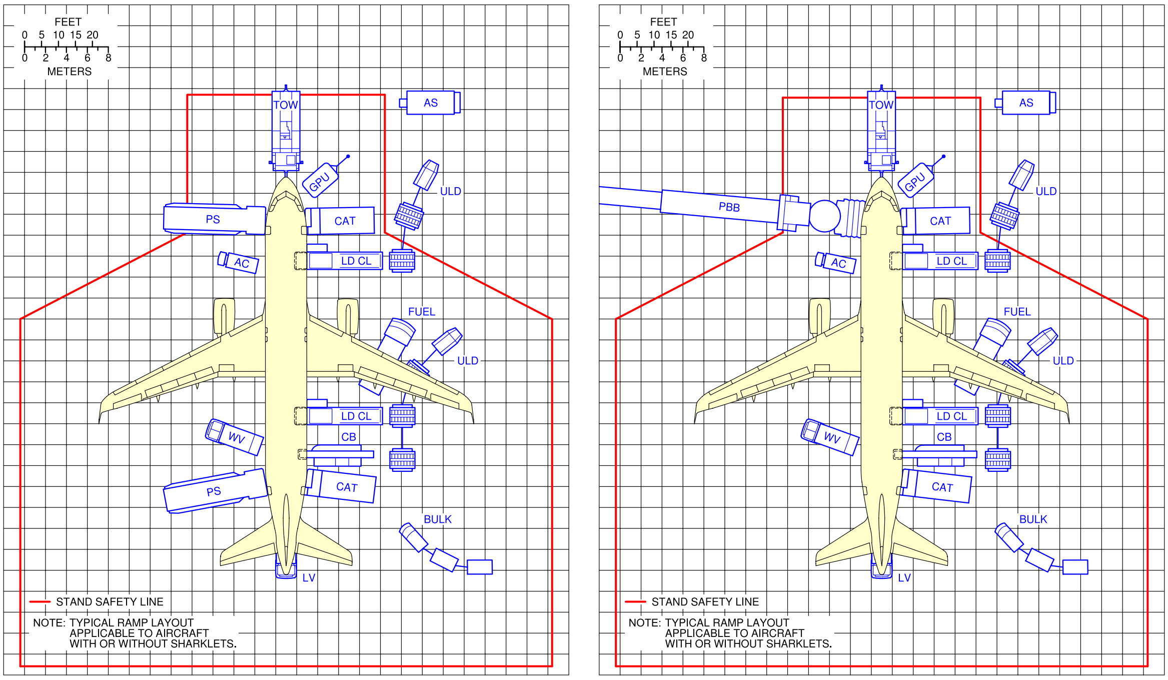


Figure 2.3: *Aircraft interface locations [74].*

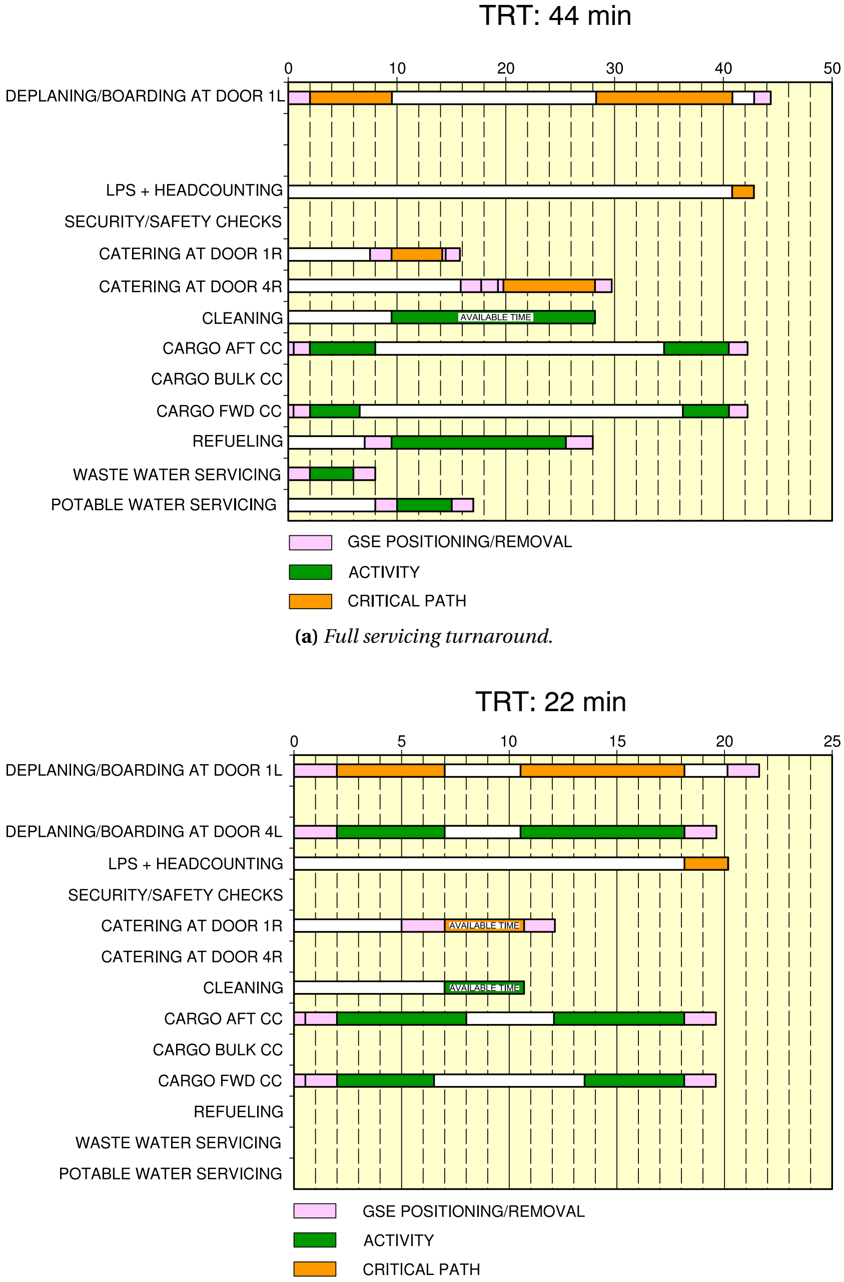
These fixed aircraft-interface locations are relevant for GSE modelling because they constrain where service vehicles can operate around the aircraft. They influence stand-side movement, vehicle positioning, service feasibility, and the interaction between simultaneous handling activities. Figure 2.4 illustrates this spatial dependency for a short-to-medium-haul aircraft. Passenger activities are generally concentrated on the left side of the aircraft, while catering and cargo handling typically take place on the right side. As a result, service-vehicle positions are partly predefined by aircraft-interface locations and stand layout [74].



(a) *Servicing arrangement with two passenger stairs.* **(b)** *Servicing arrangement with a passenger bridge.*

Figure 2.4: *Typical ramp layout for an Airbus A320neo [5].*

The duration of the turnaround process depends on aircraft characteristics, cargo volume, required service scope, stand position, and airline or airport procedures. Some activities are constrained by precedence requirements due to safety, space limitations, or operational policy [56]. The corresponding turnaround Gantt chart in Figure 2.5 shows how handling activities are sequenced and where parallel execution is possible. Activities that determine the minimum required Turnaround Time (TRT) form the Critical Path (CTP). In many cases, the CTP consists of passenger and cabin-related activities, although refuelling or baggage handling can become critical under specific operational conditions [74, 24].



(a) *Full servicing turnaround.*

(b) *Outstation turnaround.*

Figure 2.5: *Typical TRT chart for an Airbus A320neo [5].*

For the Airbus A320neo, the average full-service TRT is approximately 44 minutes [6], while the required outstation turnaround time is 22 minutes [5]. In practice, however, TRT is stochastic because passenger numbers, fuel uplift, cargo loads, stand conditions, and operational procedures vary between flights. Airlines therefore use buffer times, and actual ground-handling schedules can differ substantially from manufacturer reference timelines [74]. Not all services are required on every flight; for example, water and lavatory services are often omitted for short-haul European outstation flights [24].

Airport ground handling is also under increasing pressure due to air-traffic growth. Inefficient GSE scheduling has been identified as the second most common cause of flight delays, after air traffic control issues. Ground-handling disruptions can propagate through aircraft rotations and the wider network. Around 45% of the delays in one year at European airports are reactionary delays caused by late arrival of aircraft or crew from another flight. 15% of the delays were caused by GH disruptions, including aircraft and ramp handling (5%), passenger and baggage handling (6%), and flight operation and crewing (4%) [92, 74]. For this thesis, the key implication is that GSE operations must be represented as time-window-constrained service tasks. The fleet-sizing problem is therefore not only determined by the number of tasks, but also by task timing, stand location, service duration, travel time, vehicle availability, and operational variability.

2.1.3. Ground Support Equipment

Ground Support Equipment (GSE) refers to the vehicles and equipment used to support aircraft while they are on the ground. These assets enable the execution of turnaround activities between IB and OB, including aircraft servicing, passenger handling, baggage and cargo handling, aircraft movement, and the provision of external support functions such as power, air, water, and waste handling.

GSE can be classified according to the operational function they perform. Aircraft support equipment includes units such as GPUs, ASUs, and PCAs, which provide external power, air supply, or climate support to the aircraft. Aircraft movement equipment includes pushback tractors and towing vehicles. Aircraft servicing equipment includes potable-water trucks, toilet-service trucks, refuelling vehicles, catering trucks, cleaning equipment, and de-icing vehicles. Passenger handling equipment includes passenger stairs, buses, and passenger boarding bridges, while baggage and cargo handling equipment includes belt loaders, loaders, transporters, baggage tractors, and dollies.

These equipment groups differ substantially in their duty cycles, service locations, mobility patterns, and operational constraints. Some vehicles mainly operate at or near a fixed aircraft stand, while others repeatedly travel between depots, service points, and aircraft positions. Some GSE types have onboard resource constraints, such as water or waste capacity, while others are mainly constrained by vehicle availability, travel time, service duration, or energy state. These differences are relevant for simulation modelling because they determine which vehicle states, resource constraints, movement patterns, and service rules must be represented for each equipment type.

Electrification further increases the importance of these operational differences. Once GSE are electrified, vehicle feasibility no longer depends only on whether a vehicle is idle and compatible with a task, but also on whether it has sufficient battery capacity, access to charging infrastructure, and enough time to recover between tasks. As a result, different GSE types may face different electrification constraints depending on their utilisation pattern, travel behaviour, service duration, and charging opportunities.

The full range of GSE types is not discussed in detail in the main text, because this chapter aims to introduce the operational role of GSE rather than provide a complete equipment catalogue. A detailed overview of the main GSE categories and examples is therefore provided in Appendix A [78]. The remainder of the thesis focuses on the modelling-relevant characteristics of GSE: task timing, service duration, vehicle movement, energy use, charging behaviour, and resource constraints. The case-study implementation later concentrates on water vehicles, toilet vehicles, and loaders, because these vehicle groups represent both depot-based resource-constrained operations and stand-side service operations within the same modelling framework.

2.1.4. Characterisation of GSE Operations

Determining the required GSE types, fleet size, and service durations depends on operational factors related to the aircraft, the aircraft stand, and the procedures defined by the airport and airline. Figure 2.6 summarises these influences and shows how they jointly shape the type, number, and operation of GSE during the turnaround process.

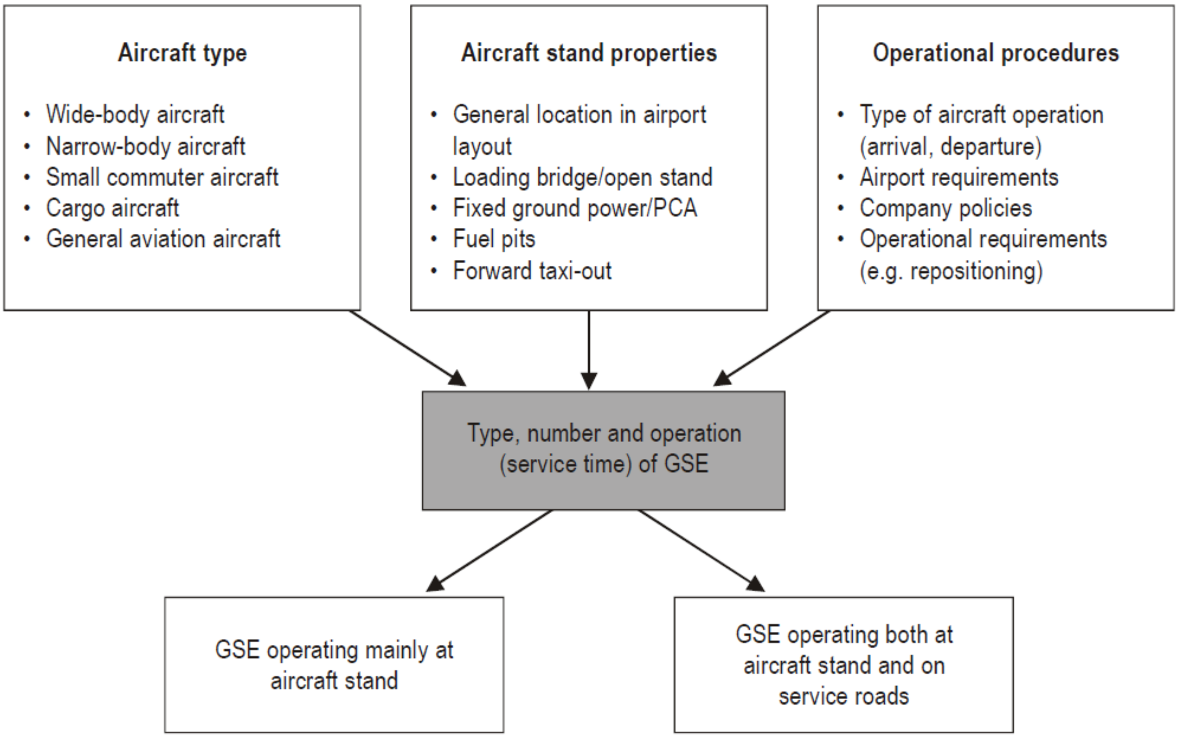


Figure 2.6: *Parameters influencing GSE operations [38].*

Three groups of characteristics influence GSE demand: aircraft characteristics, stand characteristics, and operational procedures. First, aircraft characteristics determine the required service scope. Wide-Body Aircraft (WIBO), NABO, COM, cargo aircraft, and general aviation aircraft differ in size, capacity, service interfaces, and handling requirements. Larger aircraft typically require more equipment, longer service durations, and higher resource volumes, while smaller aircraft generally require fewer resources [85, 49, 19, 38].

Second, aircraft stand characteristics influence which GSE can be used and how these vehicles operate. Relevant factors include the stand location within the airport layout, whether the aircraft is parked at a terminal gate or open stand, the availability of a Passenger Boarding Bridge (PBB), fixed ground power, pre-conditioned air, fuel pits, and local procedures such as forward taxi-out. These factors affect both the need for specific GSE types and the spatial feasibility of vehicle movements around the aircraft [38].

Third, operational procedures determine when and how GSE must be deployed. These procedures include arrival, departure, and turnaround logic, airport regulations, airline-specific service standards, company policies, and repositioning requirements. They define whether a service is required, when it should be performed, and which operational constraints apply [38].

Together, these characteristics determine the required GSE configuration, including vehicle type, number of vehicles, service duration, and operating pattern. For simulation modelling, this means that GSE demand cannot be represented only as a total number of tasks. It must also be represented through task timing, stand location, service duration, vehicle compatibility, movement requirements, and service-specific constraints.

The literature also presents several ways to characterise and classify GSE operations, reflecting differences in research objectives, modelling approaches, and levels of operational detail. Three general perspectives are relevant for this thesis: functional categories, energy- or power-based groupings, and operational working modes. Examples of these classification perspectives are shown in Figure 2.7 [17, 30, 89].

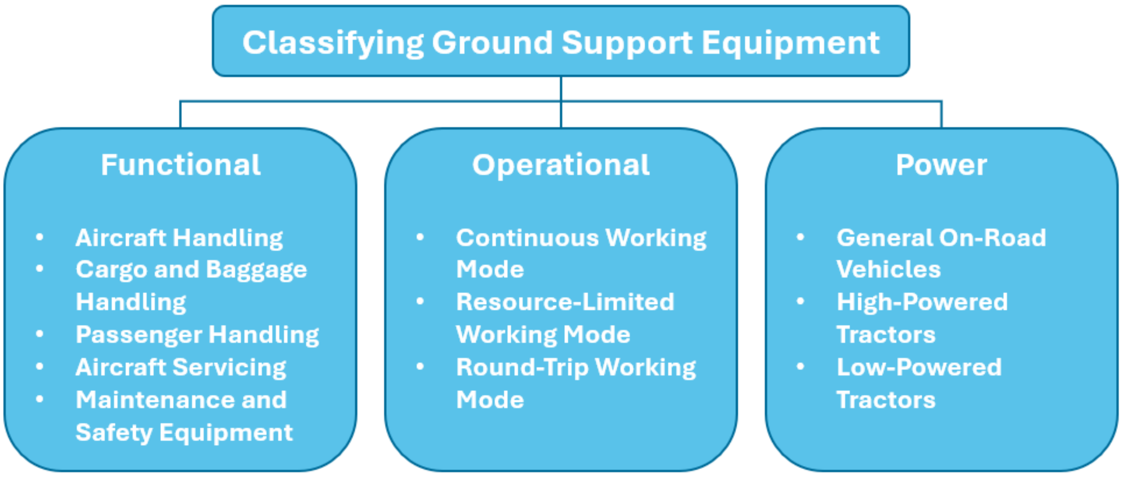


Figure 2.7: *Potential classifications of GSE [17, 30, 89].*

A functional perspective classifies GSE according to the service provided to the aircraft. Bose distinguishes aircraft handling, cargo and baggage handling, passenger handling, aircraft servicing, and maintenance and safety equipment [17]. Other studies apply a similar service-based classification with different levels of detail. Zhou reviews nine GSE types and groups them into seven categories based on operational characteristics and scheduling methods: refuelling and freshwater vehicles, passenger buses, baggage vehicles, waste-handling vehicles, catering vehicles, de-icing vehicles, and towing tractors [92]. Padrón classifies the turnaround process through the tasks themselves, including unloading and loading baggage, deboarding, boarding, catering, cleaning, fuelling, potable-water service, and toilet service [56]. Fu extends this perspective towards electrification by identifying corresponding Electric Vehicle (EV) types, such as electric passenger stairs, electric airport shuttle buses, electric refuelling trucks, electric catering trucks, electric water-service vehicles, electric sewage trucks, electric baggage tractors, electric baggage conveyor vehicles, and electric aircraft towing vehicles [27]. For this thesis, the functional perspective is relevant because it defines the service types and task categories that must be represented in the simulation model.

A second perspective classifies GSE according to energy use or power demand. This perspective has become more important as airports electrify their ground-handling fleets. Bose classifies selected GSE types into groups based on energy-consumption profiles, including pushback tractors, loaders, tug tractors, and forklifts, and shows how these profiles influence charging-infrastructure needs during large-scale electrification [18]. Gulan adopts a similar approach by organising vehicles into three classes based on required power levels: general on-road vehicles, high-powered pushback tractors, and smaller tug tractors for cargo handling [30]. These classifications are useful for analysing charging demand, charger capacity, and potential grid impacts. For this thesis, this perspective is relevant because electrification introduces battery state, charging duration, and infrastructure availability as operational constraints.

A third perspective classifies GSE according to operational working mode. This perspective describes how vehicles move between depots, aircraft stands, and service points. Yao identifies three working modes. In the continuous working mode, vehicles leave the depot, serve an aircraft, and then proceed directly to the next task without returning to a support point. Examples include towing tractors, passenger stairs, and baggage belt loaders. In the resource-limited working mode, vehicles can only serve a limited number of flights before returning to a supply point to replenish onboard resources, such as water or catering goods. Examples include potable-water trucks and catering trucks. In the round-trip working mode, vehicles repeatedly travel between a service location, such as a terminal or baggage area, and the aircraft stand. Examples include passenger buses and baggage tractors [89]. This perspective is particularly relevant for fleet-sizing simulation, because working modes determine vehicle availability, travel requirements, resource constraints, and the need for intermediate refill, dump, or depot visits.

Overall, GSE operations can be characterised from multiple complementary perspectives. Functional classifications define which service tasks must be executed, energy- and power-based classifications indicate how electrification affects charging and infrastructure requirements, and operational working modes explain how vehicle movement and resource constraints affect availability.

2.2. Electrification of GSE

This section discusses how electrification changes the operational planning problem for GSE fleets. Unlike conventional GSE, eGSE availability depends not only on task completion and vehicle location, but also on battery state, charging duration, charger access, and charging strategy. The section therefore focuses on the EV, battery, and charging concepts that are directly relevant for representing eGSE as energy-constrained operational resources.

2.2.1. Challenges of EVs

The literature identifies a broad set of challenges associated with EVs, ranging from vehicle-level limitations and battery behaviour to charging infrastructure, grid impacts, and organisational barriers. For airport GSE, these challenges are particularly relevant because operational reliability is critical. Vehicles must be available at the right time and location during tightly scheduled turnaround processes. Electrification therefore changes the operational planning problem: vehicle availability becomes dependent not only on task completion and travel time, but also on battery state, charging duration, charger access, and infrastructure capacity.

Technical and Operational Constraints

A frequently discussed limitation of battery-electric vehicles is restricted driving range. Several studies report that EVs typically offer shorter ranges than conventional vehicles, which reduces operational flexibility [21]. Battery-electric vehicles that rely on a single energy source may also face performance limitations in demanding operational contexts [46]. Historically, limited range and long charging times have been among the main barriers to widespread EV adoption [46]. For airport GSE, these limitations are operationally important because vehicles such as baggage tractors, cargo transporters, water trucks, and toilet-service vehicles may need to remain available throughout the day with only short idle periods between tasks.

Battery performance introduces additional constraints. Frequent charging and discharging cycles can increase thermal stress and accelerate battery ageing, which is relevant for high-utilisation fleets such as airport GSE [46]. Batteries also lose charge during idle periods due to self-discharge, a process influenced by ambient temperature, storage duration, and previous cycle history [31]. In addition, the battery SOC cannot be measured directly and must be estimated using measurements such as voltage, current, and temperature. Nonlinear electrochemical behaviour, temperature effects, ageing, and charge–discharge dynamics complicate this estimation process [31]. For operational planning, this matters because SOC determines whether a vehicle can be assigned to a task, should remain available, or needs to charge.

The battery State of Health (SOH) represents a related long-term challenge. Similar to SOC, SOH cannot be measured directly at the battery terminals and lacks a universally accepted definition. It is commonly approximated using capacity degradation, internal resistance increase, or a combination of both. These indicators provide useful information, but they remain indirect measures of the battery's actual condition [58]. For airport operations, SOH is especially relevant for long-term fleet planning, maintenance strategy, and technical reserve requirements, because battery degradation reduces usable capacity and can increase vehicle downtime.

Another major operational constraint is charging duration. EVs generally require substantially longer charging periods than conventional vehicles require for refuelling [21]. Charging time is therefore frequently identified as one of the main operational disadvantages of EVs [30]. For airport GSE, long charging durations can create tight scheduling constraints, because vehicles that are charging are temporarily unavailable for service. This introduces trade-offs between keeping vehicles operationally available, charging them at suitable moments, and avoiding excessive battery degradation.

Infrastructure and Grid Constraints

The operational challenges of EVs are compounded by charging infrastructure constraints. Insufficient charger availability remains a barrier to EV deployment and poses practical limitations in airport environments [21]. For eGSE, infrastructure planning is not only a question of the total number of chargers. Charger location, accessibility, available space, power connection, service-road access, and proximity to operational demand also determine whether vehicles can charge without losing too much availability. Zoutendijk notes that charging-station placement for airport surface movement remains a complex and underdeveloped optimisation problem [95]. For Electric Towing Vehicles (ETVs), suitable charging locations depend on airport layout, runway configuration, fleet size, power availability, and service-road accessibility. Similar spatial constraints apply to other eGSE types when chargers must be integrated into existing apron operations.

Large-scale eGSE charging can also create grid-related constraints. Simultaneous charging of multiple vehicles increases peak demand and can place additional load on local electrical infrastructure. Charging during peak hours may also increase operational costs under time-dependent electricity tariffs [46]. Several studies identify high charging loads, voltage drops, increased system losses, equipment overloading, and stability concerns as risks associated with large-scale EV charging [30, 86, 90]. In the airport context, transformer and switchgear overloading have been identified as important technical constraints when deploying large numbers of GSE chargers [59]. These constraints imply that airport electrification requires coordination between fleet planning, charging strategy, infrastructure capacity, and grid connection planning.

Organisational and economic barriers further complicate implementation. Introducing zero-emission vehicles at airports requires investment in vehicles, chargers, electrical infrastructure, and operational adaptation. Bose notes that limited technical expertise and complex procurement procedures can slow down implementation [17]. More generally, high upfront costs and concerns about battery lifecycle remain persistent barriers across the EV sector [90]. For airport operators and ground handlers, these factors mean that electrification decisions must be evaluated not only from an environmental perspective, but also from an operational and infrastructure-planning perspective.

Overall, the deployment of EVs in operationally intensive airport environments is constrained by vehicle range, charging duration, battery behaviour, infrastructure availability, grid capacity, and organisational implementation barriers.

2.2.2. Batteries

Batteries are the main energy storage component of EVs. For eGSE, they are operationally important because they determine how long a vehicle can operate, when it needs to charge, and whether it can remain available during peak turnaround demand. In this thesis, batteries are therefore not treated primarily as electrochemical systems, but as operational constraints that influence vehicle availability, charging behaviour, and fleet sizing.

Several battery technologies can be used in electric vehicle applications, including lead-acid, nickel-based, and lithium-based batteries. These technologies differ in energy density, power density, efficiency, cycle life, cost, self-discharge, and safety characteristics. Lead-acid batteries are relatively inexpensive and technically mature, but have lower energy density and shorter cycle life. Nickel-based batteries generally provide improved energy density and lifetime compared with lead-acid batteries, but are more expensive and often less efficient. Lithium-based batteries are currently the most relevant for transport applications because they offer higher energy density, higher power density, lower self-discharge, longer lifetime, and better fast-charging capability [46, 90, 32]. A more detailed comparison of battery technologies is provided in Section B.1.

Battery selection for eGSE depends on the operational requirements of each vehicle type. Different GSE types have different energy demand, power demand, duty cycles, utilisation patterns, and charging opportunities. In existing airport fleets, lead-acid batteries are still used because they are relatively inexpensive, familiar to operators, and already integrated in many vehicles. However, from a future-oriented operational perspective, lithium-based batteries are more suitable for intensive GH operations. In particular, Lithium Iron Phosphate (LFP) and Lithium Nickel Manganese Cobalt Oxide (NMC) batteries are relevant because they provide higher cycle life and higher energy density than lead-acid alternatives [24]. Safety is also an important selection criterion in airport environments, and LFP batteries are generally considered a relatively safe lithium-based chemistry [14]. A more detailed description of LFP battery structure and battery-management operation is provided in Section B.2.

For operational modelling, the most important battery variable is the SOC. SOC indicates the remaining available energy in the battery and therefore affects whether a vehicle can be assigned to a task, should remain available, or needs to charge. In practice, SOC cannot be measured directly. It is estimated by a Battery Management System (BMS) using measurements such as voltage, current, and temperature [31, 58]. The BMS also protects the battery against unsafe operation, including overcharging, deep discharge, overheating, and cell imbalance. For the simulation model, SOC is therefore the key battery-state variable linking energy use, charging decisions, and vehicle availability.

The SOH is also relevant because it describes battery ageing and the remaining usable capacity of the battery. However, SOH mainly affects long-term fleet planning, maintenance strategy, replacement planning, and technical reserve requirements. It is not explicitly modelled as a dynamic state in this thesis. Instead, the simulation focuses on short-term operational availability during a representative operating day. Battery degradation and detailed electrochemical behaviour are therefore outside the active simulation scope and are discussed only as technical background in Section B.1.

In the simulation model, battery behaviour is represented through operational parameters: battery capacity, energy consumption, SOC, minimum dispatch thresholds, charging thresholds, charging power, and charging rates. This abstraction is appropriate for the purpose of this thesis because the central research question concerns operational fleet capacity under uncertainty and charging constraints, not battery-health optimisation or electrochemical battery modelling. The battery model therefore captures the operational consequences of limited energy and charging-related unavailability, while excluding long-term degradation and detailed battery dynamics.

2.2.3. Charging

Charging is a central operational constraint for eGSE. Conventional GSE can be refuelled quickly, whereas electric vehicles require longer charging periods and depend on the availability of charging infrastructure. As a result, charging affects vehicle availability, idle-time use, charger occupancy, and the number of vehicles that can remain operationally active during demand peaks.

At a technical level, Battery Electric Vehicles (BEVs) are charged from the electrical grid. The grid supplies Alternating Current (AC), while batteries store energy as Direct Current (DC). A charger therefore converts AC power into DC power before energy is stored in the battery [90]. Charging time is not fixed. It depends on battery capacity, initial and target SOC, charging power, charging efficiency, temperature, and vehicle-specific charging limits [90, 22]. Technical details on charging curves, AC/DC charging, charger configurations, and indicative charging levels are provided in Section B.3.

For operational planning, an important distinction is between low-power charging during long idle periods and higher-power charging during shorter breaks. AC charging is often suitable for overnight or long-duration charging because the power level is usually limited by the vehicle's on-board charger. DC charging uses off-board equipment and can provide higher charging power, making it more suitable when vehicles have limited idle time or high daily energy demand [22, 63]. For eGSE, the appropriate charging solution depends on vehicle type, duty cycle, available infrastructure, and operational pressure at the airport [24].

Charging strategies determine when vehicles charge, how long they charge, and which vehicles receive priority when charging capacity is limited. The literature commonly distinguishes between overnight charging, opportunity charging, delayed charging, full charging, partial charging, coordinated charging, and priority-based charging [30, 53, 80]. Overnight charging uses long idle periods outside operational peaks and is suitable for vehicles that can complete a full operating day on one charge. Opportunity charging uses shorter charging sessions during idle periods between tasks and is particularly relevant for vehicles with high utilisation or limited battery capacity. Full charging keeps vehicles connected until a high or complete target SOC is reached, while partial charging stops earlier once an operationally sufficient SOC level has been reached. Partial charging can improve daytime availability because vehicles do not remain unavailable until fully charged. Technical details and examples of charging strategies are provided in Section B.4.

Charging can also be coordinated across the fleet. In uncoordinated charging, a vehicle starts charging immediately when it is connected to a charger. This rule is simple, but it can create charger congestion, high peak loads, and inefficient use of scarce charging capacity. In coordinated or priority-based charging, charging decisions are based on operational criteria such as SOC, expected idle time, task urgency, vehicle availability, or fleet-level charging demand [30, 22, 90]. This is especially relevant for airport operations, where a vehicle with low SOC and an upcoming task may be more critical than a vehicle with higher SOC or more idle time.

In this thesis, charging is not optimised as a separate scheduling problem. Instead, it is represented as part of the rule-based simulation logic. Vehicles can charge when their SOC falls below a threshold or when they are idle at a charging location. Charging infrastructure is modelled as capacity-constrained, meaning that vehicles may have to wait when all chargers are occupied. The model therefore captures the operational effects of limited charging capacity, charging sessions, charger occupancy, charging-related waiting, and charging-related vehicle unavailability.

This modelling choice matches the purpose of the thesis. The goal is not to design an optimal charging schedule, minimise electricity costs, optimise battery health, or evaluate grid services such as demand response or vehicle-to-grid operation. Instead, the aim is to assess how charging constraints interact with dispatching, vehicle availability, operational uncertainty, and fleet sizing. Charging is therefore included as an operational availability constraint that can affect service performance and infrastructure use, while detailed battery chemistry, charging-standard selection, tariff optimisation, and grid-level power-flow effects remain outside the active model scope.

2.3. Uncertainty

Uncertainty is central to airport GSE operations because aircraft timing, service tasks, vehicle movements, infrastructure use, and charging opportunities are closely connected. Small deviations in task timing or apron travel can change vehicle availability and affect whether tasks are completed within their service windows.

This section defines the uncertainty sources relevant for eGSE fleet sizing and explains how they are treated methodologically. It provides the conceptual basis for the uncertainty implementation in Chapter 4, the case-specific derivation in Chapter 5, and the stochastic simulation experiments in Chapter 6. The section first introduces the role of uncertainty and the distinction between aleatory and epistemic uncertainty, then discusses the main uncertainty sources in GSE operations, and finally reviews relevant simulation approaches such as stochastic simulation, Monte Carlo replication, scenario analysis, and empirical or parametric sampling.

2.3.1. Role of Uncertainty in Airport Ground Operations

Uncertainty is a central characteristic of airport ground operations. Aircraft turnaround is a short and tightly coordinated process in which multiple service activities are executed in parallel within strict time windows. GSE do not operate in isolation, but interacts with aircraft schedules, stand availability, service procedures, apron infrastructure, operators, and other vehicles. As a result, a disturbance in one part of the system can propagate to other parts of the turnaround process. For example, a delayed aircraft arrival can shift the timing of several ground-handling tasks, while a delayed GSE movement can reduce vehicle availability for subsequent tasks. These effects can lead to late task starts, late task completion, queueing at shared infrastructure, missed charging opportunities, and additional fleet pressure later in the day [56, 28, 92].

This is particularly relevant for fleet sizing. A deterministic model based on average arrival times, average travel times, and nominal service durations may indicate that a fleet is sufficient under planned conditions. However, such a model may fail to represent peak-load conditions, clustered demand, and operational disturbances. In practice, required fleet capacity is often determined not only by the average workload, but also by the system's ability to absorb variability and recover from disruptions. Uncertainty modelling is therefore needed when the objective is to evaluate operational robustness and determine fleet-capacity requirements under realistic airport conditions [28, 53].

A useful distinction in uncertainty modelling is the distinction between aleatory and epistemic uncertainty. Aleatory uncertainty refers to inherent variability in the system. It is caused by the natural randomness of operational processes and cannot be removed by collecting more data. Epistemic uncertainty refers to imperfect knowledge about the system, data, parameters, or model structure. This type of uncertainty can, in principle, be reduced through better measurements, improved model specification, calibration, or expert validation [23].

Both types occur in airport ground operations. Flight arrival variability, incidental apron congestion, and random travel disturbances are mainly aleatory, because they represent actual operational variability. By contrast, uncertainty about the correct delay distribution, incomplete travel-time observations, simplified task-generation rules, and expert-based disturbance assumptions are mainly epistemic. Similar distinctions apply to service durations, task demand, energy consumption, and charging behaviour. For example, variation in actual energy consumption between trips is an aleatory effect, while uncertainty about the correct battery model or charging threshold is epistemic [23, 31].

This distinction is important because different uncertainty types require different modelling treatments. Aleatory uncertainty can be represented through stochastic simulation, empirical sampling, or probability distributions. Epistemic uncertainty is more often addressed through sensitivity analysis, validation checks, expert judgement, transparent assumptions, or explicit scope limitations. Not every uncertainty source should therefore be modelled as a random variable. Some uncertainties are better treated as assumptions that affect the credibility and interpretation of the results [23, 28]. This thesis follows that logic by modelling task-timing uncertainty and travel-time uncertainty stochastically, while addressing other uncertainty sources through assumptions, validation, sensitivity analysis, or future research.

2.3.2. Sources of Uncertainty in GSE Operations

Several sources of uncertainty affect GSE operations. These sources are related to aircraft timing, task demand, service execution, vehicle movement, and energy availability. Although they can be discussed separately, they interact in practice. A delay in one element of the system can change vehicle availability, create queueing effects, reduce idle time, shift demand into an already busy period, or reduce the time available for charging.

A first major source is flight timing uncertainty. Ground-handling tasks are usually triggered by aircraft movements. Early arrivals, late arrivals, and propagated delays directly influence when aircraft-dependent services can start. This is especially relevant for tasks such as potable-water service, lavatory service, boarding-related support, towing, and other activities that depend on the aircraft being present at the stand. Recent airport operations studies therefore model stochastic aircraft arrival and departure times, delay propagation, and dynamic rescheduling, because deviations from the schedule can change the feasibility of the operational plan [28, 49, 53]. For GSE fleet sizing, flight timing uncertainty is important because it changes the temporal distribution of service demand.

A second source is task demand uncertainty. Not every flight requires the same set of services. Even when a service type is required, the amount of work may differ by aircraft type, flight route, passenger load, airline procedure, turnaround type, and stand configuration. Literature on turnaround and GSE scheduling shows that task requirements depend strongly on aircraft category, operational procedure, service type, and the spatial configuration of the aircraft stand [56, 92]. For example, potable-water and lavatory services may not be required on every flight, and the required water or waste volume may differ by aircraft size and operational context. For simulation modelling, this means that task generation and service quantities must be based on explicit operational rules or input data.

A third source is service-duration uncertainty. Even when a task is required, the time needed to complete it may vary. Service duration can be affected by aircraft size, transferred resource quantity, operator behaviour, stand layout, access conditions, sequencing with other ground-handling activities, and the position of aircraft service interfaces. Several airport scheduling studies therefore include variable service times or stochastic disturbances around nominal task durations instead of assuming one fixed duration for every task instance [28, 44, 49]. For fleet sizing, service-duration uncertainty matters because longer-than-expected tasks delay vehicle availability for subsequent assignments.

A fourth source is travel-time uncertainty. GSE travel time is not fully deterministic, even when the origin and destination are known. Vehicle movements can be affected by apron congestion, temporary blockages, crossing traffic, aircraft pushback movements, stand-side manoeuvring, route restrictions, driver behaviour, and local operational control rules. Many GSE models derive travel time from a distance matrix and an assumed vehicle speed. This provides a transparent deterministic baseline, but it does not fully capture the variability of actual apron movement. Travel uncertainty can therefore be represented through stochastic travel-time buffers, additive delays, multiplicative disturbance factors, or more detailed traffic simulation models [20, 92]. For this thesis, travel-time uncertainty is important because it directly affects vehicle arrival times and subsequent vehicle availability.

A fifth source is energy and charging uncertainty. In electrified GSE operations, vehicle availability also depends on battery state, energy consumption, charging opportunities, and charger access. SOC cannot be measured perfectly, and actual energy consumption depends on duty cycle, vehicle load, driving behaviour, auxiliary power use, temperature, and battery condition. In addition, charging infrastructure is often capacity-constrained, meaning that vehicles may compete for chargers during operational peaks. General EV literature shows that SOC estimation and battery management remain uncertain, while airport-specific eGSE studies show that charger occupancy, charging priority, limited energy capacity, and charging strategy can influence operational feasibility and fleet flexibility [58, 31, 22, 11, 48, 94]. For simulation modelling, these uncertainties matter because they affect whether a vehicle can remain available, needs to charge, or becomes unavailable at a critical moment.

Together, these sources show that GSE operations are affected by interacting uncertainties in aircraft timing, task occurrence, task execution, vehicle movement, and energy availability. For this thesis, task-timing uncertainty and travel-time uncertainty are selected as active stochastic inputs because they directly affect service-demand timing and vehicle availability during the simulated day. Other uncertainty sources, including service-duration variation, task-demand assumptions, energy consumption, charging behaviour, infrastructure availability, battery degradation, weather disruption, and operator behaviour, remain relevant but are addressed through deterministic assumptions, validation checks, sensitivity analysis, scope limitations, or future research. Excluding these sources from the stochastic implementation does not mean that they are irrelevant; it means that they are outside the active uncertainty scope of this study or are treated through another methodological mechanism [92, 80].

2.3.3. Modelling Approaches for Uncertainty in Simulation

A useful starting point for uncertainty modelling is the distinction between deterministic and stochastic models. Deterministic models assume fixed input values for aircraft arrivals, task durations, travel times, energy consumption, and charging behaviour. They are useful as a baseline because they are transparent, reproducible, and relatively easy to interpret. Many scheduling and fleet-planning studies therefore begin with deterministic assumptions before adding more operational detail [56, 16]. However, deterministic models are limited when the objective is to evaluate robustness or fleet-capacity requirements under operational variability. A fleet that is feasible under nominal input values may become infeasible when realistic disturbances are included.

Stochastic models address this limitation by treating one or more inputs as random variables. In airport operations, stochasticity is often introduced in aircraft timing, service duration, task release, travel time, or energy demand. Such models are useful when the objective is not only to generate a feasible plan, but also to evaluate how performance changes when operations deviate from the planned schedule. Robust planning, simheuristic approaches, and dynamic planning models are therefore used to test schedules under disruption or to update decisions when new information becomes available [28, 53, 49]. This is especially relevant for eGSE operations, where task execution, vehicle routing, SOC development, and charging decisions are strongly coupled.

A common method for evaluating stochastic systems is Monte Carlo simulation. In Monte Carlo simulation, the same model configuration is simulated multiple times with different random draws or random seeds. The output is then analysed as a distribution rather than as a single deterministic result. This is useful for fleet sizing because the decision should not be based only on one realised operating day or one average performance value. Instead, the fleet should meet the selected performance target across a range of plausible operational realisations. Monte Carlo simulation therefore allows the modeller to analyse mean performance, variability, confidence intervals, percentiles, and extreme outcomes. This is also relevant for eGSE operations because the timing of vehicle availability, charging opportunities, and infrastructure use can differ between replications [18, 30].

Structured experiment cases are another common way to study uncertainty. Instead of modelling all possible uncertainty sources in one complete probabilistic framework, the modeller defines a limited number of controlled cases. Examples include a deterministic baseline, flight-timing uncertainty only, travel-time uncertainty only, combined uncertainty, low- and high-delay conditions, or limited and expanded charging-infrastructure configurations. Such structured cases are useful for communicating results to practitioners because they are linked to operational questions, such as how service performance changes when delays increase, when travel times become more variable, or when charging infrastructure becomes binding [16, 94, 80]. In this thesis, this logic is used through explicit uncertainty settings, demand-case robustness checks, operational-control variants, and parameter sensitivity tests.

For the representation of stochastic inputs, three broad approaches can be distinguished: empirical, parametric, and hybrid approaches. In an empirical approach, values are sampled directly from historical observations. This is attractive when sufficient representative operational data are available, because it preserves the observed distributional shape and does not force the data into a predefined theoretical distribution [45]. In a parametric approach, a probability distribution is fitted to the data, such as a normal, lognormal, gamma, Weibull, or truncated distribution. This can be useful when a smooth mathematical representation is needed, when the available data are limited or noisy, or when the modeller wants more control over scenario design. However, estimating input distributions from finite operational data introduces input uncertainty, because the selected distribution and its parameters are themselves uncertain [87]. In a hybrid approach, event occurrence and event magnitude are modelled separately. For example, a Bernoulli trial can first determine whether a delay occurs, after which the delay magnitude is sampled from an empirical or parametric distribution. This structure is useful when the probability of an event and the size of the event represent different stochastic processes.

For flight-delay modelling, such a hybrid structure is often appropriate. Many flights are on time or close to on time, while delayed flights have positive delay magnitudes that are typically asymmetric and may include a small number of large delays. Studies on air-transportation delay data show that flight-delay distributions are non-Gaussian, time-dependent, and influenced by delay propagation across the network [25, 50]. It is therefore useful to distinguish between the probability that a delay occurs and the conditional magnitude of the delay once it occurs. A lognormal distribution can be appropriate for positive and right-skewed delay magnitudes because it is defined only for positive values and is commonly used for positively skewed empirical data [47]. A normal distribution may be suitable for some types of service-time variation, but is less suitable for strictly positive delay magnitudes unless it is truncated, because an unbounded normal distribution can generate negative values. Empirical sampling is preferable when the historical dataset is sufficiently large and representative, while parametric fitting can be useful when the dataset is limited, noisy, or when a smoother representation is required [45, 87].

The choice of stochastic representation should be combined with careful data preprocessing. Historical operational data often contain cancelled flights, missing timestamps, early arrivals, on-time flights, delayed flights, and extreme delay values. Before such data are used in a simulation, the modeller must decide how observations are classified and how outliers are treated. Typical preprocessing steps include excluding cancelled flights when the model focuses on operated flights, separating early and late movements, segmenting observations by time of day or task type, and checking whether extreme values represent real disruptions or data errors. Outliers can be retained, removed, capped, or winsorised depending on the purpose of the model. This step is important because uncertainty modelling is not only a matter of choosing a probability distribution, but also of defining which historical observations are relevant for the simulated system [53, 49].

Recent airport-operations literature increasingly combines multiple modelling approaches. Some studies combine optimisation with simulation, while others use rolling-horizon control, digital twins, or simheuristics to evaluate robustness under changing operational conditions [28, 54, 49]. This is relevant for eGSE operations because dispatching, routing, charging, and task timing are interdependent. There is therefore no single best uncertainty-modelling method for all cases. A deterministic baseline is useful for transparency and comparison, stochastic simulation is useful for evaluating robustness, Monte Carlo replication is useful for analysing output variability, and structured experiment cases are useful for comparing operational assumptions.

This thesis combines these approaches. A deterministic baseline is used to establish nominal fleet performance. Task-timing uncertainty and travel-time uncertainty are then introduced as stochastic mechanisms and evaluated through repeated simulation replications. Empirical and hybrid uncertainty representations are used where operational data are available, while other uncertain assumptions are handled through validation checks, sensitivity analysis, or explicit scope limitations. This approach allows the model to assess how uncertainty affects service-window performance, vehicle availability, charging-related resource use, and fleet-capacity requirements.

2.4. Implications for the Simulation Model

This section synthesises the operational, electrification-related, and uncertainty-related foundations required for eGSE fleet simulation. The purpose of this chapter is not to provide the full modelling literature review; that is addressed in Chapter 3. Instead, this chapter identifies what an eGSE fleet-sizing model must be able to represent before existing modelling approaches are evaluated. It therefore forms the link between the operational foundations in this chapter, the literature review in Chapter 3, and the simulation framework developed in Chapter 4.

The airport-operations literature shows that GSE demand is created within the aircraft turnaround process. The turnaround window between IB and OB determines when ground-handling tasks become relevant, where they must be executed, and when they should be completed. This implies that the simulation model must represent airside service demand at the level of individual tasks. Each task should therefore include a service type, aircraft stand, service duration, release time, and service window.

The model does not represent the full aircraft turnaround as an integrated precedence network. Instead, it isolates selected handling tasks and represents them through task-specific service windows within the turnaround process. This means that each task is positioned at the correct operational moment relative to the aircraft movement, while dependencies with other handling tasks are not explicitly modelled. This modelling choice fits the objective of this thesis, which is to evaluate vehicle availability and task completion rather than full turnaround sequencing or aircraft departure-delay causality.

The same literature also shows that GSE demand is not uniform across flights. It depends on aircraft type, stand configuration, airport layout, airline-specific procedures, and service requirements. Different GSE types also follow different operating patterns. Some vehicles mainly operate at or near the aircraft stand, while others repeatedly move between aircraft stands, depots, refill points, dump points, or other support locations. Some vehicles have onboard resource constraints, such as potable-water capacity or waste-storage capacity, while others are mainly constrained by availability, travel time, service duration, or energy state. For this reason, the simulation model must distinguish between vehicle groups and include vehicle-specific operating rules, including aircraft compatibility, service duration, travel behaviour, depot-based or stand-side operation, onboard resource depletion, refilling, dumping, and post-task repositioning.

The electrification literature adds a second modelling requirement. eGSE availability cannot be treated in the same way as conventional GSE availability. Conventional vehicles can often be assumed to return to service after task completion or short refuelling, whereas eGSE may become unavailable because of limited battery capacity, charging duration, charger access, or operational charging rules. Electrification therefore changes the modelling problem from a purely time- and location-based dispatching problem into an energy-constrained vehicle-availability problem.

For this thesis, this means that each vehicle must be represented as an individual operational resource with a dynamic state. The model must track vehicle-level SOC, battery capacity, energy consumption, minimum dispatch thresholds, charging thresholds, charging sessions, charger occupancy, and charging-related waiting. These elements determine whether a vehicle can be assigned to a task, whether it should charge instead, and whether charging infrastructure can become a temporary bottleneck during the operational day. Charging is therefore included within the operational simulation horizon rather than treated only as an external overnight assumption. This allows charging behaviour to affect dispatching feasibility, vehicle availability, infrastructure use, and service-level performance.

The uncertainty literature adds a third modelling requirement. Airport ground operations are affected by variability in flight timing, task occurrence, service duration, travel time, energy use, charging behaviour, infrastructure availability, operator behaviour, and other disturbances. For fleet sizing, uncertainty is important because it can change both when service demand occurs and when vehicles become available again. A deterministic model may therefore indicate that a fleet is sufficient under nominal conditions, while the same fleet may perform insufficiently under operational variability.

Different uncertainty sources require different modelling treatments. Short-term operational variability can be represented through stochastic simulation, while uncertain assumptions are better addressed through validation checks, sensitivity analysis, structured experiment cases, or explicit scope limitations. This thesis therefore combines a deterministic baseline, stochastic uncertainty settings, and sensitivity analysis rather than modelling all uncertainty sources probabilistically.

Task-timing uncertainty and travel-time uncertainty are selected as the active stochastic inputs. Task-timing uncertainty changes the temporal distribution of service demand, while travel-time uncertainty affects vehicle arrival times and subsequent vehicle availability. These two sources are modelled stochastically because they directly influence service-window feasibility and fleet-size requirements. Other uncertainty sources remain relevant but are treated differently. Service-duration variation, task-demand assumptions, water and waste quantities, energy consumption, charging power, battery capacity, charging behaviour, and infrastructure capacity are represented through deterministic assumptions, validation checks, structured experiment cases, or sensitivity analysis. Long-term battery degradation, detailed operator behaviour, weather disruption, vehicle breakdowns, charger failures, and real-time stand changes are acknowledged as relevant but remain outside the active simulation scope.

Together, these insights define the modelling basis of this thesis. An eGSE fleet-sizing model should represent fleet capacity as a dynamic vehicle-availability problem. Required fleet capacity is not determined only by the number of tasks. It also depends on when and where tasks occur, how long vehicles spend travelling and servicing, whether vehicles have sufficient battery energy and onboard resources, whether charging and support infrastructure are available, and how uncertainty changes vehicle availability during the day.

These implications are used in two ways. First, they define the requirements that guide the simulation framework in Chapter 4. Second, they provide the evaluation lens for the literature review in Chapter 3. The literature review therefore assesses whether existing studies combine task execution, individual vehicle states, charging behaviour, infrastructure constraints, operational uncertainty, and fleet-size evaluation within one coherent operational modelling approach. This leads to the research gap addressed in this thesis: the need for an integrated simulation framework that links these elements in a dynamic eGSE fleet-sizing approach.

3
Literature Review and Research Gap

The literature review builds directly on the modelling needs identified in Chapter 2. The operational foundations define the elements that an eGSE fleet-sizing model should capture: task execution, service windows, vehicle movement, fleet heterogeneity, resource constraints, charging, infrastructure availability, and uncertainty. The purpose of this chapter is therefore to assess whether existing GSE and eGSE models integrate these elements in a way that supports operational fleet sizing.

This creates the link between the background chapters and the model developed later. If an element is identified as operationally relevant in Chapter 2, the literature review evaluates how it is represented in prior work. If prior work treats that element only partially, the gap becomes a modelling requirement for the simulation framework in Chapter 4.

The chapter is structured as follows. Section 3.1 reviews the main modelling objectives in the literature and shows how fleet sizing is treated across studies. Section 3.2 compares optimisation, simulation, and heuristic approaches and discusses their suitability for operational decision support. Section 3.3 evaluates how studies represent time and uncertainty. Section 3.4 reviews whether fleets are represented as aggregated vehicle groups or as individual vehicles with vehicle-level states. Section 3.5 analyses the representation of turnaround task execution, travel, service windows, and energy consumption. Section 3.6 reviews how charging is integrated into operational models. Finally, Section 3.7 brings these findings together, identifies the research gap, and positions the simulation framework developed in this thesis. An overview is presented in Table 3.1.

3.1. Modelling Objectives

Across the reviewed literature, three dominant modelling objectives can be distinguished: energy management, operational routing and scheduling, and environmental impact. These objectives differ in how the modelling problem is formulated, which decision variables are included, and whether fleet size is treated as an input or an outcome.

A first group of studies focuses on operational routing and task scheduling. In these studies, the core objective is to allocate vehicles to tasks and sequence operations efficiently within turnaround constraints. Typical objectives include minimising travel distance, completion time, delay, operational cost, or workload imbalance [16, 15, 20, 27, 28, 44, 49, 77, 80, 84, 88]. These studies are directly relevant for GSE operations because they address how limited vehicle resources are assigned to time-sensitive ground-handling tasks.

A second group of studies focuses primarily on energy and charging management for eGSE. These studies model battery dynamics, charging strategies, grid interaction, energy flexibility, or charging-infrastructure use [11, 18, 30, 39, 48, 53, 94]. In these papers, routing and task allocation are often simplified or treated as exogenous inputs. Their main contribution is therefore not detailed operational dispatching, but the representation of energy use, charging behaviour, and infrastructure constraints.

A third perspective concerns environmental impact. Bosma explicitly quantifies emissions and environmental effects rather than optimising daily operational decisions [19]. This perspective is relevant because electrification is often motivated by emission reduction, but it usually provides less detail on operational feasibility, vehicle availability, and day-of-operations decision-making.

Regarding fleet size, most reviewed studies treat the number of vehicles as a fixed input. However, some studies explicitly aim to determine the required number of vehicles, either as a primary objective or as an outcome of the model [19, 28, 30, 39, 80, 88]. Bosma provides an example of a hybrid approach, in which fleet size is used both as an input and as an output: system performance is evaluated for a given number of vehicles, after which the fleet is scaled up or down to identify more efficient configurations [19].

For this thesis, this distinction is important because fleet size should not only be treated as a fixed input to an operational model. Instead, the model must be able to evaluate how different fleet sizes affect service-level performance under operational, charging, and uncertainty constraints. This motivates the use of simulation experiments in which fleet size is varied and assessed as an outcome of system performance.

3.2. Model Types

Across the reviewed literature, three main modelling approaches can be distinguished: optimisation-based models, simulation-based models, and (meta)heuristic methods. Each approach addresses the GSE and eGSE planning problem from a different perspective. Optimisation-based models are designed to identify optimal or near-optimal decisions under explicit constraints. Simulation-based models evaluate how a system behaves over time under given rules and assumptions. (Meta)heuristic methods are often used when exact optimisation becomes computationally difficult due to problem size or operational complexity.

3.2.1. Optimisation-Based Models

Optimisation-based models formulate the planning problem explicitly using mathematical programming or related optimisation techniques. They aim to find optimal solutions with respect to defined objectives and constraints. Examples include Bao [16], Fu [27], Tang [77], and Timmermans [80]. These models differ in their mathematical structure, computational properties, and ability to handle operational complexity, uncertainty, and multiple objectives.

Mixed-Integer Linear Programming (MILP) is the most frequently applied exact optimisation technique in the reviewed literature. It formulates decision problems using linear objective functions and constraints, with both continuous and integer decision variables. Alruwaili applies MILP to maximise airport profitability by allowing eGSE to participate in energy markets [11]. Bao uses MILP to minimise total operational costs for a mixed fleet [16]. Kuipers embeds MILP in a Model Predictive Control (MPC) framework, while Tang applies it to balance workload and energy consumption [44, 77]. Van Oosterom and Zoutendijk use MILP for fuel-saving analysis and planning under energy constraints, respectively [53, 94].

Constraint programming is another optimisation approach. It focuses on satisfying logical and temporal constraints rather than optimising a single linear objective function. Gök uses constraint programming to manage complex planning rules and to restrict the solution space to schedules that remain robust under operational disruptions [28].

Optimisation can also be applied in a rolling-horizon setting. In this approach, the optimisation problem is repeatedly solved over a moving time window, allowing decisions to be updated as new operational information becomes available. Kuipers applies MPC to dynamically reschedule tasks in real time when flight delays or other disturbances occur [44].

Column generation is a decomposition technique designed for large-scale optimisation problems that are too complex for standard solvers. Tang uses this method to obtain optimal or near-optimal solutions for cooperative planning problems involving many interacting resources [77]. Pareto analysis is used in multi-objective optimisation to identify trade-offs between conflicting objectives. Gulan applies this technique to analyse the balance between vehicle investment costs and fuel savings [30].

Overall, optimisation-based models are useful when the objective is to find a best feasible plan under explicit constraints. However, their operational realism often depends on how much detail can be included without making the optimisation problem computationally intractable. This is particularly relevant for eGSE, where vehicle routing, task timing, SOC, charging, infrastructure capacity, and uncertainty can become strongly interdependent.

3.2.2. Simulation-Based Models

Simulation-based models focus on evaluating system behaviour rather than directly optimising decisions. Studies such as Gulan [30], Kirca [39], Bose [18], and van Oosterom [53] use simulation to analyse energy demand, charging behaviour, operational feasibility, or system robustness. These models do not guarantee optimality, but they are well suited for representing dynamic interactions, operational variability, and resource constraints.

Different simulation paradigms are used in the reviewed literature. Agent-Based Modelling (ABM) simulates system behaviour from the bottom up by modelling autonomous agents, such as vehicles, aircraft, or operational entities. Chen and Kuipers use ABM to capture interactions between operational entities [20, 44]. Luo extends this approach by integrating digital twins, enabling real-time monitoring and detection of deviations between the physical airport and its virtual representation [49].

Monte Carlo simulation evaluates system performance by repeatedly sampling random scenarios or input realisations. Gulan and Bose use Monte Carlo simulation to quantify the impact of uncertainty in flight schedules and energy demand [18, 30]. This approach is relevant for fleet sizing because it allows performance to be evaluated across multiple stochastic realisations rather than under a single deterministic input set.

Discrete-Event Simulation (DES) represents system evolution as a sequence of discrete events in time. This makes it suitable for modelling processes such as task release, vehicle dispatching, travel, service execution, charging, and queueing. Gök applies DES, based on Petri nets, to study how delays propagate through the GH process [28].

Simulation can also be combined with optimisation. Simheuristics embed simulation directly into a search or optimisation process, allowing candidate solutions to be evaluated under stochastic or dynamic conditions. Gök uses this hybrid approach to assess solution robustness and guide optimisation based on simulated performance [28]. Kirca develops a Multi-Input Multi-Output Airport Energy Model as a dedicated simulation model to estimate eGSE charging requirements under different turnaround procedures and operational conditions [39].

For this thesis, simulation-based modelling is particularly relevant because the objective is not to generate an optimal daily schedule, but to evaluate how task execution, dispatching rules, vehicle states, charging behaviour, infrastructure constraints, and uncertainty interact during daily operations.

3.2.3. (Meta)heuristic Models

A third group of studies relies on heuristic or metaheuristic methods to solve large or complex scheduling, routing, and coordination problems that are difficult to handle with exact optimisation alone. Examples include Fu [27] and Luo [49]. These approaches trade optimality guarantees for computational tractability, flexibility, and the ability to handle operational complexity.

Genetic algorithms are population-based search methods inspired by natural evolution. Luo proposes an enhanced genetic algorithm with a three-layer structure for vehicle assignment and service sequencing [49]. Simulated annealing is a probabilistic search method that allows temporary acceptance of worse solutions to escape local optima. Bao, Tang, and Zoutendijk apply simulated annealing to improve solution quality in energy-aware scheduling problems [16, 77, 94]. Tabu search prevents cycling by temporarily forbidding recently visited solutions. Bao and Zoutendijk use tabu search to enhance exploration in fleet-planning problems [15, 94].

Adaptive Large Neighborhood Search (ALNS) iteratively destroys and reconstructs parts of a solution to explore the search space efficiently. Bao and Zoutendijk apply ALNS to solve large-scale fleet-scheduling problems with reduced computation time [16, 94]. Chen introduces a Temporal Sequential Single Item auction, in which vehicle agents bid for tasks based on their own costs and schedules, enabling decentralised task allocation [20]. Chen also uses Safe Interval Path Planning and its learning-augmented variant to compute conflict-free GSE routes in time-expanded airport networks [20].

These methods are useful when exact optimisation becomes computationally expensive or when decentralised decision-making is required. However, in this thesis, heuristics are not used as a separate optimisation method. Instead, rule-based heuristics are embedded within the simulation model to represent operational decision-making, including task selection, vehicle selection, charging decisions, and repositioning behaviour.

For this thesis, simulation is therefore the most suitable primary modelling approach. The objective is to evaluate how task execution, charging, resource constraints, dispatching logic, and uncertainty interact during daily operations, rather than to produce an optimal schedule. A rule-based DES can represent these interactions explicitly and can compare deterministic and stochastic fleet-size performance. Within this simulation, heuristic decision rules are used to represent operational control logic. The thesis therefore combines a simulation-based model structure with embedded rule-based decision heuristics.

3.3. Representation of Time and Uncertainty

Across the reviewed literature, studies differ in how they represent time and uncertainty. With respect to time, a distinction can be made between static and dynamic models. Static models assume fixed schedules, fixed inputs, and decisions that are made for the complete planning horizon in advance. They are useful for strategic analysis or simplified planning problems, but they are less suitable for operational settings where flight timing, vehicle availability, and charging opportunities evolve throughout the day. Dynamic models explicitly represent temporal system evolution and allow decisions or system states to change over time. This makes them more suitable for airport ground operations, where tasks are released at specific moments, vehicles move between locations, and operational feasibility depends on the state of the system at the time a task must be executed.

Regarding uncertainty, some studies remain deterministic and assume known flight schedules, service times, travel times, and energy requirements. Other studies explicitly include uncertainty or variability through stochastic formulations, simulation-based randomness, or disturbance factors. Gulan uses stochastic simulation to assess the impact of flight-schedule variability and energy demand uncertainty on eGSE operations [30]. Xu employs a stochastic dynamic programming model with chance constraints to represent uncertainty in bus arrival times [88]. Chen combines deterministic and stochastic modelling elements: the baseline schedule is constructed using fixed flight times and service durations, while travel-time uncertainty is represented through a stochastic distance-buffer coefficient that accounts for variability in vehicle movements [20]. Kuipers applies ABM and introduces a stochastic offset of 20% to nominal task duration to emulate operational disturbances within the simulation [44]. Luo represents uncertainty through parameters that capture fluctuations in service and flight times, often using normal distributions derived from historical delay data [49].

A notable observation is that several recent studies combine multiple modelling paradigms. Examples include optimisation embedded in simulation, rolling-horizon control, agent-based simulation, and digital-twin-supported heuristic approaches [49, 44]. This indicates a trend towards hybrid models that aim to balance operational realism, scalability, and decision quality. Such approaches are relevant for eGSE planning because dispatching, routing, charging, vehicle availability, and task timing are interdependent and evolve during the operational day.

For this thesis, a dynamic and stochastic representation is required. Required fleet capacity depends on how individual vehicles become available over time, where they are located, whether they have sufficient battery energy, and whether service tasks can still be completed within their time windows. The model must therefore represent temporal system evolution and allow operational uncertainty to affect task timing, travel time, vehicle availability, and service-window feasibility. A deterministic baseline remains useful for comparison, but it is insufficient to assess robust fleet-capacity requirements under operational variability.

3.4. Fleet Representation

Across the reviewed studies, clear differences can be observed in how the fleet is represented. These differences relate to vehicle technology, fleet composition, vehicle heterogeneity, and the level of detail at which vehicle states are modelled.

First, studies differ in the vehicle technology they consider. Some papers focus exclusively on eGSE fleets, such as Alruwaili [11], Bose [18], and Tang [77]. Other studies consider conventional fuel-based fleets or explicitly analyse mixed fleets of electric and conventional vehicles to study transitional phases or hybrid operations, such as Bao [16], Kirca [39], Wang [84], and Bosma [19]. This distinction is important because conventional and electric fleets are constrained by different operational factors. Conventional fleets are mainly constrained by vehicle availability, task timing, and travel time, while eGSE fleets are additionally constrained by battery state, charging duration, charger availability, and charging strategy.

Second, the number of vehicle types included varies widely. Some studies focus on a single GSE type, such as passenger buses [88], while others include heterogeneous fleets with multiple vehicle classes [39, 48]. A heterogeneous fleet representation is more complex because each vehicle group may have different service tasks, duty cycles, energy consumption, charging requirements, and operational constraints. However, this level of detail is necessary when the objective is to compare fleet requirements across different GSE types or to analyse interactions between vehicle groups and shared infrastructure.

Third, studies differ in whether vehicles are represented in an aggregated or individual manner. In aggregated fleet models, vehicles of the same type are treated as interchangeable units, and fleet states are represented through totals, averages, or distributions. This can be appropriate for high-level energy analysis, strategic fleet planning, or infrastructure-load estimation. However, it limits the ability to evaluate whether a specific vehicle can perform a specific task at a specific moment. In individual vehicle models, each vehicle has its own state, such as location, availability, assigned task, battery level, and charging status. Studies such as Kuipers [44] and Tang [77] use this more detailed representation to capture operational feasibility at vehicle level.

For this thesis, individual vehicle representation is required because vehicle-level states determine whether a task can be executed. A vehicle can only be assigned to a task if it is available at the right time, located within feasible reach, compatible with the task, has sufficient SOC, and satisfies service-specific constraints such as onboard water or waste capacity. Aggregated fleet representations are therefore insufficient for the operational questions addressed in this study. The simulation model must track individual vehicles over time, including their location, availability, battery state, charging status, resource level, and task history. This allows fleet-capacity requirements to be evaluated as an outcome of dynamic vehicle availability rather than as a fixed aggregate fleet assumption.

3.5. GSE Turnaround Operations

Across the reviewed literature, the level of operational detail in GSE turnaround modelling differs substantially. Routing- and scheduling-oriented studies typically include task allocation, travel times between locations, service time windows, and service times at the aircraft stand. Energy-focused studies, by contrast, often simplify operational processes and mainly use them to estimate energy consumption, charging demand, or vehicle availability [30, 53]. This distinction is important because the level of operational detail determines whether a model can evaluate service feasibility at task level or only approximate aggregate energy and fleet requirements.

One important modelling choice concerns the representation of service time. Service time is defined here as the time a vehicle physically spends servicing an aircraft after arriving at the stand. Across the reviewed papers, service time is either explicitly modelled or omitted. When included, it is represented in different ways. Some studies use fixed service times per task, while others define service duration as a function of aircraft characteristics, vehicle type, resource type, or stochastic disturbance. For example, service times may depend on aircraft size or category [19], be specified per task type [20], be differentiated by vehicle or resource type [27], or be modelled as stochastic or otherwise variable parameters [28]. These differences matter for fleet sizing because service duration directly affects when a vehicle becomes available for its next task.

Energy consumption is represented with a similar range of modelling detail. Some studies use simplified average-based formulations, while others apply more detailed mechanical or activity-based models. Bao uses a mechanical energy-consumption model that accounts for rolling resistance, aerodynamic drag, and gradient resistance [16]. This formulation distinguishes between operational modes such as unloaded driving, towing with aircraft weight, and energy use related to Auxiliary Power Unit (APU) substitution. Bosma decomposes energy consumption into driving, pumping, and idling [19]. Driving energy is calculated using resistance-based equations, while pumping and idling are represented as constant-power demands, such as 40 kW for pumping and 32 kW for idling. Kirca applies a more aggregated approach by estimating energy consumption as the product of nominal vehicle power, a load factor, and duration of use [39]. The load factor represents the ratio between average operating power and maximum power and differs across operational modes such as loaded transit, unloaded transit, and high-auxiliary operation.

These modelling choices show that operational detail is often traded off against model tractability. Detailed representations of routing, service execution, resource use, and energy consumption can improve operational realism, but they also require more input data and increase model complexity. Simplified representations are easier to implement and interpret, but may fail to capture the mechanisms that determine whether a vehicle is actually available when a task must be performed.

For this thesis, task allocation, travel time, service windows, service duration, vehicle-specific resource constraints, and energy consumption must be represented together. These elements jointly determine whether an eGSE vehicle can reach a task in time, complete the service within its window, retain sufficient onboard resources, and maintain enough battery energy for subsequent operations. The simulation model therefore represents GSE turnaround operations at task level, while abstracting from more detailed physical vehicle dynamics that are not required for fleet-capacity evaluation.

3.6. Charging

The reviewed papers show substantial differences in how charging is integrated into eGSE models. A broad distinction can be made between implicit and explicit charging integration. Implicit charging integration refers to approaches in which charging is treated outside the operational model, for example through an initial SOC assumption or an overnight recharge assumption. In these cases, charging does not dynamically affect vehicle availability, task feasibility, routing, dispatching, or scheduling during the modelled operational horizon. Timmermans provides an example of such an approach, where charging is not represented as an operational decision within the daily planning problem [79].

Explicit charging integration, by contrast, represents charging as part of the system dynamics or decision-making process. These models typically include vehicle SOC, charging power limits, charger availability, and sometimes Bidirectional Charging/Vehicle-to-Grid (V2G) flows. Charging can therefore directly influence task feasibility, vehicle availability, routing, scheduling, and system performance. Examples include Gulan [30], Alruwaili [11], Tang [77], and van Oosterom [53]. This distinction is important for eGSE fleet sizing because charging is not only an energy-management issue, but also a source of operational unavailability.

Building on this distinction, the literature considers several charging strategies. Overnight charging assumes that vehicles are charged during long idle periods, often until they are fully charged before the next operating day. Opportunity charging allows vehicles to charge during shorter idle periods between tasks and is therefore more relevant for high-utilisation vehicles or vehicles with limited battery capacity. SOC threshold-based charging starts or stops charging when predefined battery thresholds are reached, which can be used to maintain operational availability, reduce peak demand, or limit battery degradation [44, 27, 18]. These strategies differ in how they use idle time and in how strongly they affect vehicle availability during the operational day.

When charging infrastructure is capacity-constrained, priority rules become relevant. Priority-based charging determines which vehicle should charge first when the number of vehicles requiring charge exceeds the number of available chargers. The priority can be based on criteria such as lowest SOC, expected idle time, expected future availability, or operational urgency [30, 19]. Such priority mechanisms are important in airport environments because a vehicle with low SOC and an upcoming task may be more critical than a vehicle with more remaining energy or longer idle time.

Some studies go further by integrating charging decisions with operational scheduling. This can be referred to as integrated charging scheduling, where task assignment, routing, and charging are jointly planned rather than treated as separate decisions. Tang and Fu provide examples of models in which charging is integrated with task allocation and scheduling decisions [77, 27]. Wang applies a deep-reinforcement-learning-based scheduling paradigm that integrates electric and fuel-powered vehicles and provides interpretable insights through Gantt-chart visualisations [84]. In this approach, an agent learns when a vehicle should charge or execute a task, guided by rewards and penalties related to battery health, energy consumption, and operational performance.

A smaller subset of the literature extends charging integration towards demand response or V2G. In these models, eGSE are not only operational vehicles but also flexible electrical assets that can interact with the grid. Alruwaili and Liu, for example, explore how eGSE can provide grid services or participate in energy-market mechanisms [11, 48]. These approaches are relevant for long-term airport energy-system planning, but they require additional assumptions about market participation, grid constraints, charging infrastructure, and battery degradation.

For this thesis, charging must be represented explicitly within the operational simulation horizon. Charging is not treated only as an overnight reset or external energy assumption, because vehicles that are charging are temporarily unavailable for service. The model therefore needs to include vehicle-level SOC, charging thresholds, charging duration, charger capacity, charger occupancy, and charging-related waiting. At the same time, charging is kept as an operational constraint rather than formulated as a separate charging-schedule optimisation, grid-optimisation, or V2G problem. This modelling choice fits the objective of this thesis: to evaluate how charging constraints interact with dispatching, vehicle availability, task completion, and fleet-size requirements.

3.7. Research Gap and Positioning

The reviewed literature shows that GSE and eGSE operations have been studied from several complementary perspectives. Routing and scheduling studies provide methods for assigning vehicles to tasks, sequencing operations, and reducing travel time, delay, or workload imbalance. Energy-management studies represent battery behaviour, charging demand, charging strategies, grid interaction, and energy flexibility. Simulation-based studies are useful for evaluating system behaviour under dynamic or uncertain conditions. Environmental assessment studies quantify the potential emission benefits of electrification. Together, these streams show that eGSE fleet planning requires the joint consideration of task execution, vehicle movement, energy availability, charging infrastructure, and operational variability.

However, the review also shows that these elements are often treated separately. Studies that include detailed charging or grid interaction frequently simplify routing, task execution, or turnaround service requirements. Routing and scheduling studies often represent task allocation and travel times in detail, but include charging behaviour, charger capacity, or stochastic day-of-operations variability only in a limited way. Several studies also represent fleets as aggregated vehicle classes or assume fleet size as a fixed input. This limits their ability to explain how individual vehicle states, locations, battery levels, charging status, and operational-control rules affect service feasibility during the operational day.

A first gap concerns the coupling between operational task execution and charging. Several studies model charging behaviour or energy demand in detail, but do not fully connect charging-related unavailability to task-level service feasibility during daily operations. For eGSE, this coupling is important because a vehicle that is charging is temporarily unavailable for dispatch, and charger capacity can therefore affect whether service tasks can still be completed within their time windows.

A second gap concerns the role of fleet size in the modelling framework. In much of the reviewed literature, fleet size is defined before the model is evaluated. This is useful for scheduling, energy analysis, or infrastructure-load estimation, but it provides limited insight into how many vehicles are required to meet a given service level. For fleet-sizing decision support, fleet size should instead be varied and evaluated as an operational outcome.

A third gap concerns the treatment of uncertainty. Airport ground operations are inherently variable, but uncertainty is not always linked to vehicle availability, charging behaviour, and service-window feasibility. Deterministic models can show whether a fleet is sufficient under nominal conditions, but they do not show whether that fleet remains sufficient when task timing or travel times vary during the operational day.

Taken together, these gaps indicate the need for an operational simulation framework that integrates task execution, individual vehicle states, charging behaviour, infrastructure constraints, fleet-size evaluation, and operational uncertainty. This thesis addresses that gap by developing a rule-based Discrete-Event Simulation (DES) framework for eGSE fleet sizing. The model represents individual vehicles, task time windows, dispatching decisions, travel between airport locations, SOC development, charging behaviour, charger capacity, depot-return behaviour, and service-specific resource constraints such as refilling and dumping. It also includes task-timing and travel-time uncertainty to evaluate how fleet performance changes under stochastic operating conditions.

This thesis builds on the closely related work of Timmermans, who developed an optimisation-based model to analyse the impact of eGSE implementation on fleet capacity and energy demand [80]. That work provides an important strategic basis for eGSE fleet planning. The present study extends this perspective towards operational execution and robustness. Instead of primarily analysing strategic fleet-capacity implications across a broad set of vehicle types, this thesis focuses on a narrower set of vehicle groups and represents their operational behaviour in more detail.

The case-study application focuses on water vehicles, toilet vehicles, and loaders. Water and toilet vehicles represent depot-based, resource-constrained operations with refill or dump requirements, while loaders represent stand-side service equipment without onboard service-resource constraints. This scope allows the model to analyse not only how many vehicles are required, but also why vehicles become unavailable, where bottlenecks occur, and how operational-control assumptions affect service performance.

Overall, this thesis contributes a simulation-based decision-support approach for eGSE fleet sizing under operational uncertainty and charging constraints. The contribution is not only a set of fleet-size estimates for the tested vehicle groups, but also a modelling framework for understanding how fleet capacity, demand timing, vehicle positioning, charging, resource constraints, infrastructure use, and robustness interact in electrified airport ground-handling operations.

3.8. Summary

This chapter reviewed the modelling literature on GSE and eGSE operations using the modelling needs identified in Chapter 2. The review focused on how existing studies represent the elements required for operational eGSE fleet sizing: task execution, vehicle movement, fleet representation, time, uncertainty, charging behaviour, and infrastructure constraints. The reviewed studies differ in their primary objective. Some focus on routing and scheduling, others on energy management, charging, or environmental impact. They also differ in modelling paradigm, ranging from optimisation and simulation to (meta)heuristics and hybrid combinations. In general, routing- and scheduling-oriented studies tend to represent operational task assignment and travel in more detail, while energy-focused studies often include charging and battery behaviour more explicitly. However, these modelling perspectives are frequently treated separately. The review also showed important differences in how studies represent time, uncertainty, fleets, and charging. Some models are static and deterministic, while others include dynamic system evolution or stochastic variability. Fleet representation ranges from aggregated vehicle groups to individual vehicle tracking. Charging integration also varies strongly.

The structured comparison in Table 3.1 summarises these differences across the reviewed studies. The table compares each study along the main modelling dimensions discussed in this chapter: objective, model type, time representation, uncertainty treatment, fleet representation, turnaround-operation detail, energy-consumption modelling, charging integration, and charging strategy. It therefore provides the basis for comparing existing approaches with the modelling scope of this thesis. Implicit charging integration refers to charging that is treated outside the operational model, for example through an initial SOC assumption or an overnight recharge assumption. Explicit charging integration refers to charging that is represented within the model through elements such as battery state, charging duration, charger availability, charging location, or charging status. This distinction is separated from charging strategy: overnight and opportunity charging describe when charging takes place, while full and partial charging describe the charging target or stopping rule. Integrated charging scheduling is identified only when charging decisions are represented within the operational modelling horizon and can affect vehicle availability, task feasibility, routing, dispatching, or scheduling.

The comparison shows that existing studies provide important building blocks, but that the full integration of fleet sizing, task execution, individual vehicle availability, charging behaviour, infrastructure constraints, and operational uncertainty remains limited. This is the research gap addressed in this thesis. The simulation framework developed in Chapter 4 is therefore positioned as an operational decision-support model that links these elements in one dynamic eGSE fleet-sizing approach.

For this thesis, Table 3.1 also clarifies the distinction between the general model structure and the validated case-study scope. The model framework can incorporate thirteen GSE vehicle groups using the same underlying structure, provided that vehicle-specific input data and operating rules are available. However, the simulation experiments and case-study validation in this thesis are limited to three vehicle groups: water vehicles, toilet vehicles, and loaders. The table therefore shows both how the thesis is positioned relative to existing literature and how its broader modelling potential differs from the narrower validated case-study application.

Table 3.1: *A comparison of different studies that include the modelling of GSE operations.*

	Alruwaili (2022) [11]	Bao (2023) [16]	Bao (2024) [15]	Bose (2025) [18]	Bosma (2022) [19]	Chen (2022) [20]	Fu (2025) [27]	Gök (2022) [28]	Gulan (2019) [30]	Kirca (2020) [39]	Kuipers (2024) [44]	Liu (2024) [48]	Luo (2024) [49]	Tang (2025) [77]	Timmermans (2025) [80]	van Oosterom (2024) [53]	Wang (2025) [84]	Xu (2023) [88]	Zoutendijk (2024) [94]	This Thesis (2026)
Objective	🔋	📍	📍	🔋	🍃	📍	📍	📍	🔋	🔋	📍	🔋	📍	📍	📍	🔋	📍	📍	🔋	📍
– Determines number of vehicles					✓			✓	✓	✓					✓			✓		✓
– Number of vehicles as an input	✓	✓	✓	✓	✓	✓	✓				✓	✓	✓	✓		✓	✓		✓	✓
Model																				
– Optimisation	✓	✓	✓			✓	✓	✓			✓		✓	✓	✓		✓	✓	✓	
– Simulation				✓	✓	✓		✓	✓	✓	✓	✓	✓			✓	✓		✓	✓
– (Meta)heuristic		✓	✓				✓	✓			✓		✓					✓	✓	✓
Time																				
– Static						✓		✓												
– Dynamic	✓	✓	✓	✓	✓		✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Uncertainty																				
– Deterministic	✓	✓			✓	✓	✓			✓		✓		✓	✓				✓	✓
– Stochastic			✓	✓		✓		✓	✓		✓		✓			✓	✓	✓		✓
Fleet	eGSE	Mix	GSE	eGSE	Mix	GSE	eGSE	GSE	Mix	Mix	eGSE	Mix	GSE	eGSE	Mix	eGSE	Mix	GSE	eGSE	eGSE
– Number of vehicle types	4	1	1	4	1	5	3	7	3	16	5	17	8	2	16	1	4	1	1	13
– Modelled individually	✓	✓	✓		✓		✓		✓		✓		✓	✓	✓		✓	✓		✓
GSE Turnaround Operations																				
– Task allocation		✓	✓			✓	✓	✓	✓		✓		✓	✓	✓		✓	✓		✓
– Travel time between locations	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
– Service time window		✓	✓			✓	✓	✓		✓	✓	✓	✓	✓	✓		✓	✓		✓
– Service time at aircraft stand	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
– Individual service times	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
– Vehicle energy consumption	✓	✓		✓	✓			✓	✓	✓	✓	✓		✓	✓	✓	✓		✓	✓
Charging Integration	Expl	Expl		Expl	Expl		Expl		Expl	Expl	Expl	Expl		Expl	Impl	Expl	Expl		Expl	Expl
Charging Strategy																				
– Overnight Charging										✓					✓	✓			✓	✓
– Opportunity Charging	✓	✓		✓	✓		✓		✓	✓	✓	✓		✓		✓	✓		✓	✓
– Full Charging							✓		✓	✓					✓	✓	✓		✓	✓
– Partial Charging	✓	✓		✓	✓					✓	✓	✓		✓		✓	✓		✓	✓
– Integrated Charging Scheduling	✓	✓		✓	✓		✓		✓	✓	✓	✓		✓		✓	✓		✓	✓
– Demand Response / V2G	✓											✓								

🔋 = *Energy Management*, 📍 = *Routing*, 🍃 = *Environmental Impact*

4
Simulation Framework

The purpose of this chapter is not only to describe the model, but to show how the earlier modelling requirements are operationalised. The system definition, input data, conceptual model, simulation logic, operational policies, and uncertainty implementation together form the bridge between the theoretical problem definition and the case-study experiments in Chapter 5 and Chapter 6.

The chapter is structured as follows. Section 4.1 defines the system boundary and main model components. Section 4.2 describes the required input data. Section 4.3 presents the conceptual model structure, while Section 4.4 explains the main simulation logic, including dispatching, task execution, charging, replenishment, dumping, and depot return decisions. Finally, Section 4.5 explains how uncertainty is implemented, and Section 4.6 presents the verification approach used before applying the framework to the case study.

4.1. System Definition

The simulation framework represents day-of-operations eGSE service execution at aircraft stands. The model focuses on the allocation of individual eGSE vehicles to time-window-constrained service tasks under vehicle-availability, energy, onboard-resource, and infrastructure-capacity constraints. Its purpose is to support fleet-sizing decisions and to evaluate operational-control assumptions while providing insight into service reliability, vehicle utilisation, driven distance, charging behaviour, and supporting-infrastructure use.

The system boundary is defined at the level of airside service execution. The model includes service demand at aircraft stands, individual eGSE vehicles, vehicle movements between airport locations, task execution, charging, replenishment or disposal actions, depot return, and the use of shared supporting infrastructure. Upstream processes, such as flight scheduling, air traffic control, stand allocation, and airline planning decisions, are treated as exogenous. Detailed airside traffic interactions, human decision variability, weather disruption, vehicle breakdowns, charger failures, and full turnaround precedence modelling are excluded. The model therefore evaluates operational control and resource allocation within a given demand pattern, rather than optimising the wider airport or airline planning system.

The system consists of three main entity groups: service tasks, eGSE vehicles, and supporting infrastructure. Tasks represent realised service demand at aircraft stands with specific timing, location, service duration, resource requirement, and service-window constraints. Vehicles execute these tasks while operating under dynamic state constraints, including location, availability, SOC, onboard resource level, charging status, and operational status. Supporting infrastructure consists of capacity-constrained resources such as chargers, refill points, dump points, depots, and service locations. These infrastructure elements can create waiting times and affect vehicle availability when multiple vehicles require the same resource.

Although the model is applied to the KLM Schiphol case study, its structure is designed as a generic operational simulation framework. Airport-specific characteristics, such as demand patterns, spatial layout, fleet composition, aircraft compatibility, service rules, charging assumptions, and operational parameters, are introduced through input data. This separation between model structure and case-study input allows the same framework to be adapted to different airports, vehicle groups, and operational assumptions, provided that the required input data and validation evidence are available.

The model structure is applicable to GSE vehicle types that can be represented as mobile service resources assigned to time-window-constrained tasks. Depending on the vehicle type, additional constraints can be activated, such as onboard resource capacity, intermediate replenishment or disposal visits, depot-based operation, stand-side positioning, and vehicle-specific charging behaviour. In this thesis, these mechanisms are implemented and evaluated for three vehicle groups: water vehicles, toilet vehicles, and loaders. Water and toilet vehicles represent depot-based, resource-constrained operations with refill or dump requirements, while loaders represent stand-side service equipment without onboard service-resource constraints. Other GSE types can be represented by the same general modelling principles, but they require additional vehicle-specific input data, service rules, operational logic, and validation before they can be used for fleet-sizing conclusions.

4.2. Input Data

The model integrates multiple input layers, combining flight-based service demand with process-specific service rules and system-level operational parameters. Together, these inputs define both the temporal distribution of demand and the constraints under which vehicles operate. The primary data source is a flight schedule, representing service demand at the level of individual aircraft movements. Each record contains scheduled arrival and departure times, stand assignment, aircraft subtype, operational classification, and information on connected flights. These attributes determine when and where services are required and how flight-level demand is translated into operational service tasks.

Service data define how this demand is operationalised. For each combination of aircraft subtype and operational classification, service rules specify nominal service durations and time-window offsets relative to the scheduled flight time. Operational classifications include arrival, departure, and turnaround movements. The service rules are defined at a detailed level, allowing variation across aircraft types and operational contexts. In addition, process identifiers distinguish between different service types, such as replenishment, disposal, platform service, towing, or aircraft-support tasks.

Resource requirements are derived using a combination of aircraft subtype and aircraft class. The aircraft subtype represents the specific aircraft model and is used to define service characteristics and compatibility rules. Aircraft class, such as commuter, narrow-body, or wide-body, is used to approximate service quantities where detailed aircraft-specific resource data are not available. Depending on the vehicle type, these quantities may represent consumable resources, collected return resources, load-handling demand, towing requirements, or other service-specific quantities. For water and toilet operations, this includes approximations of potable-water demand and lavatory-waste quantities.

Input datasets from different sources are first harmonised and then processed into task-level simulation records. During this preprocessing step, the model derives service duration, time-window constraints, location, aircraft subtype, operation type, and resource requirements for each task. A key preprocessing step is the treatment of connected flights. Turnaround operations are represented as a single logical entity by removing duplicate records and assigning a combined operational classification. This prevents service demand from being double-counted while preserving the correct service logic for arrival, departure, and turnaround operations.

After preprocessing, each instance of service demand is converted into a standardised task record. These records include timing information, aircraft stand, location group, aircraft subtype, service duration, required resource quantities, and service-window constraints. Additional metadata capture the operational context of each task, including whether the task is associated with an arrival, departure, or turnaround operation. The resulting service tasks are then treated as independent simulation tasks. No explicit precedence relations between different ground-handling tasks are imposed beyond their individual service windows. This reflects the model scope: the framework evaluates vehicle availability and task completion, not full turnaround sequencing or aircraft departure-delay causality.

Not all potential tasks are retained. Tasks may be filtered out during preprocessing if they are operationally irrelevant. For example, tasks corresponding to aircraft subtype–operation-type combinations for which no service is required are excluded, and duplicate demand resulting from linked flights is removed as described above. In addition, a deterministic demand-reduction rule can be applied to predefined aircraft subtypes for which water or toilet service is not required after every turnaround. For these subtypes, every second task is removed based on the original input sequence. This rule is not intended to represent stochastic cancellation, but to approximate an observed operational servicing frequency while keeping demand generation reproducible. This assumption is particularly relevant for water and toilet eGSE operations.

Once defined, tasks are released into the simulation prior to their planned start time by a fixed offset. This early-release mechanism allows vehicles to initiate travel in advance, while service execution remains constrained by the defined service window. As a result, dispatching decisions can account for travel time without violating operational constraints at the point of service.

Movement across the airport is represented using an aggregated spatial structure. Individual aircraft stands are assigned to location groups, and travel distances between these groups are defined through a distance matrix. Travel times are then calculated from these distances and vehicle-specific speeds. This abstraction preserves the main spatial differences across the apron while avoiding the complexity of a microscopic apron-traffic model. The case-specific construction of the distance matrix is described in Chapter 5.

Vehicle behaviour is governed by parameters describing fleet composition and operational capability. These include fleet size, vehicle speed, battery capacity, initial SOC, energy consumption, charging behaviour, minimum dispatch thresholds, charging thresholds, and onboard resource limits such as water or waste capacity. Supporting infrastructure is modelled as a set of capacity-constrained resources, including charging stations, refill points, dump points, and depot locations. These resources introduce potential queueing and competition between vehicles, which can affect vehicle availability and system performance.

The model further incorporates stochastic inputs to reflect operational variability. Task-timing uncertainty and travel-time variability are defined through externally specified probability profiles, empirical magnitude arrays, or modelling assumptions. Task-timing uncertainty is applied during preprocessing, before tasks are released into the simulation environment, because it changes the timing of service demand. Travel-time variability is applied dynamically during simulation execution, because it affects individual vehicle movements and therefore subsequent vehicle availability. The uncertainty implementation is described in Section 4.5.

Finally, the model is designed to support controlled simulation experiments. Key parameters, such as demand dataset, vehicle group, fleet size, operational-control policy, charging settings, uncertainty setting, and random seed, can be varied between runs. An experiment configuration is therefore defined as a combination of a fixed demand dataset and a specific parameter configuration. This enables systematic comparison of deterministic baselines, stochastic uncertainty settings, operational-control variants, and sensitivity tests without changing the underlying simulation logic. The resulting input-data structure is summarised in Figure 4.1.

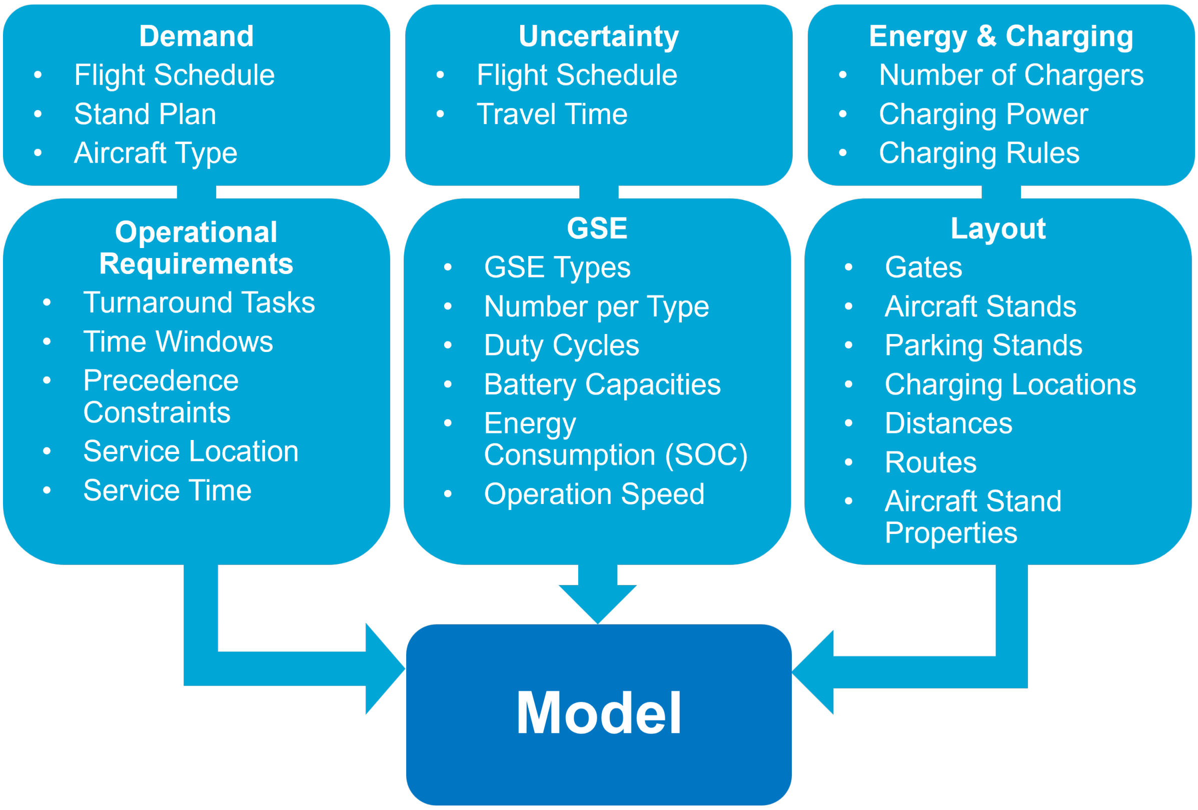


Figure 4.1: *Input-data structure of the simulation framework*

4.3. Conceptual Model

The conceptual model represents the simulation framework as a rule-based, discrete-event service system in which service demand, individual eGSE vehicles, and supporting infrastructure interact over time. Its purpose is to define the main system entities, their state variables, and the high-level decision structure through which service tasks are matched to available vehicles. At this level, the model describes the structure and intended behaviour of the system, while the detailed implementation logic is described in Section 4.4.

The system consists of three main entity groups: service tasks, eGSE vehicles, and supporting infrastructure. Service tasks represent operational demand at aircraft stands. Each task is characterised by a service type, aircraft stand, location group, service duration, required resource quantity, and service time window. The opening of the service window defines the earliest allowable service start, while the closing of the service window defines the latest allowable service completion under the modelled operational requirements. Vehicles represent mobile eGSE resources that can execute compatible tasks. Supporting infrastructure consists of capacity-constrained resources such as chargers, refill points, dump points, depots, hubs, or other service-specific support locations.

The system state evolves through the interaction of these entities. Each vehicle has dynamic state variables, including its current location, availability, SOC, onboard resource level, charging status, and operational status. Operational status can include states such as idle, travelling, servicing, charging, refilling, dumping, or returning to depot. Tasks enter the system over time, remain available while they are operationally relevant, and leave the system once they have been completed or when the simulation reaches an explicit end-of-run or stalemate condition. Infrastructure resources can be occupied by one or more vehicles depending on their capacity, which can create queueing and affect the timing of future vehicle availability.

At a high level, the system follows a recurring operational cycle. Service demand is released into the simulation over time. The dispatcher evaluates the available tasks and vehicles, selects a task to be addressed, identifies a suitable vehicle, and assigns the task when the relevant operational constraints are satisfied. The assigned vehicle then travels to the task location, waits if it arrives before the service window opens, performs the service, and updates its state after task completion. Depending on its post-task state, the vehicle may become available for a new task or may first need to charge, refill, dump, return to depot, or perform another service-specific support action.

The decision structure of the model is organised into four interconnected components: Dispatcher, Task Selection, Vehicle Selection, and Task Execution. The Dispatcher acts as the central coordination mechanism that links released service demand to fleet availability. Task Selection determines which waiting task should be prioritised, based on urgency and operational feasibility. Vehicle Selection determines which available vehicle should execute the selected task, based on compatibility, location, timing, battery state, and service-specific constraints. Task Execution represents the operational realisation of the assignment, including travel, waiting, service delivery, infrastructure interaction, and subsequent vehicle-state changes.

The conceptual model assumes adaptive dispatching. Tasks that cannot currently be assigned remain in the task queue and can be reconsidered when vehicle availability, vehicle state, or infrastructure conditions change. Similarly, vehicles are not modelled only as task-execution resources, but as operational resources that must also maintain their own feasibility through charging, replenishment, disposal, and repositioning. This is important for eGSE fleet sizing because vehicle availability depends not only on task completion, but also on battery state, onboard resources, spatial position, and access to supporting infrastructure.

Supporting infrastructure shapes system performance because it can impose additional capacity constraints on otherwise available vehicles. A vehicle may be idle but not immediately operationally useful if it first needs to charge, refill, dump, or return to a service location. Likewise, a vehicle may require access to a charger or service point that is already occupied by another vehicle. The conceptual model therefore treats infrastructure use and queueing as part of the vehicle-availability problem rather than as external assumptions.

Overall, the conceptual model defines eGSE operations as an event-driven interaction between time-window-constrained service demand, mobile vehicle resources, and capacity-constrained supporting infrastructure. This structure reflects the central modelling assumption of this thesis: required fleet capacity is not determined only by the number of service tasks, but by the dynamic availability of suitable vehicles at the required time and location. The simulation design in Section 4.4 translates this conceptual structure into concrete processes, decision rules, and control logic.

4.4. Simulation Design

The simulation model operationalises the conceptual model as a rule-based Discrete-Event Simulation (DES) in which service tasks, individual eGSE vehicles, and supporting infrastructure interact through time-dependent events. Building on the conceptual structure introduced in Section 4.3, the implementation is organised around four main components: Dispatcher, Task Selection, Vehicle Selection, and Task Execution. This section describes how these components are translated into concrete simulation processes, decision rules, and state transitions. An overview of the overall model logic is shown in Figure 4.2.

At the start of each simulation run, the simulation environment is created, the central task store is initialised, and the fleet is generated. Vehicle initialisation is defined per vehicle group and depends on the service type. Each vehicle is assigned an initial location, which may be a depot, charging location, stand-side hub, or another operational base. For depot-based vehicle groups, the assigned charging locations also serve as operational depots and starting positions. This allows the same simulation structure to represent both depot-based vehicles and vehicles distributed across the apron. The specific starting locations, charging locations, and service-specific base locations used in the KLM Schiphol case are defined in the case-study input configuration.

After vehicle initialisation, the core simulation processes are launched: task generation, stop-schedule control, central dispatching, and one worker process for each eGSE vehicle. The stop-schedule control represents the operational end-of-shift logic of the vehicle group. Each vehicle is implemented as a separate SimPy process that manages its own behaviour, including waiting, travelling, servicing, charging, refilling, dumping, and returning to its base. As a result, the simulation consists of multiple concurrent interacting processes rather than one sequential procedure. System performance therefore emerges from the interaction between task releases, dispatcher decisions, vehicle states, and infrastructure availability.

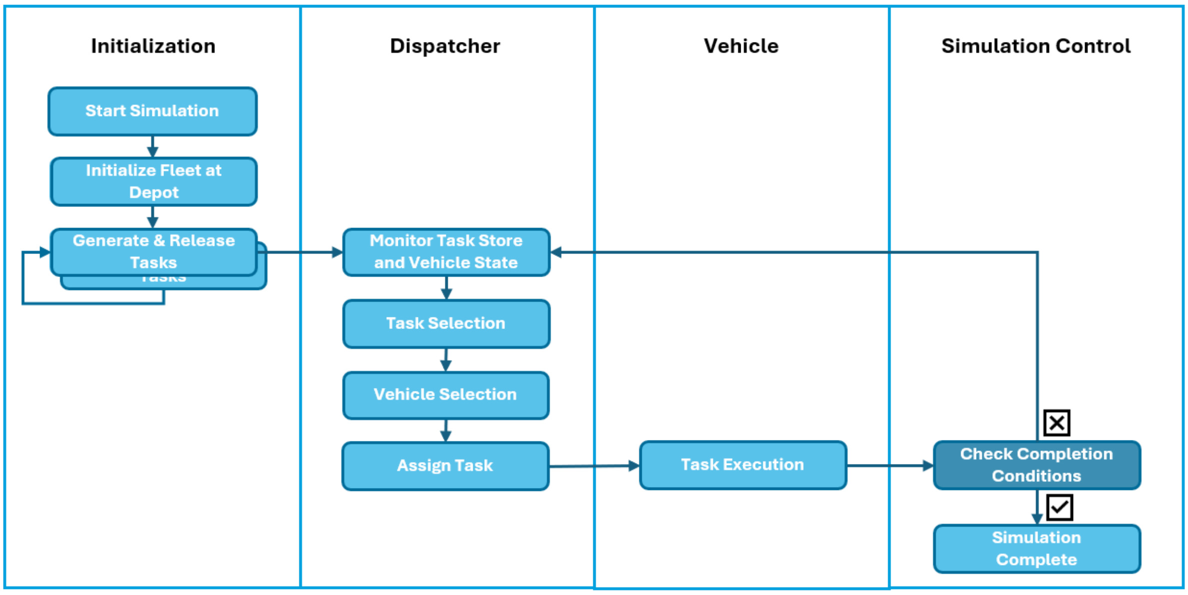


Figure 4.2: *Flowchart of the simulation model.*

Tasks do not enter the system all at once, but are released over time through a dedicated task-generation process. Each task is inserted into the central task store when its release time is reached. To account for the travel time required before service can begin, the model applies an early-release mechanism: tasks become available for dispatch a fixed number of minutes before their planned service-window opening. This allows vehicles to be dispatched in advance, while actual service execution remains constrained by the task-specific service window.

The simulation continues until one of several completion conditions is met. Under normal circumstances, a run ends when all tasks have been generated and completed, no pending assignments remain, the task store is empty, all vehicles have returned to their required end location, and any end-of-shift obligations related to tank filling or emptying have been satisfied. In addition, the model can terminate under end-of-shift conditions when cutoff schedules become binding and no operationally relevant work remains. The model can also terminate under stalemate conditions, where tasks remain in the system but none can still be dispatched because of eligibility restrictions, stop-cutoff rules, or other operational constraints. These termination rules prevent the simulation from continuing indefinitely in states where no further meaningful operational activity can occur.

4.4.1. Dispatcher

The dispatcher is implemented as the central coordination process of the simulation model. It continuously links released service demand to the available fleet and assigns tasks to vehicles when the relevant operational constraints are satisfied. The dispatcher is hybrid in nature. It reacts to relevant events, such as task releases and vehicles becoming available, but also performs periodic retries when no assignment can be made immediately. This ensures that waiting tasks are reconsidered as vehicle availability, vehicle state, or infrastructure conditions change, without requiring continuous polling at a very fine time resolution. The dispatcher logic is summarised in Figure 4.3.

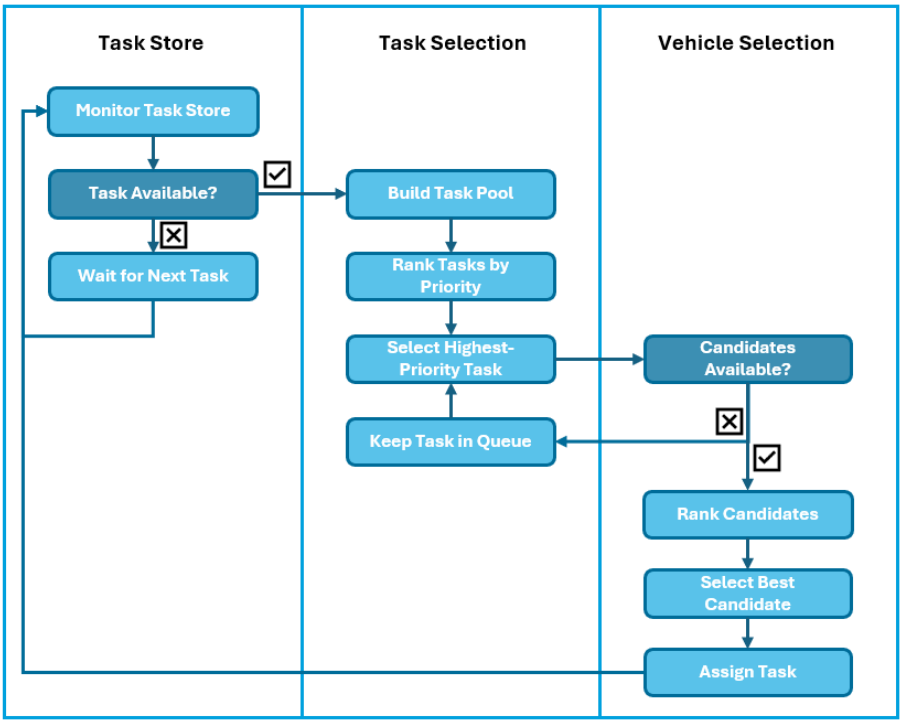


Figure 4.3: *Flowchart of the dispatcher.*

Dispatching is performed centrally for the fleet as a whole rather than independently per vehicle. In each decision round, the dispatcher evaluates tasks that have already been released and are still waiting in the queue. If at least one operationally suitable vehicle is available, the dispatcher assigns one of the queued tasks according to the task-selection and vehicle-selection rules. If no suitable vehicle is available, the task remains in the queue and is reconsidered in a later dispatching round. Task release and task dispatch are therefore not necessarily simultaneous: a task can enter the system before a vehicle is available to execute it.

Tasks are not removed from the system solely because they can no longer be completed on time. If a task remains operationally assignable but its service window can no longer be met, the task may still be executed and is recorded as a late task. This distinction is important because the model evaluates both task completion and on-time completion. Late tasks therefore represent internal service-window violations rather than deleted or unserved demand.

4.4.2. Task Selection

Within the dispatcher, task selection determines which queued task should be considered first. The baseline dispatcher operates under `DISPATCH_PRIORITY_MODE = latest_start`. In this mode, tasks are ranked by the sorting key $(slack_i, -\tau_i^{min})$, where $slack_i$ is the remaining time until the latest feasible service start of task i, and τ_i^{min} is the minimum travel time required to reach that task from the currently available vehicle set. The slack of task i is defined as:

$$slack_i = (t_i^{close} - d_i) - t$$

where t_i^{close} is the service-window closing time of task i, d_i is the task service duration, and t is the current simulation time. The term $(t_i^{close} - d_i)$ represents the latest possible service start if the task is to be completed within its service window, excluding vehicle-specific travel time. A smaller slack value therefore indicates a more urgent task.

If two tasks have equal slack, the tie-break rule selects the task with the larger minimum travel time first. This gives priority to tasks that are harder to reach when their latest feasible service start is otherwise identical. The rule should be interpreted as a local dispatching heuristic rather than a globally optimal scheduling rule. For two tasks with identical unadjusted slack, the task with the larger minimum travel time requires more lead time before service can start. Prioritising this task is therefore a conservative way to reduce the immediate risk of time-window infeasibility for less accessible tasks.

Before ranking is performed, task filtering is applied. Tasks for which no currently available vehicle satisfies the relevant compatibility, state, or operational constraints remain in the queue until the next dispatch cycle. Once a suitable vehicle becomes available, all waiting tasks are reconsidered, and the task with the smallest slack is given priority. The task-selection process used by the dispatcher is shown in Figure 4.4.

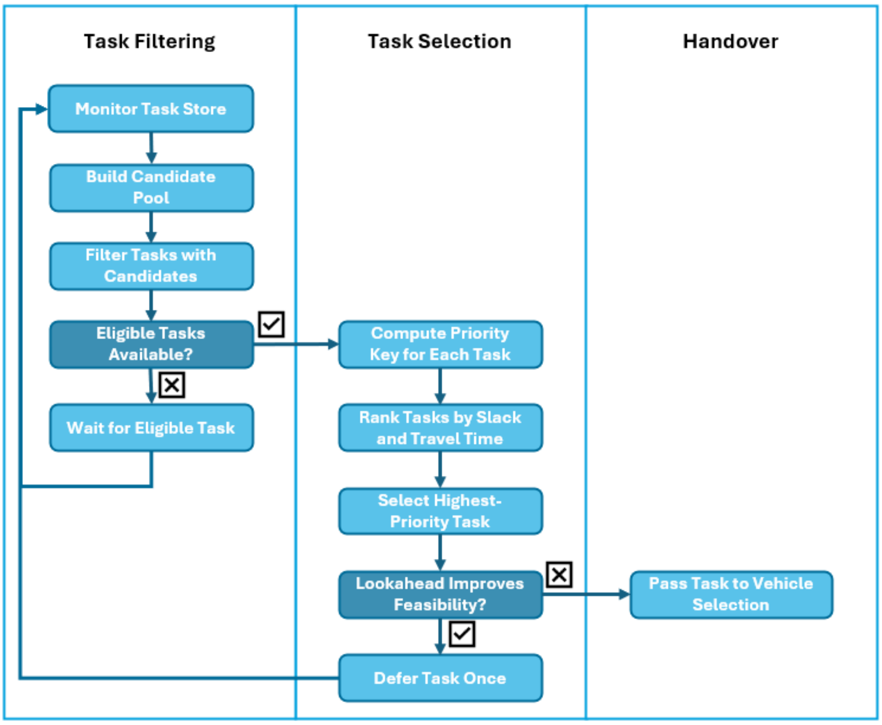


Figure 4.4: *Flowchart of the task-selection process.*

The task-selection logic therefore prioritises tasks based on the latest feasible service start rather than only on the service-window closing time. A task with a later window close can still be more urgent if it has a longer service duration. The implementation also includes a lookahead-based deferral rule. If the currently available vehicle cannot complete a task within its service window, but another vehicle is expected to become available shortly and can complete the task more feasibly, the dispatcher may defer assignment. This prevents premature assignment to an unsuitable vehicle when a better feasible assignment is expected within the short lookahead period.

4.4.3. Vehicle Selection

Once a task has been selected, the vehicle-selection component determines which vehicle should execute it. The dispatcher first constructs a candidate pool from vehicles that are operationally available for assignment. Candidate vehicles are then screened using operational constraints, including vehicle availability, SOC, depot-return status, reassignment restrictions, and service-specific constraints. For resource-constrained vehicles, candidates may be excluded if the task would violate onboard capacity limits, replenishment requirements, disposal requirements, or other vehicle-specific policy checks. The vehicle-selection process is shown in Figure 4.5.

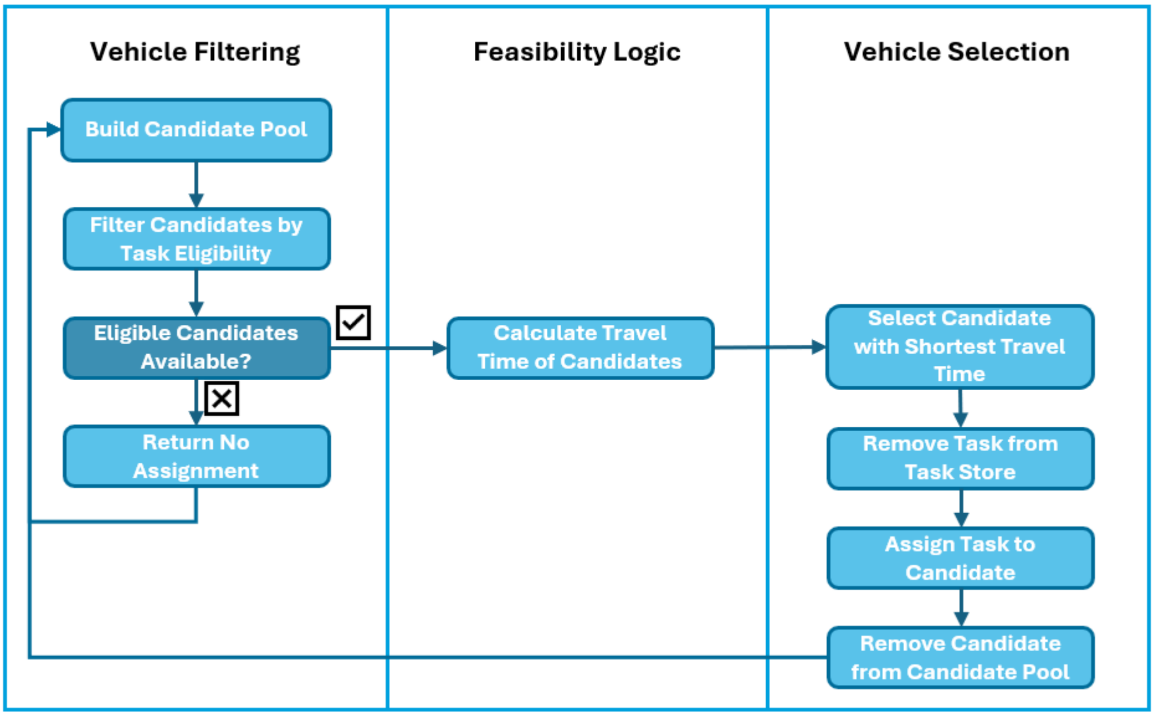


Figure 4.5: *Flowchart of the vehicle-selection process.*

Among the eligible candidates, vehicle selection is based on travel efficiency and service-window feasibility. The model first checks whether each candidate vehicle can travel to the task location, wait if necessary until the service window opens, perform the service, and complete the task within the service window. If one or more vehicles can complete the task on time, the dispatcher selects the feasible vehicle with the shortest travel time.

If no eligible candidate can complete the task within the service window, the task is not automatically discarded. Instead, the dispatcher can still assign the task to an operationally eligible vehicle when service execution remains possible, even though completion will be late. In that case, the task is completed and the service-window violation is recorded in the output metrics. This mechanism allows the model to distinguish between insufficient punctuality and non-execution of tasks.

4.4.4. Task Execution

Task execution is implemented through one worker process per vehicle. Each vehicle alternates between idle and active states. At the start of the simulation, each vehicle is idle at its assigned starting location. When idle, the vehicle waits for an assignment while also checking whether operational support actions should be initiated. These actions may include charging, refilling, dumping, stop-cutoff handling, depot return, or other service-specific reset actions. The task-execution process, including travel, service execution, and vehicle-state updates, is shown in Figure 4.6.

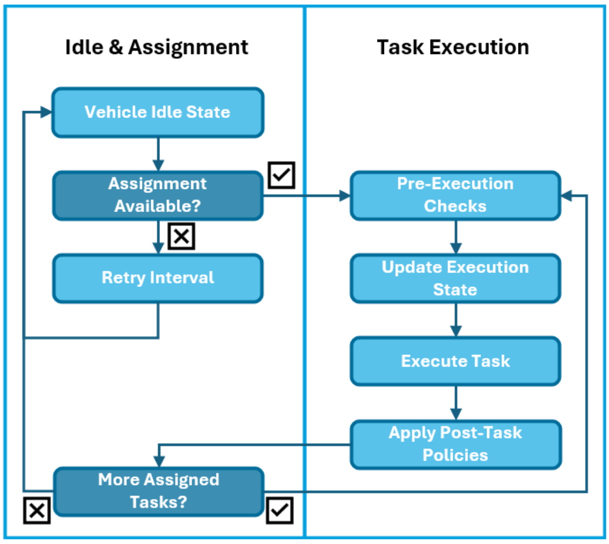


Figure 4.6: *Flowchart of the task-execution process.*

After receiving an assignment, the vehicle travels to the task location. If it arrives before the service window opens, it waits until the opening time. Service starts only after the vehicle has arrived and the service window is open. The standard execution sequence is therefore `TASK_ARRIVE`, `GSE_DISPATCH`, `GSE_ARRIVE`, `TASK_START`, and `TASK_DONE`. This sequence ensures that travel occurs before service, that service does not start before vehicle arrival, and that tasks assigned to the same vehicle cannot overlap.

After task completion, vehicles do not simply return to a generic available state. Instead, post-task behaviour is governed by a sequence of operational checks. Stop-cutoff conditions may prevent the vehicle from accepting new work. Resource-related actions, such as refilling, unloading, dumping, or other service-specific reset actions, may be initiated when threshold conditions are met. If the SOC is below the relevant charging threshold, the vehicle may be sent to charge. Depot-return logic can also direct the vehicle back to its depot or base location when no immediate assignment is available. In the baseline logic, this depot-return action can be triggered after an idle period of at least 1 min following a dispatcher waiting period of 1 min without assignment.

This post-task logic makes vehicle behaviour endogenous to the system state. Vehicle availability is therefore not determined only by whether the previous task has been completed, but also by battery state, onboard resource level, infrastructure access, shift constraints, and repositioning decisions. This is central to the fleet-sizing logic of the model: a fleet is sufficient only if suitable vehicles are available at the required time and location, not merely if the total number of vehicles is large enough.

4.4.5. Operational Policies and Constraints

Several cross-cutting operational policies and constraints shape the behaviour of the simulation model. These policies determine when vehicles remain available for task assignment, when they must interact with supporting infrastructure, and when they are temporarily removed from the assignable fleet. They therefore affect vehicle availability directly and are central to the interpretation of fleet-size requirements.

Charging is explicitly modelled as an operational process within the simulation horizon, rather than as an external overnight assumption. Vehicles can enter charging events during the simulated day based on their SOC, location, and operational status. Charging can occur reactively when SOC falls below an operational threshold. Vehicles located at their depot or assigned charging location can also charge opportunistically when they are idle and below a higher depot-charging threshold. Charging behaviour is vehicle-group specific and location-dependent. Because chargers are represented as capacity-constrained resources, a vehicle may have to wait before charging can begin. During charging and charging-related waiting, the vehicle is not available for task execution unless the charging process is interrupted by the operational logic.

Charging is represented through rule-based control rather than through a separate charging-schedule optimisation. The model therefore does not optimise electricity cost, battery degradation, grid interaction, or charger assignment at fleet level. Instead, charging is included as an operational availability constraint: it affects whether a vehicle can remain available for service, whether it needs to leave the task-assignment pool, and whether limited charger capacity can create waiting time.

Onboard resource replenishment and disposal are also modelled through threshold-based rules. Vehicles with consumable onboard resources, such as potable-water vehicles, visit a replenishment point when their remaining resource level falls below a predefined threshold. Vehicles that collect return resources, waste, or other service-specific loads visit a disposal or unloading point when the onboard load exceeds a predefined share of capacity. Where required, vehicles may also be forced to replenish or dispose of resources before the end of the shift, so that they return in the required operational state. These rules ensure that onboard resource levels are actively managed during the simulation, rather than being treated only as passive bookkeeping variables.

Supporting infrastructure is represented as a set of shared resources with limited capacity. Charging stations, refill points, dump points, depots, and other service-specific support locations can only be used when capacity is available. If the relevant infrastructure is occupied, vehicles wait in a queue. Charging locations are modelled with location-specific capacities, allowing multiple vehicles to charge simultaneously when sufficient chargers are available. Replenishment and disposal infrastructure is modelled in the same way using predefined capacity limits. Through this structure, the model captures congestion and waiting effects when multiple vehicles compete for the same operational resource.

Task-assignment feasibility is determined by compatibility, timing, vehicle state, and service-specific constraints. A vehicle can only be considered for a task when it is compatible with the task type, operationally available, and able to execute the task within its applicable constraints. Vehicles may be excluded from the candidate pool because of insufficient SOC, stop-cutoff status, depot-return status, reassignment restrictions, incompatible service type, or vehicle-specific onboard-resource constraints. These conditions ensure that dispatching decisions reflect both the service requirements of the task and the operational limitations of the vehicle.

A distinction is made between operational eligibility and on-time feasibility. Operational eligibility determines whether a vehicle is allowed to execute a task at all. On-time feasibility determines whether the task can still be completed within its service window. If no eligible vehicle can complete a task on time, the task is not necessarily removed from the system. It can still be executed by an eligible vehicle and recorded as a late task. This distinction allows the model to measure service-window violations explicitly, rather than hiding them through task deletion.

Stop-cutoff logic represents the operational end of the working period for a vehicle group. After the cutoff becomes active, vehicles may stop accepting new regular assignments and may instead complete required end-of-shift actions, such as returning to depot, refilling, dumping, or charging. This prevents vehicles from being assigned to tasks beyond the modelled operating period and ensures that end-state requirements are represented consistently.

Finally, the simulation design is parameterised to support controlled simulation experiments. Key configurable parameters include fleet size, vehicle-group settings, charging thresholds, minimum dispatch thresholds, infrastructure capacities, uncertainty toggles, random seeds, and operational-control policies. These parameters are defined through configuration files and environment settings, allowing experiments to be varied systematically without changing the underlying simulation logic. As a result, the model can represent a baseline operating system and test how changes in fleet composition, charging assumptions, infrastructure capacity, uncertainty settings, and operational policies affect system performance.

4.5. Uncertainty Modelling

This section describes how uncertainty is represented in the simulation model. The uncertainty structure is generic: stochastic inputs can be introduced to affect task timing, travel time, service duration, energy consumption, charging behaviour, or infrastructure availability. In this thesis, however, the active stochastic implementation is limited to task-timing uncertainty and travel-time uncertainty. These two mechanisms are selected because they directly affect when service demand occurs and when vehicles become available for subsequent tasks.

The section is organised in three parts. First, Subsection 4.5.1 describes the general uncertainty mechanisms used in the simulation framework. Second, Subsection 4.5.2 explains how the task-timing profiles are derived for the KLM Schiphol case using historical flight data. Third, Subsection 4.5.3 describes the implementation of stochastic replications, random seeds, and confidence intervals. The resulting uncertainty parameters should therefore be interpreted as case-study inputs for this thesis, not as universal airport-wide uncertainty parameters.

4.5.1. Sources and Modelling of Uncertainty

Two active sources of uncertainty are included in the simulation experiments: task-timing uncertainty and travel-time uncertainty. Task-timing uncertainty changes when service tasks occur during the simulated day. Travel-time uncertainty changes how long individual vehicle movements take during simulation execution.

Task-timing uncertainty covers both late and early arrivals. A service task can therefore be shifted later or earlier than originally planned. Task cancellations due to uncertainty are not modelled explicitly. This modelling choice is conservative for fleet sizing, because the analysis aims to evaluate system performance under realised demand. Including stochastic cancellations would reduce the effective task volume and could therefore underestimate the required fleet size and supporting-infrastructure capacity.

Task-timing uncertainty is modelled by shifting the planned timing of service tasks linked to an aircraft arrival. For each relevant flight leg, the model first samples whether the arrival is late, early, or on time using empirically estimated probabilities. If a late or early arrival is sampled, the corresponding magnitude is drawn through uniform resampling from a capped empirical magnitude array for the relevant time-of-day bucket. The sampled magnitude is rounded to whole minutes and applied as a signed shift: positive for a late arrival and negative for an early arrival. The relevant time-of-day bucket is based on the scheduled local time of the flight event. The dayparts are configurable. In the case-study implementation, the morning period is defined as before 11:00, the afternoon as 11:00–17:00, and the evening as after 17:00. This segmentation allows the uncertainty profiles to reflect differences in arrival-time variability across the operating day.

Only arrival timing is used as an active stochastic input in the simulation. This is because arrival timing directly determines when aircraft-dependent ground-handling tasks can start. Departure delay, by contrast, does not necessarily imply delayed service execution, because the relevant ground service may already have been completed before the departure delay occurs. Consequently, departure-delay profiles are not used to shift service tasks in the stochastic experiments. Timing deviations are applied at the flight-arrival level. When an arrival is sampled as late or early, the planned arrival time is shifted by the sampled value. The service tasks linked to that arrival are then shifted by the same amount. Both the planned task timing and the opening of the service window move accordingly. Because the closing time of the service window is defined relative to the opening time, the full service window shifts while its duration remains unchanged. Task-timing uncertainty therefore changes the temporal demand pattern during the day, but it does not change service duration or intrinsic task characteristics.

Travel-time uncertainty is modelled as a stochastic multiplicative increase of deterministic travel times. For each travel leg, a Bernoulli trial determines whether a travel disturbance occurs. The disturbance probability is set to $p = 30\%$. If a disturbance is triggered, the deterministic travel time, calculated from distance and vehicle speed, is increased by 15%. If no disturbance is triggered, the travel time remains unchanged. These travel-time uncertainty parameters are not estimated directly from empirical travel-time observations. Instead, they are selected in consultation with operational experts as a plausible order of magnitude for incidental apron movement disturbances, such as congestion, routing inefficiencies, waiting time, or interaction with other apron traffic. Travel-time uncertainty is applied independently to each vehicle movement during simulation execution. Its effect is dynamic: if a travel leg takes longer, the corresponding eGSE vehicle remains in transit for a longer period through the simulation timeout mechanism. As a result, vehicle availability is delayed. This can affect subsequent dispatching decisions, queue formation, task completion times, and service-window compliance.

4.5.2. Case-Specific Empirical Derivation

This subsection instantiates the generic task-timing uncertainty logic for the KLM Schiphol case. It does not define a general airport-wide uncertainty model. The delay parameters are derived from historical Schiphol flight data using local timestamps. The dataset consists of approximately 74,000 flight records over a three-month period covering May, June, and July of 2025. This provides a case-specific empirical basis for estimating arrival-time variability.

For each operated arrival leg, the signed arrival deviation d_i is calculated as the difference between actual and scheduled arrival time in minutes. Positive values indicate late arrivals, while negative values indicate early arrivals. Early arrivals are included because they can shift service demand into an earlier part of the day and thereby affect vehicle availability in the same way as late arrivals can shift demand later.

Late-arrival and early-arrival probabilities are estimated separately. Cancelled flights are excluded from both the numerator and denominator, because the baseline simulation represents operated demand and does not model stochastic flight cancellation. A minimum threshold of $\tau = 1$ minute is used to distinguish meaningful timing deviations from negligible timestamp noise. For a given daypart, the late-arrival probability is defined as:

$$p^{late} = \frac{\#\{i \in \mathcal{O} : d_i \geq \tau\}}{\#\{i \in \mathcal{O}\}}, \tag{4.1}$$

where $\mathcal{O}$ is the set of operated, non-cancelled arrival legs in the corresponding daypart. The early-arrival probability is defined as:

$$p^{early} = \frac{\#\{i \in \mathcal{O} : d_i \leq -\tau\}}{\#\{i \in \mathcal{O}\}}. \tag{4.2}$$

The remaining probability mass is treated as on-time or negligible timing deviation:

$$p^{on-time} = 1 - p^{late} - p^{early}. \tag{4.3}$$

Across the analysed time-of-day buckets, the empirical late- and early-arrival probabilities range approximately between 35.0% and 60.7%. The resulting arrival-delay probabilities are shown in

Figure 4.7.

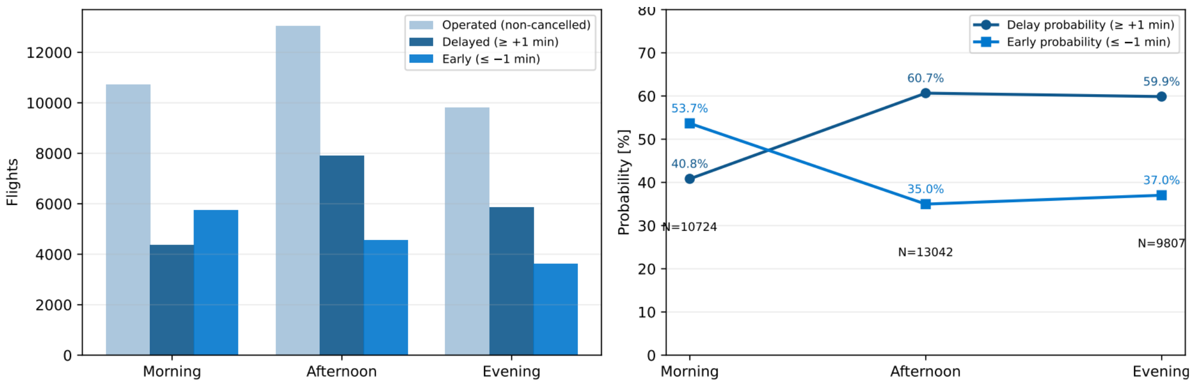


Figure 4.7: *Arrival-timing probability profiles derived from the empirical dataset*

Arrival-delay magnitudes are not represented by fitted parametric distributions. Instead, empirical magnitude arrays are constructed separately for late and early arrivals in each daypart. For late arrivals, the empirical magnitude array is constructed from observations satisfying $d_i \geq 1$, with magnitude $x_i = d_i$. For early arrivals, the array is constructed from observations satisfying $d_i \leq -1$, with positive magnitude $x_i = |d_i|$. In both cases, the empirical magnitude array is denoted by:

$$\mathcal{X} = \{x_1, x_2, \ldots, x_n\}. \tag{4.4}$$

Before percentile capping is applied, observations with an absolute signed deviation larger than 1440 minutes are excluded as a sanity filter against likely timestamp artefacts:

$$|d_i| > 1440. \tag{4.5}$$

This filter removes only a very small share of observations. For late arrivals, 31 out of 18,260 delayed observations exceeded 1440 minutes and were excluded, corresponding to 0.17% of the delayed-arrival sample. These excluded values ranged from 1,481 minutes to 87,853 minutes, with a median of 43,254 minutes, or approximately 30 days. For early arrivals, 5 out of 13,947 observations were smaller than −1440 minutes and were excluded, corresponding to 0.04% of the early-arrival sample. These observations had unrealistic magnitudes of approximately 209,000 to 298,000 minutes, with a median of approximately 252,000 minutes. These values are interpreted as timestamp artefacts rather than realistic operational deviations. The purpose of the 1440-minute filter is therefore to remove extreme data-quality errors without materially altering the normal empirical delay distribution.

After this sanity filter, the magnitude distributions are capped at the 99th percentile. For each daypart and arrival category, the cap is defined as:

$$c_{99} = P_{99}(\mathcal{X}), \tag{4.6}$$

where P_{99} denotes the 99th percentile of the filtered empirical magnitude array. Each magnitude is then transformed as:

$$\tilde{x}_i = \min\{x_i, c_{99}\}, \qquad i = 1, \ldots, n. \tag{4.7}$$

This procedure retains all observations in the filtered empirical array, while replacing magnitudes above the 99th percentile with the daypart- and category-specific cap value. The capped empirical arrays are exported for simulation and are also used to compute descriptive profile statistics. The capped mean is computed as:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} \tilde{x}_i, \tag{4.8}$$

the capped standard deviation as:

$$\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (\tilde{x}_i - \bar{x})^2}, \tag{4.9}$$

and the capped maximum as:

$$x_{\max} = \max_{i=1,\ldots,n} \tilde{x}_i. \tag{4.10}$$

These statistics describe the capped empirical magnitude distributions. They are not used to parameterise a fitted distribution. During simulation, magnitudes are sampled directly from the capped empirical arrays through uniform resampling. This preserves the empirical distributional shape of the historical data, while limiting the effect of extreme tail observations and potential timestamp artefacts.

For example, in the morning late-arrival sample, the filtered empirical array contained $n = 4371$ operated delayed arrival legs. The 99th-percentile cap was approximately 195.9 minutes. In total, 44 observations exceeded this threshold, including one observation of 1408 minutes. The capped mean was approximately 18.0 minutes, compared with a raw mean of approximately 20.5 minutes. The mean of the values above the 99th percentile was 444.1 minutes, with a median of 306.0 minutes. This indicates that the capped tail does not consist of marginal overshoots just above the cap, but of a small number of severe arrival delays and extreme observations.

The empirical arrival-time uncertainty distributions are shown both before and after filtering. The raw late-arrival and early-arrival distributions are shown in Figure 4.8 and Figure 4.10. The filtered late-arrival and early-arrival distributions used as simulation input are shown in Figure 4.9 and Figure 4.11.

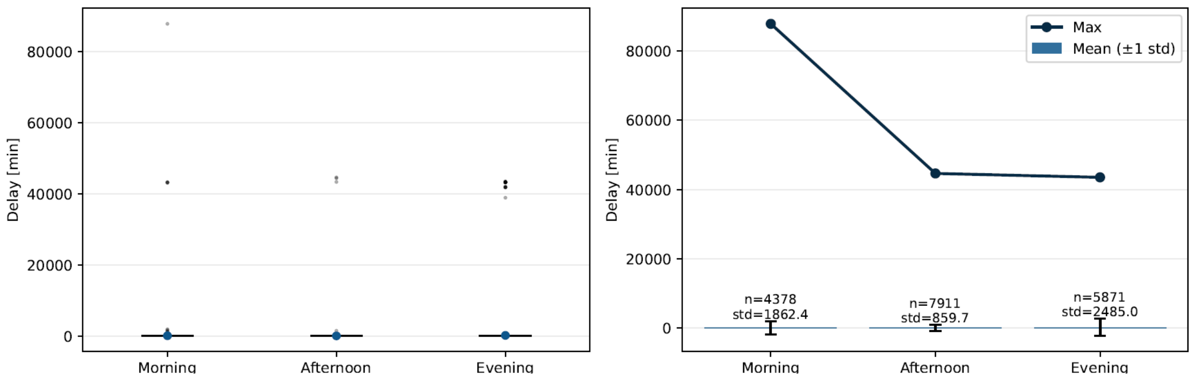


Figure 4.8: *Raw arrival-delay distribution derived from the empirical dataset*

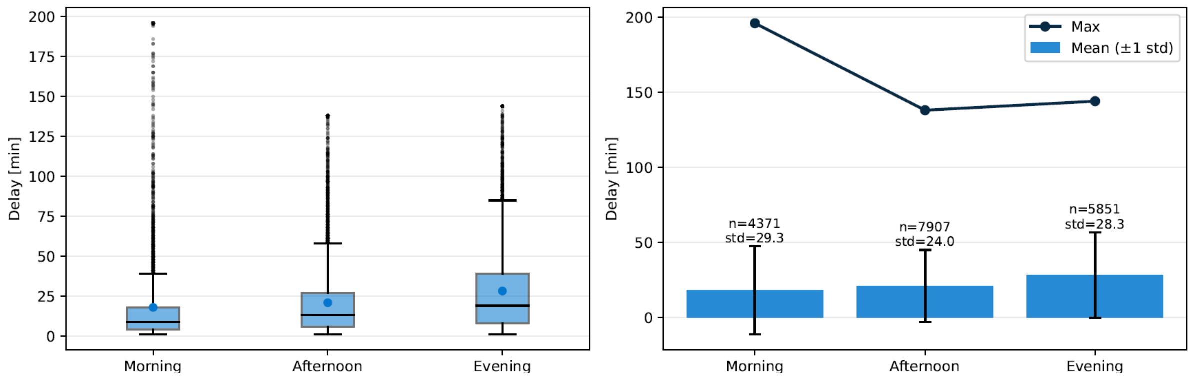


Figure 4.9: *Filtered arrival-delay distribution used as simulation input*

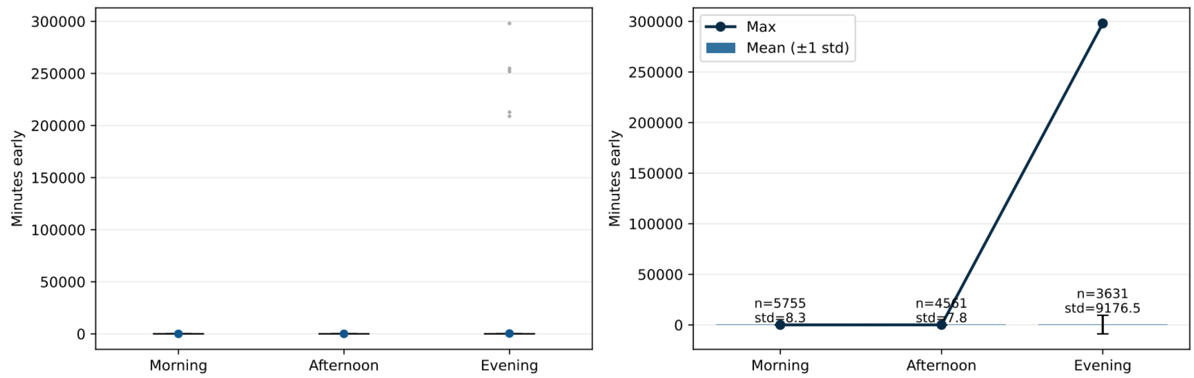


Figure 4.10: *Raw early-arrival distribution derived from the empirical dataset*

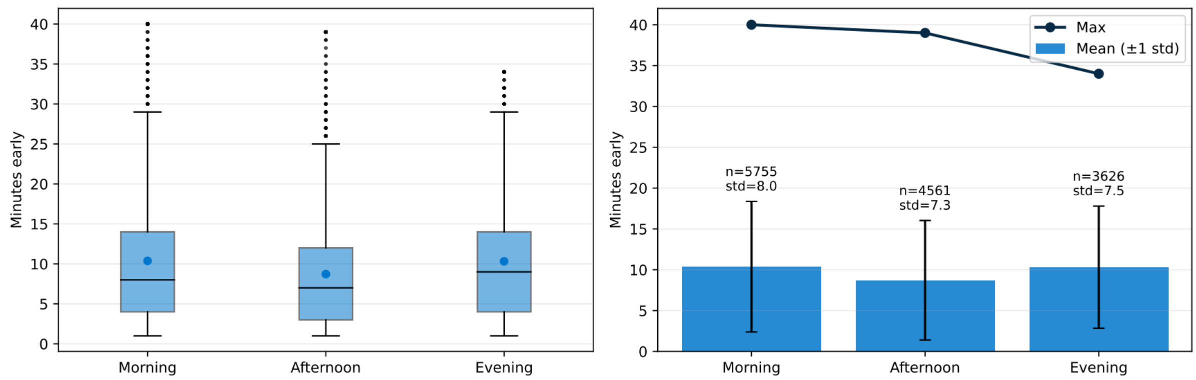


Figure 4.11: *Filtered early-arrival distribution used as simulation input*

The stochastic parameters are externally defined and provided to the simulation through configuration settings. The simulation itself does not estimate or calibrate the delay profiles during runtime. Instead, it relies on precomputed probability profiles and capped empirical magnitude arrays. Task-timing uncertainty is sampled independently for each arrival leg given its time-of-day profile, and travel-time uncertainty is sampled independently for each vehicle movement. Within the task-timing mechanism, occurrence and magnitude are treated as separate stochastic processes.

4.5.3. Implementation

Uncertainty is applied at different stages of the simulation. Task-timing uncertainty is applied during preprocessing, before tasks are released into the simulation environment. It therefore modifies the temporal demand structure before simulation execution starts. Travel-time uncertainty is applied dynamically during simulation execution. It affects individual vehicle movements in real time and can therefore delay subsequent vehicle availability.

The inclusion of uncertainty is controlled through independent configuration toggles. Task-timing uncertainty and travel-time uncertainty can each be enabled or disabled separately. This allows the model to represent a deterministic baseline, task-timing uncertainty only, travel-time uncertainty only, and combined uncertainty. The same underlying demand dataset and model logic can therefore be evaluated under different uncertainty settings.

All stochastic runs are controlled through fixed random seeds. A single seed is used per run and governs all random draws within that run. This ensures that identical configurations produce identical outcomes, which is necessary for reproducibility and controlled experiment comparison. Different seeds are used across replications to evaluate the variability of system performance under repeated stochastic realisations.

For the stochastic experiments, each configuration is evaluated using 30 independent replications with different random seeds. This number of replications was selected for both methodological and practical reasons. First, each replication represents one complete operational sample path of the simulated day, rather than one independent task observation. Tasks within a single run are not independent, because they are linked through shared vehicle availability, dispatching decisions, queueing effects, charging behaviour, and infrastructure use [64]. Second, task-timing uncertainty is sampled from empirical arrival-delay profiles. The historical dataset defines the input distributions, while the replications evaluate how repeated samples from these profiles propagate through the operational system. Third, validation of the uncertainty representation shows that the aggregated behaviour over 30 replications reproduces the direction and approximate scale of the empirical late- and early-arrival profiles sufficiently well for the purpose of the simulation experiments. Finally, 30 replications provide a practical balance between statistical precision and computational effort, because each experiment is repeated across multiple vehicle groups, fleet sizes, uncertainty settings, operational-control variants, and sensitivity cases.

The adequacy of 30 replications is assessed in two ways. First, the sampled uncertainty profiles are compared with the empirical input profiles in the validation section. Second, replication-level 95% confidence intervals are calculated for the main stochastic fleet-sizing thresholds in the result synthesis. These checks indicate whether the selected number of replications provides sufficiently stable estimates for comparing the tested fleet-size thresholds. They do not imply a formal guarantee of robustness over all possible operational realisations.

For each vehicle group, uncertainty setting, and evaluated fleet size, on-time completion rates are first calculated at replication level. The replication-level outputs are then used to calculate sample means, sample standard deviations, and 95% confidence intervals. Confidence intervals are used to quantify the precision of the estimated mean performance across a finite number of stochastic replications. They should not be interpreted as evidence that a selected fleet size is robust under all possible uncertainty realisations. Confidence intervals are a standard method in simulation-output analysis, and the confidence-interval half-width can be used to assess whether the number of replications provides sufficient precision for the output of interest [33]. For a performance indicator Y, the confidence interval is calculated as:

$$\bar{Y} \pm t_{0.975,n-1} \frac{s}{\sqrt{n}}, \tag{4.11}$$

where $\bar{Y}$ is the sample mean across replications, s is the sample standard deviation, and $n = 30$ is the number of replications. The term $t_{0.975,n-1}$ corresponds to a two-sided 95% confidence interval with $\alpha = 0.05$, following the standard Student-t formulation for simulation-output means [64].

The confidence intervals are interpreted as approximate statistical intervals and are bounded by the logical range of 0–100% when reported for on-time completion. They are especially relevant for the pragmatic service-level interpretation, where fleet sizes are evaluated based on near-perfect average on-time performance across stochastic replications.

The strict robust threshold is evaluated separately from the confidence interval of the mean. A fleet size satisfies the strict robust criterion only if it achieves 100% on-time task completion in all 30 sampled stochastic replications. This is a count-based condition, not a mean-performance condition. A narrow confidence interval for the mean on-time rate therefore does not by itself prove strict robustness. The strict robust threshold should be interpreted as robustness across the sampled stochastic realisations, not as a formal guarantee over all possible operational realisations.

4.6. Verification

Verification assesses whether the implemented simulation model behaves according to its intended logic. It is therefore distinct from validation. Verification checks whether the model is built correctly, while validation assesses whether the model is sufficiently representative of the real case.

The verification is structured around targeted test cases that isolate specific parts of the simulation logic. The tasks, timestamps, flight numbers, aircraft types, stands, and service characteristics used in these verification cases are synthetic and should not be interpreted as representative operational data. They are deliberately constructed to trigger specific model behaviours, such as task generation, dispatching, vehicle selection, task execution, charging, refilling, dumping, queueing, depot return, stop-cutoff handling, and uncertainty propagation.

For each verification case, event-log fragments are used as evidence. These logs are then interpreted to assess whether the observed model behaviour matches the expected logic. Although the verification section is presented before the case-study results, it verifies generic implementation logic using simplified and, where required, case-informed test settings. The purpose is not to prove that the case-study assumptions are realistic, but to show that the implemented simulation mechanisms operate consistently with the model design described in Chapter 4.

4.6.1. Input Data and Preprocessing

Before analysing the dynamic behaviour of the simulation model, the input-data pipeline and preprocessing logic were verified. This step checks whether the model correctly reads, maps, filters, and transforms input data into internally consistent simulation objects. It does not validate whether the case-study input values are operationally representative; that assessment is part of the validation in Section 5.6.

Although the verification examples use water-service vehicles as the illustrative resource-constrained vehicle type, the verified mechanisms represent generic model logic. These mechanisms include task generation, dispatching, vehicle assignment, task execution within service windows, travel and service sequencing, onboard resource depletion, threshold-based replenishment, charging, infrastructure queueing, stop-cutoff handling, depot return, redeployment, and uncertainty propagation.

First, the integrity of the input data was checked. Flight records were read and parsed correctly, and all time-related fields were interpreted consistently in local time. The transformation from flight records to service tasks was verified by checking that relevant attributes, including aircraft subtype, stand location, operation type, scheduled timing, release time, service-window opening, and service-window closing, were correctly transferred to the task representation. The release times and corresponding service-window openings were confirmed to be consistent with the defined early-release offset.

Second, the correctness of task-generation rules and parameter mappings was verified. Service durations and service-window parameters were checked against the values defined in the service dataset. Aircraft-dependent service quantities, such as potable-water demand, were confirmed to be assigned according to aircraft subtype and classification rules. The mapping from aircraft stands to location groups was also verified, as were the inter-group distance values used to calculate travel times. Travel times derived from these distances were confirmed to follow the implemented distance, speed, minimum-time, and rounding logic.

Third, the operational constraints embedded in the input configuration were checked. Battery capacity, initial SOC, energy consumption, onboard water capacity, refill thresholds, charging thresholds, and charger parameters were confirmed to be read and applied consistently. Charging characteristics, such as charging power, charging rate, and resulting SOC increase, were checked against the configured values. Minimum travel-time constraints and distance-related constraints were also confirmed to be correctly implemented.

Overall, the verification confirms that the input data are correctly read, processed, and transformed into a valid and internally consistent set of simulation inputs. This provides the foundation for verifying the dynamic simulation behaviour in the following test cases.

4.6.2. Verification Case A — Core Logic

The core logic of the model was verified using small, manually designed test cases with artificial flight schedules. These cases were constructed to isolate specific decision rules and to make the expected model behaviour directly visible in the event log. The verification focused on five main elements: task generation, dispatching, task selection, vehicle selection, and task execution.

Task Generation

Task generation was verified by checking whether valid tasks became active at the expected release moments and whether filtered tasks were excluded. Table 4.1 shows that regular tasks generated `TASK_ARRIVE` events, while cancelled and zero-service tasks were removed. This verifies the task-release rule, the zero-service filter, the cancellation logic, and the handling of connected flights.

Table 4.1: *Event log fragment for task-generation verification*

Time	A/D time	TW open	TW close	Event	Flight	Aircraft	A/D	Stand	Service time (min)
10:49	12:25	11:00	11:43	TASK_ARRIVE	KL0001	788	D	C13	28
11:00	11:10	11:11	11:50	TASK_ARRIVE	KL0002	73H	AD	A54	5
11:10	–	–	–	TASK_CANCELLED	KL0003	73H	AD	B16	5
12:29	–	–	–	TASK_CANCELLED	KL0004	32Q	A	A54	0

Dispatching

Dispatching was verified by checking whether released tasks remained available until a feasible vehicle could be assigned. Table 4.2 shows both immediate dispatching and delayed dispatching after depot return. The delayed assignment of task KL0005 verifies that the task remained in the queue until the vehicle completed its depot return and cooldown period.

Table 4.2: *Event log fragment for dispatch verification*

Time	A/D time	TW open	TW close	Event	Vehicle	Flight	Aircraft	A/D	Stand	Service time (min)
10:49	12:25	11:00	11:43	TASK_ARRIVE	–	KL0001	788	D	C13	28
10:50	12:25	11:00	11:43	GSE_DISPATCH	WATER_1	KL0001	788	D	C13	28
14:38	–	–	–	GSE_ARRIVE_DEPOT	WATER_2	–	–	–	E24	–
14:45	14:40	14:56	15:28	TASK_ARRIVE	–	KL0005	332	A	A54	16
14:54	14:40	14:56	15:28	GSE_DISPATCH	WATER_2	KL0005	332	A	A54	16

Task Selection

Task selection was verified using a dedicated synthetic verification case. The dispatcher operated under `DISPATCH_PRIORITY_MODE = latest_start`. In this mode, task priority is determined by the sorting key $(slack, -min\_travel)$, where $slack = (window\_close - service\_time) - now$. This means that the dispatcher first selects the task with the smallest remaining slack. If two tasks have equal slack, the task with the larger minimum travel time is selected first.

Four checks were used to verify this logic. Table 4.3, Table 4.4, and Table 4.5 verify the basic task-priority rules. The first check confirms that, with equal release time and service time, the task with the earlier window close is selected first. The second check confirms that, with equal release time and window close, the task with the longer service duration is prioritised because it has less remaining slack. The third check confirms the tie-break rule: when slack is equal, the task with the larger minimum travel time is selected first.

Table 4.3: *Event log fragment for task-selection verification based on window close*

Time	A/D time	TW open	TW close	Event	Vehicle	Flight	Aircraft	A/D	Stand	Service time (min)
16:31	16:26	16:42	17:14	`TASK_ARRIVE`	–	KL0006	332	A	C13	16
16:31	16:26	16:42	17:06	`TASK_ARRIVE`	–	KL0007	332	A	C13	16
16:33	16:26	16:42	17:06	`GSE_DISPATCH`	WATER_1	KL0007	332	A	C13	16
16:59	16:26	16:42	17:14	`GSE_DISPATCH`	WATER_1	KL0006	332	A	C13	16

Table 4.4: *Event log fragment for task-selection verification based on service time*

Time	A/D time	TW open	TW close	Event	Vehicle	Flight	Aircraft	A/D	Stand	Service time (min)
18:31	18:26	18:42	19:14	`TASK_ARRIVE`	–	KL0008	332	A	C13	16
18:31	18:26	18:42	19:14	`TASK_ARRIVE`	–	KL0009	788	A	C13	28
18:33	18:26	18:42	19:14	`GSE_DISPATCH`	WATER_1	KL0009	788	A	C13	28
19:11	18:26	18:42	19:14	`GSE_DISPATCH`	WATER_1	KL0008	332	A	C13	16

Table 4.5: *Event log fragment for task-selection verification based on minimum travel time under equal slack*

Time	A/D time	TW open	TW close	Event	Vehicle	Flight	Aircraft	A/D	Stand	Service time (min)
20:31	20:26	20:42	21:14	`TASK_ARRIVE`	–	KL0010	332	A	A54	16
20:31	20:26	20:42	21:14	`TASK_ARRIVE`	–	KL0011	332	A	D27	16
20:33	20:26	20:42	21:14	`GSE_DISPATCH`	WATER_1	KL0011	332	A	D27	16
20:59	20:26	20:42	21:14	`GSE_DISPATCH`	WATER_1	KL0010	332	A	A54	16

Table 4.6 verifies the lookahead-based dispatch deferral rule. In this case, the first available vehicle could not complete the selected task within its service window, while another vehicle became available shortly afterwards from a more favourable location. The dispatcher therefore delayed assignment and selected the later but feasible vehicle. Together, these four checks verify that task selection follows the intended slack-based priority logic and that the dispatcher can defer assignment when a better feasible assignment is expected within the lookahead period.

Table 4.6: *Event log fragment for lookahead-based dispatch deferral verification*

Time	TW open	TW close	Event	Vehicle	Flight	Aircraft	A/D	Stand	Service time (min)	Candidate set
17:25	17:36	17:55	`TASK_ARRIVE`	–	KL0012	781	A	A54	19	–
17:32	17:16	17:48	`TASK_DONE`	WATER_1	KL0013	332	A	D27	16	–
17:33	17:17	17:49	`TASK_DONE`	WATER_2	KL0014	332	A	A54	16	–
17:34	17:36	17:55	`GSE_DISPATCH`	WATER_2	KL0012	781	A	A54	19	`WATER_1:6,WATER_2:1`
17:55	17:36	17:55	`TASK_DONE`	WATER_2	KL0012	781	A	A54	19	–

Vehicle Selection

Vehicle selection was verified by checking whether the closest feasible vehicle was selected from the candidate pool. Table 4.7 shows that task KL0017 was assigned to WATER_1, which was located closer to the task than WATER_2. This verifies the shortest-travel-time selection rule among feasible candidates.

Table 4.7: *Event log fragment for vehicle-selection verification based on closest feasible vehicle*

Time	A/D time	TW open	TW close	Event	Vehicle	Flight	Aircraft	A/D	Stand	Service time (min)
16:05	16:00	16:16	16:48	`TASK_ARRIVE`	–	KL0015	332	A	A54	16
16:05	16:00	16:16	16:48	`TASK_ARRIVE`	–	KL0016	332	A	D27	16
16:06	16:00	16:16	16:48	`GSE_DISPATCH`	WATER_1	KL0015	332	A	A54	16
16:06	16:00	16:16	16:48	`GSE_DISPATCH`	WATER_2	KL0016	332	A	D27	16
16:13	16:08	16:24	16:56	`TASK_ARRIVE`	–	KL0017	332	A	A54	16
16:33	16:08	16:24	16:56	`GSE_DISPATCH`	WATER_1	KL0017	332	A	A54	16

Task Execution

Task execution was verified by checking whether the event sequence followed the expected order: `TASK_ARRIVE`, `GSE_DISPATCH`, `GSE_ARRIVE`, `TASK_START`, and `TASK_DONE`. Table 4.8 shows this sequence for task KL0001. The task is released, dispatched, reached by the vehicle, started at the window opening, and completed within the service window.

Table 4.8: *Event log fragment for task-execution verification*

Time	A/D time	TW open	TW close	Event	Vehicle	Flight	Aircraft	A/D	Stand	Service time (min)
10:49	12:25	11:00	11:43	`TASK_ARRIVE`	–	KL0001	788	D	C13	28
10:50	12:25	11:00	11:43	`GSE_DISPATCH`	WATER_1	KL0001	788	D	C13	28
10:54	12:25	11:00	11:43	`GSE_ARRIVE`	WATER_1	KL0001	788	D	C13	28
11:00	12:25	11:00	11:43	`TASK_START`	WATER_1	KL0001	788	D	C13	28
11:28	12:25	11:00	11:43	`TASK_DONE`	WATER_1	KL0001	788	D	C13	28

Overall, Verification Case A confirms that the core operational logic behaves as intended. The event-log fragments verify task generation, dispatching, task selection, lookahead-based deferral, vehicle selection, and task execution before the model is applied to larger experimental tests.

4.6.3. Verification Case B — Resource and Energy Stress Case

Verification Case B was used to test resource-related and energy-related behaviour under stress conditions. The case focused on refill behaviour, charging behaviour, stop-cutoff handling, charging queueing, depot cooldown, and return-to-depot behaviour. Several parameter settings were intentionally made more restrictive than in the main experiments. This made water depletion and battery depletion clearly visible within a short simulation run.

Refill Behaviour

Refill behaviour was verified by checking whether a vehicle travelled to a fill point after its onboard water level dropped below the refill threshold. In the test case, the threshold was set at 1000 L. Table 4.9 shows that the vehicle completed task KL0023 with 875.1 L remaining, travelled to fill point F04, refilled for the configured 14 minutes, and returned to normal operation with a full 4000.0 L tank.

Table 4.9: *Event log fragment for refill-behaviour verification*

Time	Event	Vehicle	Flight	Aircraft	A/D	Stand	Service time (min)	Water task (L)	Water remaining (L)
18:46	TASK_START	WATER_1	KL0023	332	A	C13	16	–	1321.5
19:02	TASK_DONE	WATER_1	KL0023	332	A	C13	16	446.4	875.1
19:02	GSE_TO_FILL	WATER_1	–	–	–	–	–	-3124.9	875.1
19:06	GSE_ARRIVE_FILL	WATER_1	–	–	–	F04	–	–	875.1
19:06	GSE_FILL_START	WATER_1	–	–	–	F04	–	–	875.1
19:20	GSE_FILL_DONE	WATER_1	–	–	–	F04	–	-3124.9	4000.0
19:21	GSE_DISPATCH	WATER_1	KL0024	332	A	D27	16	446.4	4000.0

Charging Behaviour

Charging behaviour was verified by increasing the energy-consumption parameters so that battery depletion became visible. Table 4.10 shows that the vehicle started charging after arriving at the depot below the charging threshold, interrupted charging when a task became available, and later travelled to the charging station again at a low SOC. This verifies depot charging, charging interruption, battery depletion, and charging-station behaviour in one stress case.

Table 4.10: *Event log fragment for charging-behaviour verification*

Time	Event	Vehicle	Flight	Stand	Service time (min)	SOC (%)	Charging time (min)	Added SOC (%)	Energy (kWh)
18:43	GSE_ARRIVE_DEP	WATER_1	–	E24	–	39.8	–	–	–
18:43	GSE_CHG_REQUEST	WATER_1	–	E24	–	39.8	–	–	–
18:43	GSE_CHG_START	WATER_1	–	E24	–	39.8	–	–	–
18:50	TASK_ARRIVE	–	KL0025	A54	16	–	–	–	–
18:50	GSE_CHG_INTER	WATER_1	KL0025	A54	16	43.5	7	3.7	2.310
18:50	GSE_DISPATCH	WATER_1	KL0025	A54	16	43.5	–	–	–
18:56	GSE_ARRIVE	WATER_1	KL0025	A54	16	28.6	–	–	–
19:01	TASK_START	WATER_1	KL0025	A54	16	28.6	–	–	–
19:17	TASK_DONE	WATER_1	KL0025	A54	16	27.8	–	–	–
19:37	GSE_TO_CHG_STATION	WATER_1	–	–	–	12.6	–	–	–
19:41	GSE_ARRIVE_CHG_STATION	WATER_1	–	E24	–	4.3	–	–	–
19:41	GSE_CHG_REQUEST	WATER_1	–	E24	–	4.3	–	–	–
19:41	GSE_CHG_START	WATER_1	–	E24	–	4.3	–	–	–

Stop-cutoff Handling

Stop-cutoff handling was verified by checking whether the vehicle stopped accepting regular work after the cutoff time and completed the required end-of-day actions. Table 4.11 shows that the vehicle interrupted charging at the cutoff, travelled to the fill point, restored the tank level to 4000.0 L, returned to the depot, and resumed charging.

Table 4.11: *Event log fragment for stop-cutoff handling with fill and depot return*

Time	Event	Vehicle	Stand	SOC (%)	Water remaining (L)	Destination
22:00	GSE_STOP_CUTOFF	WATER_2	E24	45.2	3380.2	–
22:00	WT_STOP_SCHED_DONE	WATER_2	E24	45.2	3380.2	–
22:00	GSE_CHG_INTER	WATER_2	E24	45.2	3380.2	–
22:00	GSE_TO_FILL	WATER_2	E24	45.2	3380.2	F04
22:03	GSE_FILL_START	WATER_2	F04	38.8	3380.2	–
22:17	GSE_FILL_DONE	WATER_2	F04	38.8	4000.0	–
22:17	GSE_RETURN_DEP	WATER_2	F04	38.8	4000.0	E24
22:21	GSE_ARRIVE_DEP	WATER_2	E24	30.7	4000.0	E24
22:21	GSE_CHG_START	WATER_2	E24	30.7	4000.0	E24

Charging & Fill Queue

Queueing behaviour was verified for both charging and filling by restricting the relevant infrastructure capacity. Table 4.12 shows that WATER_2 waited for the charger while WATER_1 was charging, and only started charging once the charger became available. Table 4.13 shows the same logic for the single fill point: WATER_2 waited while WATER_1 was filling and started filling after WATER_1 completed the fill action. These checks verify that shared infrastructure capacity constraints create waiting times when resources are occupied.

Table 4.12: *Event log fragment for charging-queue verification*

Time	Event	Vehicle	Stand	SOC (%)	Charge wait (min)	Queue length (-)	Chargers in use (-)
18:43	`GSE_ARRIVE_DEP`	WATER_1	E24	19.8	–	–	–
18:43	`GSE_CHG_REQUEST`	WATER_1	E24	19.8	–	0	1
18:43	`GSE_CHG_START`	WATER_1	E24	19.8	0.0	0	1
18:47	`GSE_ARRIVE_DEP`	WATER_2	E24	19.1	–	–	–
18:47	`GSE_CHG_REQUEST`	WATER_2	E24	19.1	–	1	1
18:58	`GSE_CHG_INTER`	WATER_1	E24	27.7	–	–	–
18:58	`GSE_CHG_START`	WATER_2	E24	19.1	11.0	0	1

Table 4.13: *Event log fragment for filling-queue verification*

Time	Event	Vehicle	Stand	Water Remaining (L)	Fill wait (min)	Queue length (-)	Fill points in use (-)
19:32	`TASK_DONE`	WATER_1	D27	875.1	–	–	–
19:32	`GSE_TO_FILL`	WATER_1	F04	875.1	–	–	–
19:32	`TASK_DONE`	WATER_2	C13	875.1	–	–	–
19:32	`GSE_TO_FILL`	WATER_2	F04	875.1	–	–	–
19:33	`GSE_ARRIVE_FILL`	WATER_1	F04	875.1	–	–	–
19:33	`GSE_FILL_REQUEST`	WATER_1	F04	875.1	–	0	0
19:33	`GSE_ARRIVE_FILL`	WATER_2	F04	875.1	–	–	–
19:33	`GSE_FILL_REQUEST`	WATER_2	F04	875.1	–	0	1
19:33	`GSE_FILL_START`	WATER_1	F04	875.1	0.0	1	1
19:47	`GSE_FILL_DONE`	WATER_1	F04	4000.0	–	1	0
19:47	`GSE_FILL_START`	WATER_2	F04	875.1	14.0	0	1
20:01	`GSE_FILL_DONE`	WATER_2	F04	4000.0	–	0	0

Return to Depot & Redeployment

Return-to-depot and redeployment behaviour was verified by checking whether a vehicle returned to the depot after task completion when no immediate follow-up task was available. Table 4.14 shows that WATER_1 returned to depot E24 after completing KL0027 and was later redeployed from the depot to task KL0028. This verifies that depot return does not remove the vehicle from the system, but returns it to an available state for later assignment.

Table 4.14: *Event log fragment for return-to-depot and redeployment-from-depot verification*

Time	A/D time	TW open	TW close	Event	Vehicle	Flight	Aircraft	A/D	Stand	Service time (min)	SOC (%)
16:05	16:00	16:16	16:48	TASK_ARRIVE	–	KL0027	332	A	A54	16	–
16:06	16:00	16:16	16:48	GSE_DISPATCH	WATER_1	KL0027	332	A	A54	16	100.0
16:12	16:00	16:16	16:48	GSE_ARRIVE	WATER_1	KL0027	332	A	A54	16	97.0
16:16	16:00	16:16	16:48	TASK_START	WATER_1	KL0027	332	A	A54	16	97.0
16:32	16:00	16:16	16:48	TASK_DONE	WATER_1	KL0027	332	A	A54	16	96.9
16:33	–	–	–	GSE_RETURN_DEP	WATER_1	–	–	–	A54	–	96.9
16:39	–	–	–	GSE_ARRIVE_DEP	WATER_1	–	–	–	E24	–	93.9
17:10	17:05	17:21	17:53	TASK_ARRIVE	–	KL0028	332	A	A54	16	–
17:12	17:05	17:21	17:53	GSE_DISPATCH	WATER_1	KL0028	332	A	A54	16	93.9
17:18	17:05	17:21	17:53	GSE_ARRIVE	WATER_1	KL0028	332	A	A54	16	90.9
17:21	17:05	17:21	17:53	TASK_START	WATER_1	KL0028	332	A	A54	16	90.9
17:37	17:05	17:21	17:53	TASK_DONE	WATER_1	KL0028	332	A	A54	16	90.8

Overall, Verification Case B confirms that the resource and energy mechanisms behave as intended under stress conditions. The event-log fragments verify refill triggering, refill duration, battery depletion, charging triggers, charging interruption, infrastructure queueing, stop-cutoff handling, depot return, and redeployment.

4.6.4. Verification Case C – Uncertainty Behaviour

Verification Case C was used to verify the stochastic uncertainty mechanisms in the simulation model. The purpose of this case is not to validate whether the uncertainty profiles are representative of the case study operation, but to verify whether the implemented stochastic mechanisms behave according to the intended model logic. The case focuses on task-timing uncertainty, travel-time uncertainty, seed reproducibility, and the combined propagation of both uncertainty mechanisms.

In this verification case, task-timing uncertainty was based on empirical resampling rather than parametric delay sampling. A synthetic schedule with 100 water-service tasks and four water vehicles was used. The same synthetic schedule was applied in all runs, so that differences in the results could be attributed to the uncertainty settings and random seed rather than to differences in demand input.

For each task, the model first sampled one mutually exclusive timing outcome from the configured daypart-specific probabilities: on time, late, or early. When a late or early outcome was selected, the corresponding magnitude was sampled from the precomputed capped empirical magnitude array for the relevant daypart and timing category. These empirical arrays were derived from the arrival-delay analysis described in Subsection 4.5.2. Magnitudes were capped at the 99th percentile after applying the 1440-minute sanity filter. The applied timing shift was calculated as $\lceil \text{magnitude} \rceil$ minutes, with positive values for late arrivals and negative values for early arrivals.

Travel-time uncertainty was implemented as a separate Bernoulli mechanism. For each travel leg, the model sampled whether a travel disturbance occurred. If a disturbance was triggered, the deterministic travel time was increased by the configured multiplicative factor. This mechanism was tested independently from task-timing uncertainty and in combination with it.

The verification consisted of three checks. First, a deterministic baseline was used to confirm that the model behaved correctly when both uncertainty mechanisms were disabled. Second, task-timing uncertainty and travel-time uncertainty were enabled separately to verify that both mechanisms operated independently and affected the expected part of the simulation. Third, combined uncertainty runs were used to verify that both mechanisms could operate together, that runs with the same seed were reproducible, and that different seeds produced different but plausible outcomes. The tested runs are summarised in Table 4.15.

Table 4.15: *Overview of Verification Case C runs*

Run type	Seed	Task timing uncertainty	Travel time uncertainty
Deterministic baseline	–	Disabled	Disabled
Task-only	123	Enabled	Disabled
Task-only	124	Enabled	Disabled
Travel-only	123	Disabled	Enabled
Travel-only	124	Disabled	Enabled
Combined	123	Enabled	Enabled
Combined repeated	123	Enabled	Enabled
Combined	124	Enabled	Enabled

The deterministic baseline provides the reference case. As shown in Table 4.16, all 100 tasks were generated and completed without delay events, early-arrival events, or service-window violations.

Table 4.16: *Deterministic baseline results for Case C*

Metric	Baseline value
Generated tasks	100
`TASK_ARRIVE_DELAY` events	0
`TASK_ARRIVE_EARLY` events	0
Completed tasks	100
Tasks started within window (%)	100
Tasks finished within window (%)	100
Average time outside window (min)	0.0

Table 4.17 and Table 4.18 verify the task-timing uncertainty mechanism. The task-only runs generated both delayed and early-arrival events, and different seeds produced different but plausible timing patterns. The bucket-specific results show that morning, afternoon, and evening tasks followed different late and early-arrival profiles, confirming that the daypart-specific empirical profiles were applied rather than one global uncertainty process.

Table 4.17: *Summary of task timing uncertainty in task-only runs*

Run	Seed	Early tasks	Delayed tasks	Total delay (min)	Max delay (min)
Task-only	123	36	60	1596	121
Task-only	124	32	64	1609	138

Table 4.18: *Bucket-specific task timing uncertainty results*

Seed	Daypart	Late share (%)	Mean late delay (min)	Early share (%)	Mean early magnitude (min)
123	Morning	50.0	14.53	50.0	5.00
123	Afternoon	61.9	21.69	31.0	8.08
123	Evening	67.9	42.84	28.6	9.75
124	Morning	53.3	18.06	36.7	9.64
124	Afternoon	67.4	22.55	32.6	7.57
124	Evening	70.4	35.05	25.9	11.43

Table 4.19 verifies the travel-time uncertainty mechanism. With task timing uncertainty disabled, selected travel-start legs were affected by the multiplicative travel-time disturbance. Affected 6-

minute trips increased to 7 minutes, which is consistent with the configured rule $\lceil 6 \cdot 1.15 \rceil = 7$. The difference between `GSE_DISPATCH` and `GSE_ARRIVE` matched the logged travel time.

Table 4.19: *Observed travel-time uncertainty effects in travel-only runs*

Run	Seed	Travel-start legs	Affected travel-start legs	Affected task-dispatch legs	Observed change
Travel-only	123	149	16	9	6 min → 7 min
Travel-only	124	149	14	9	6 min → 7 min

Same-seed reproducibility was verified by repeating the combined uncertainty run with seed 123. The repeated run produced identical logs and performance outcomes. The combined uncertainty runs were then used to verify that task-timing uncertainty and travel-time uncertainty can operate together. Table 4.20 shows that combined uncertainty changes operational performance compared with the deterministic baseline, while preserving the task-shift pattern for the same seed.

Table 4.20: *Operational propagation of uncertainty*

Run	Seed	Delayed tasks	Early tasks	Tasks finished within window (%)	Avg. finish outside window (min)
Baseline	–	0	0	100	0.00
Task-only	118	63	33	54	22.48
Combined	118	63	33	52	23.40
Task-only	119	60	37	61	21.51
Combined	119	60	37	60	21.56

Overall, Verification Case C confirms that the stochastic uncertainty mechanisms behave as intended. Task-timing uncertainty generated delayed and early task-arrival events from the capped empirical magnitude pools, travel-time uncertainty increased selected travel durations according to the configured multiplicative rule, same-seed runs were reproducible, and combined uncertainty propagated into operational performance. The results of Verification Case C are summarised in Table 4.21.

Table 4.21: *Summary of Verification Case C results*

Verification check	Observed evidence	Result
Deterministic baseline	100 tasks completed; no delay or early-arrival events; 100% within window; 0.0 min outside window.	✓
Empirical task timing	`TASK_ARRIVE_DELAY` and `TASK_ARRIVE_EARLY` events were generated; logged shifts matched observed arrival-time changes.	✓
Empirical magnitude sampling	Late and early magnitudes matched the capped daypart-specific empirical arrays after ceiling.	✓
Seed variation	Different seeds produced different but plausible late and early-arrival patterns.	✓
Daypart-specific profiles	Morning, afternoon, and evening buckets showed different late and early-arrival behaviour.	✓
Travel time uncertainty	Selected travel-start legs were affected; affected 6-minute trips increased to 7 minutes.	✓
Travel-time consistency	`GSE_ARRIVE` minus `GSE_DISPATCH` matched the logged `travel_min`.	✓
Same-seed reproducibility	The repeated combined run with seed 123 produced identical logs and outcomes.	✓
Combined uncertainty	Task timing and travel time uncertainty operated together without inconsistent behaviour.	✓
Operational propagation	Uncertainty reduced service-window performance compared with the deterministic baseline.	✓

5

Case Study

This chapter instantiates the generic simulation framework developed in Chapter 4 for the KLM Schiphol case and validates whether the resulting model behaviour is sufficiently plausible for simulation experiments and fleet-capacity assessment. The model structure remains generic, but the input data, operational rules, fleet configurations, infrastructure locations, and validation evidence are case-specific. Consequently, the numerical fleet-size outcomes in Chapter 6 should be interpreted as results for the tested KLM Schiphol assumptions and vehicle groups, not as generic airport-wide fleet requirements.

The validated case-study application is limited to three vehicle groups: water vehicles, toilet vehicles, and loaders. These groups were selected because they represent two relevant operational behaviours within the same modelling structure. Water and toilet vehicles are depot-based, resource-constrained service vehicles with onboard capacity limits and possible intermediate visits to refill, dump, or depot locations. Loaders represent stand-side platform equipment without onboard service-resource constraints or replenishment requirements. This scope allows the case study to test both resource-constrained depot-based operations and stand-side aircraft-service operations, while keeping the validation scope manageable.

The broader potential model scope includes thirteen GSE vehicle types: water vehicles, toilet vehicles, tank-service vehicles, loaders, transporters, loader-transporters, three types of conveyor belts, three-wheel tractors, aircraft stairs, dispensers, and pushback tractors. These vehicle types share a common operational structure: a vehicle is assigned to a time-window-constrained task, travels to the aircraft stand or service location, performs the service, and then either becomes available for a next task or moves to an operational support location. However, not all thirteen vehicle types are implemented, experimentally analysed, or validated in this thesis. Table 5.1 therefore distinguishes the broader extension scope of the framework from the narrower validated case-study scope.

Table 5.1: *Overview of vehicle types considered for broader model extension*

Vehicle type	Compatible aircraft	Capacity	Intermediate stop	Base
Conveyor belt 9M	320, 332, 333, 789, 781, 772, 77W	No	No	Near VOP
Powerstow SLH	E75, E90, 295	No	No	Near VOP
Powerstow LLH	73W, 73H, 73J	No	No	Near VOP
Loader	332, 333, 789, 781, 772, 77W	No	No	Near VOP
Transporter	332, 333, 789, 781, 772, 77W	No	No	Near VOP
Loader-transporter	32Q	No	No	Near VOP
Three-wheel tractor	NABO, WIBO	No	No	Near VOP
Aircraft stairs	COM, NABO, WIBO at buffer positions	No	No	Near VOP
Dispenser	COM, NABO, WIBO if a hydrant is available	No	No	Depot
Pushback tractor	COM, NABO, WIBO	No	No	Depot
Water truck	COM, NABO, WIBO	Yes	Yes	Depot
Toilet truck	COM, NABO, WIBO	Yes	Yes	Depot
Tank truck	COM, NABO, WIBO if no hydrant is available	Yes	Yes	Depot

Note The table defines the broader potential extension scope of the simulation framework. Only water vehicles, toilet vehicles, and loaders are implemented, experimentally analysed, and validated in this thesis.

The vehicle types differ mainly in aircraft compatibility, operational department, onboard capacity constraints, intermediate service-point requirements, and starting location. Platform vehicles are generally positioned near the Vliegtuigopstelplaats (aircraft stand) (VOP), while water, toilet, tank, dispenser, and pushback-related vehicles are more strongly linked to depot-based operations. Water, toilet, and tank vehicles also require onboard capacity tracking and may need intermediate replenishment or disposal visits, whereas most platform vehicles do not.

Extension of the model to additional vehicle types therefore requires vehicle-specific input parameters and validation evidence. This includes service rules, aircraft compatibility, movement patterns, infrastructure locations, energy and charging parameters, resource constraints where applicable, operational-control rules, and validation data. The case-study results in this thesis should therefore be interpreted as validated for water vehicles, toilet vehicles, and loaders only, while the broader framework indicates how additional GSE types could be incorporated.

5.1. Amsterdam Airport Schiphol

AAS is one of Europe's major hub airports and functions as a key operational node within the European air transport network. The airport handles a high volume of traffic and operates with pronounced peak structures throughout the day, with the largest concentration of movements occurring in the morning, followed by additional peaks in the afternoon and evening. Between these peak periods, activity levels decrease, resulting in alternating periods of high and moderate operational pressure [65, 70].

Unlike airports dominated by a single carrier, Schiphol accommodates a wide range of airlines and stakeholders, contributing to a complex multi-actor operational environment. This diversity increases coordination requirements during aircraft turnaround processes, as multiple parties, such as airlines, ground handlers, and airport operations, must interact within shared time and resource constraints. Schiphol is particularly relevant as a case study due to a combination of operational and strategic characteristics. First, the airport operates at a high traffic density. In 2024, Schiphol handled approximately 473,000 aircraft movements and 66.8 million passengers, placing it among the busiest airports in Europe. This high throughput results in a system where timing, coordination, and resource allocation are critical to maintaining operational performance [70, 66].

Second, Schiphol operates under significant capacity constraints. With a policy-driven cap on annual aircraft movements close to current traffic levels, the system functions near its operational limits. This leads to strong competition for critical resources such as stands, gates, and ground handling capacity. Empirical observations indicate that operational disruptions, such as last-minute gate changes, can have substantial downstream effects on departure performance, further highlighting the importance of efficient ground operations [70, 68].

Third, the airport exhibits considerable spatial and operational complexity. Schiphol consists of multiple piers and a large number of aircraft stands distributed across the apron. The physical scale of the airport, including long taxi distances and multiple traffic flows, increases the complexity of vehicle routing and dispatching. This makes the airport particularly suitable for studying logistics processes involving mobile ground support equipment [71, 72].

Finally, Schiphol distinguishes itself through its ambitious sustainability agenda, particularly in the electrification of ground operations. The airport aims to achieve emission-free ground operations by 2030, supported by large-scale investments in charging infrastructure and energy capacity. This transition towards electric ground support equipment introduces additional constraints and decision variables related to charging behaviour, energy availability, and infrastructure usage. As such, Schiphol provides a highly relevant context for analysing the operational implications of electric fleet deployment [69].

Taken together, the combination of high traffic density, capacity pressure, operational complexity, and an advanced electrification agenda makes Schiphol a representative and challenging case for studying ground support operations and fleet sizing in a realistic airport environment.

5.2. Research Background KLM Ground Services (KLM GS)

KLM Royal Dutch Airlines (KLM), founded in 1919, is the oldest airline in the world still operating under its original name. Since 2004, it has been part of the Air France–KLM Group (AFKL). KLM operates an extensive global network, with AAS as its main hub. As a full-service international airline, it plays a key role in global air transport by connecting the Netherlands with destinations worldwide, serving 92 European cities and 70 intercontinental destinations. Each year, KLM transports approximately 34.1 million passengers and 621.000 tons of cargo [7].

KLM places strong emphasis on safety, operational reliability, and sustainability. The airline is investing in modernising its aircraft fleet, improving operational efficiency, and reducing its environmental impact. KLM also supports nature restoration through projects in countries such as Colombia, Panama, and Uganda. In addition, the airline uses Sustainable Aviation Fuel (SAF)/Alternative Aviation Fuel (AAF), which can reduce CO_2 emissions by up to 65% compared with conventional kerosene [8]. On the ground, AAS and so KLM aim to achieve emission-free GH by 2030, a target that requires major changes in both processes and GSE [67].

The organisational structure of KLM can be summarised as shown in Figure 5.1. The structure consists of four layers, each containing different divisions. The following section focuses on the division relevant to this project.

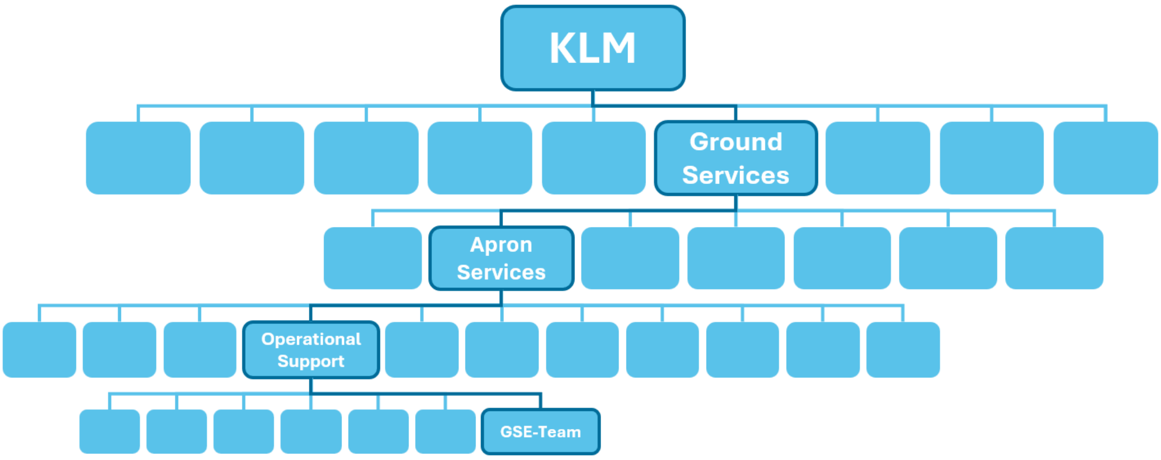


Figure 5.1: *Simplified Organisational Structure KLM [7].*

Within the organisation, KLM GS is responsible for all GH activities that take place between an aircraft's arrival and departure. The operational work is carried out by Passenger Services (PS), Baggage Services (BS), and Apron Services (AS). AS manages all below-wing handling processes at the VOP. These activities include loading and unloading baggage and cargo; refilling potable water and fuel; conducting pushback and towing operations; providing lavatory service; supporting catering and cleaning; performing de-icing operations; placing boarding stairs; and delivering Ground Power Unit (GPU), Air Starter Unit (ASU), and Pre-Conditioned Air Unit (PCA). Depending on the aircraft type and the required services, a single turnaround may require the use of up to 12 different GSE units [41].

Within AS, the Operational Support (OS) department plays a central role in ensuring that the GSE fleet is available, and reliable. Since OS does not operate the vehicles itself, some influence on efficiency can be achieved through vehicle specifications and, where relevant, the placement of charging points, but this is limited. The department is responsible for fleet management, data analysis, monitoring performance and reliability, and ensuring that sufficient equipment is available for daily operations. In general, OS functions as a support department for the operational divisions of AS (such as K1, K4, K5, ARD KE, ATD KN, and AHSU KF), which depend on this equipment to carry out their tasks. OS/Ground Support Equipment Team (GSE-Team) manages the vehicles, while K1, K4, K5, etc. are the users. This coordination includes a central coordinator who directs the operators within these departments. Operators use the vehicles and can therefore be considered to indirectly control them. However, depending on the department, operators have some autonomy to select any available vehicle.

Several teams operate within OS. Focusing on the GSE-Team, this team is responsible for ensuring the availability of all GSE required for aircraft handling. In addition, the team is involved in evaluating new equipment and potential modifications, for example in relation to the ongoing electrification of the fleet. GSE refers to all vehicles and equipment used to support aircraft during GH. This includes belt and lower deck loaders (no higher deck loaders), potable-water trucks, fuel trucks, baggage trucks, cargo transporters, lavatory trucks, disabled-passenger lift trucks (owned by AAS), towing tractors, passenger stairs, and various service units like GPU, ASU, and PCA as mentioned earlier in this page.

5.3. KLM Operations

KLM functions as the dominant hub carrier at AAS. As one of the two hub airports of the Air France–KLM group, Schiphol plays a central role in the airline's network, with the group accounting for approximately half of all aircraft movements at the airport. This dominant position implies that KLM's operational structure significantly shapes overall traffic patterns and ground operations at Schiphol [73].

KLM operates Schiphol as a transfer hub within a hub-and-spoke network. Feeder traffic from European destinations is routed into the hub and connected to outbound flights, including intercontinental services. Public Air France–KLM documentation describes this hub structure as organised around synchronized transfer waves, in which inbound arrival flows are aligned with outbound departure banks to maximise connectivity. As a result, aircraft movements are not evenly distributed over time, but instead exhibit clustered patterns associated with these transfer waves [42, 3].

The strong reliance on transfer passengers leads to a highly time-critical operation. Minimum connection times of approximately 40 minutes for Schengen flights and 50 minutes for non-Schengen flights indicate that transfer processes are tightly constrained. Schengen flights refer to flights within the Schengen area, where passengers typically do not undergo regular passport control at internal borders. Although exact turnaround durations are not publicly benchmarked against other airports, the hub structure implies a need for short and, more importantly, predictable turnaround times. Operational performance therefore depends on maintaining reliable ground processes, particularly during peak transfer periods [43, 73].

Aircraft turnaround at Schiphol can be characterised as a tightly managed, multi-process operation. Rather than a single sequential activity, turnaround consists of numerous interrelated service processes, including unloading, loading, cleaning, refuelling, and technical servicing. Industry standards and operational documentation indicate that many of these activities are executed in parallel to minimise ground time. At the same time, strong interdependencies exist between processes, meaning that delays in one activity can propagate through the entire turnaround chain. This interdependence increases the sensitivity of operations to disruptions [66, 37].

The temporal structure of KLM's operations further amplifies operational pressure. Because traffic is organised around transfer waves, demand for ground handling services is concentrated in relatively short time intervals. As the dominant hub carrier, KLM both contributes to and is affected by these peaks, resulting in a traffic pattern that is strongly shaped by these peak periods [3, 73].

Overall, the operational pressure within KLM's Schiphol operations is concentrated in three dimensions. First, strict time windows impose tight constraints on service execution, particularly in relation to transfer connectivity. Second, limited and shared resources such as stands, ground support equipment, and handling capacity must be allocated efficiently under high demand. Third, the large number of interdependent processes requires intensive coordination between multiple actors. Together, these characteristics create a complex and time-critical operational environment in which efficient planning and execution of ground support services are essential [43, 66, 73].

5.4. Model Input

This section describes how the generic simulation framework is instantiated for the AAS case study. The case-study input consists of operational demand, aircraft characteristics, service parameters, spatial representation, fleet configuration, infrastructure settings, and operational-control parameters. Together, these inputs define the demand pattern and the operational constraints under which the simulated eGSE vehicles operate.

5.4.1. Operational Demand

The primary source of operational demand is a flight-schedule dataset from a summer peak period in June 2026. From this dataset, individual days are selected as operational instances for the simulation experiments. The selected week is used as a high-demand reference period for robust fleet-capacity assessment. This choice is intentional: fleet-capacity decisions for operationally critical eGSE should be robust to peak demand conditions, rather than optimised only for average workload.

Figure 5.2 shows the number of arrivals and departures during the week from 08-06-2026 to 14-06-2026. The figure shows that flight demand is distributed relatively evenly across the week. For the baseline configuration in this study, 14-06-2026 is selected as the reference day. This day contains 811 flight records, consisting of 407 arrivals and 404 departures. After preprocessing and filtering, not every flight record generates a service task for every vehicle group. The number of resulting tasks depends on the service type and the operational rules applied during preprocessing. The final service dataset therefore contains arrival-related tasks, departure-related tasks, and turnaround-related tasks. A more detailed analysis of the selected baseline day and the resulting service tasks per vehicle group is provided in Subsection 6.3.2.

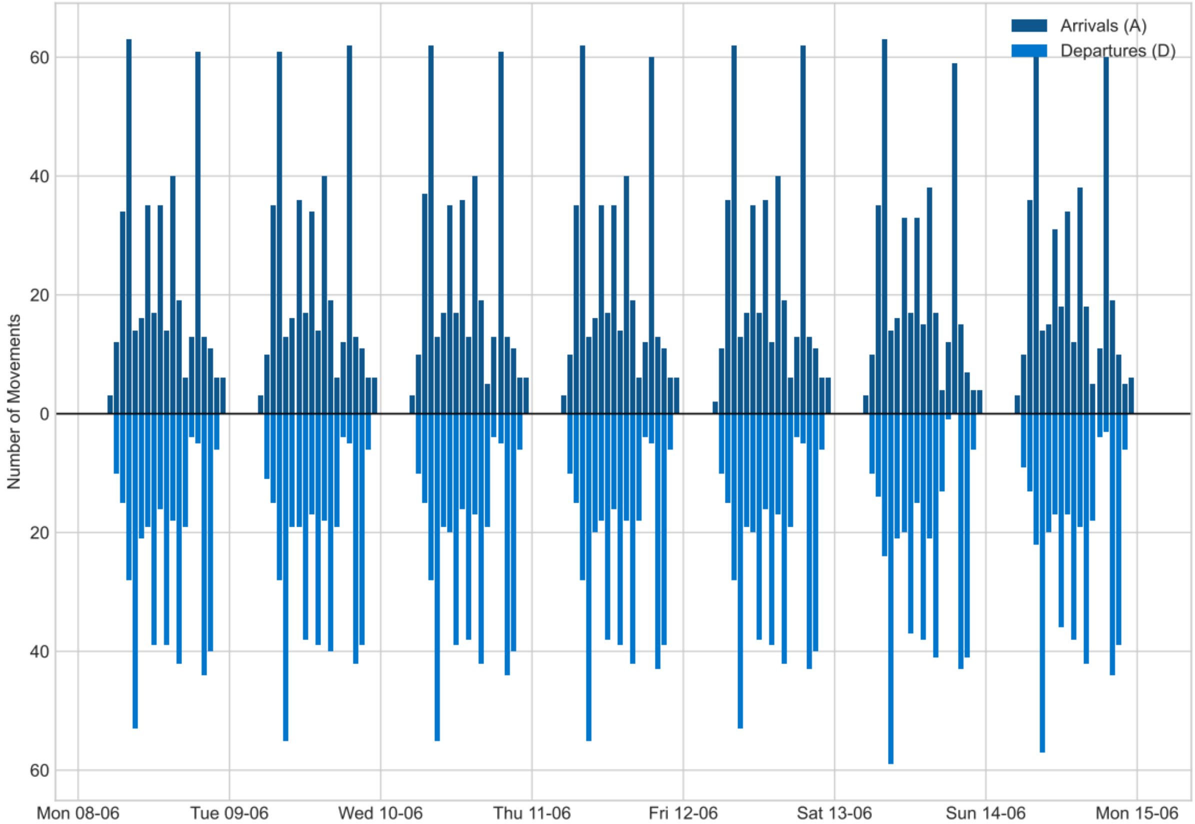


Figure 5.2: *Flights from 08-06-2026 until 14-06-2026*

The temporal structure of demand on the selected baseline day covers approximately 18 hours of operation, from the early morning until the late evening. To capture differences in operational intensity during the day, demand is divided into three time-of-day buckets: morning, afternoon, and evening. The morning period is defined as before 11:00, the afternoon as 11:00–17:00, and the evening as after 17:00. These buckets are also used in the stochastic uncertainty modelling, because they define the empirical delay profiles used for task-timing uncertainty.

Aircraft characteristics are represented at the level of aircraft subtype and body category. The dataset includes a mix of regional aircraft, narrow-body aircraft, and wide-body aircraft, as shown in Figure 5.3. Aircraft subtype is important because it determines service durations, aircraft compatibility, and service-resource requirements. Body category is used where a more aggregated representation is sufficient, for example when approximating resource quantities for specific service types.

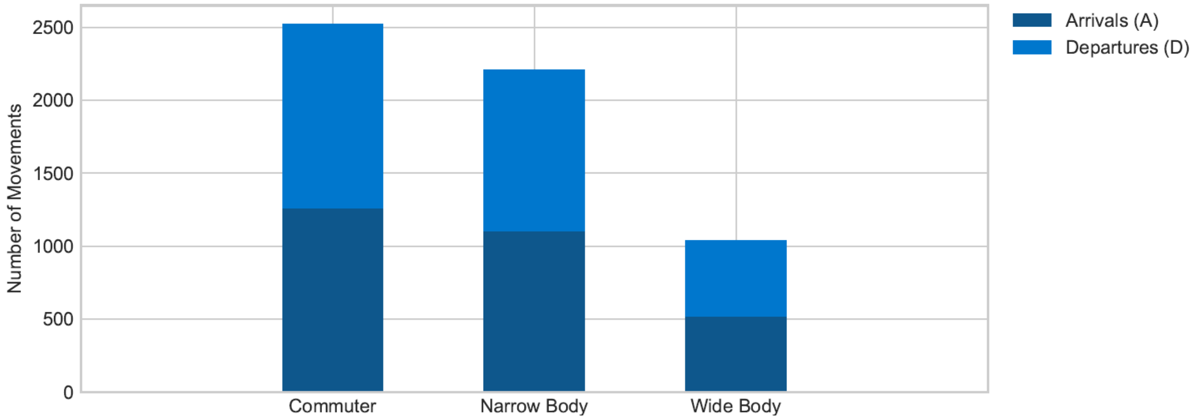


Figure 5.3: *Body types in week 08-06-2026 until 14-06-2026*

Service parameters are derived from predefined service rules. Service durations and time-window offsets are specified for each combination of aircraft subtype and operational classification, namely arrival, departure, or turnaround. The time windows are defined relative to the planned service moment using early and late offsets. Their length remains fixed throughout the simulation, unless task-timing uncertainty shifts the full service window in time. Resource requirements are also linked to aircraft characteristics. Depending on the vehicle group, these requirements may represent potable-water demand, lavatory-waste quantities, load-handling demand, or other service-specific quantities.

5.4.2. Case-Specific Fleet and Infrastructure Specifications

This subsection defines the case-specific fleet and infrastructure configuration used to instantiate the simulation framework for the KLM Schiphol case. The configuration translates the generic model structure into operational input values for the three analysed vehicle groups: water vehicles, toilet vehicles, and loaders. These inputs define vehicle starting locations, charging locations, supporting service locations, energy parameters, onboard resource constraints, and infrastructure capacities. Together, they determine how vehicles become available, unavailable, or constrained during the simulated operating day.

The configuration reflects the different operational roles of the vehicle groups. Water and toilet vehicles are modelled as depot-based service vehicles with onboard resource constraints. Their availability is therefore affected not only by task timing and travel time, but also by intermediate refill or dump requirements. Water vehicles may need to visit a refill point when their remaining potable-water level falls below the configured threshold. Toilet vehicles may need to visit a dump location when the collected waste volume approaches the onboard tank capacity. For instance, for toilet vehicles, the collected waste volume is differentiated by aircraft type. Smaller aircraft, such as the E70 and 73H, generate approximately 70 litres of waste. Medium-sized aircraft, such as the 332 and 333, generate approximately 250 litres. Larger wide-body aircraft, such as the 772 and 789, can generate up to 400 litres. This differentiation affects how quickly the onboard tank fills up and therefore how often a toilet vehicle must leave the task-assignment pool to visit the dump location.

Loaders are modelled differently. They are distributed across the airport and do not require intermediate resource replenishment or disposal. Their availability is mainly determined by stand-side positioning, task execution, battery use, charging access, and the timing of linked arrival and departure operations. Loader operations also include aircraft-type-specific turnaround slack between linked arrival-related and departure-related tasks. When an arrival-related task is delayed, this delay is therefore not transferred one-to-one to the associated departure-related task. Instead, a buffer is applied before the delay affects the departure-side loader task. For example, the applied buffer is 40 minutes for the 332 and 22 minutes for the 77W. This reflects that loader planning contains operational slack before an arrival delay directly affects the departure-side handling process.

Table 5.2 summarises the main fleet, infrastructure, energy, and resource parameters used in the case-study configuration. These values define the reference configuration for the baseline simulation. Fleet sizes are varied separately in the experiment design to determine deterministic and stochastic fleet-size requirements.

Table 5.2: *Case-study fleet, infrastructure, and energy configuration per vehicle group*

Parameter	Water vehicles	Toilet vehicles	Loaders
Baseline fleet size (-)	11	11	25
Depot	E24	E20	19 hubs distributed across the airport
Charging locations	E24: 11	E20: 5; E22: 4	28 chargers distributed over 19 locations
Refill/dump location and duration	Refill at F04; 14 min	Dump at A34; 15 min	–
Tank capacity (L)	4,000 L water	3,300 L waste	–
Tank threshold	Refill when ≤ 1,000 L	Dump when ≥ 80% full	–
Driving speed	300 m/min (≈ 18 km/h)	300 m/min (≈ 18 km/h)	183.3 m/min (≈ 11 km/h)
Energy per task (kWh)	0.1 kWh	0.1 kWh	3.2 kWh
Energy per km (kWh/km)	1.1 kWh/km	1.1 kWh/km	1.6 kWh/km
Battery capacity (kWh)	70 kWh	70 kWh	80 kWh
Usable battery share (%)	90% = 63 kWh	90% = 63 kWh	90% = 72 kWh
Charger power (kW)	22 kW	22 kW	16 kW
Minimum SOC before task (%)	20%	20%	20%
Depot charging threshold (%)	≤ 70% SOC	≤ 70% SOC	≤ 70% SOC
Minimum depot stay after return (min)	15 min	15 min	n.a.
End-of-shift slots	21:00–03:00	21:00–03:00	–

Note The table reports the case-study configuration used to instantiate the simulation model. The baseline fleet sizes define the reference configuration; alternative fleet sizes are varied separately in the experiment design.

5.4.3. System Resources and Infrastructure

Movement across the airport is modelled using an aggregated spatial representation. Instead of modelling every individual stand, road segment, and possible vehicle route, aircraft stands are assigned to location groups. Travel distances between these location groups are defined in a precomputed distance matrix. This abstraction reduces the spatial detail required in the simulation while preserving the main differences in travel distance across the apron.

The distance matrix distinguishes between outbound and return movements. This is necessary because parts of the airport network may be affected by one-way traffic rules or route restrictions. As a result, the distance from location group i to location group j does not necessarily equal the distance from j to i. The spatial layout of Schiphol used to define the case-study context is shown in Figure 5.4. The network points used to represent Schiphol in the simulation model are shown in Figure 5.5. The distance matrix also includes an approximation for movements within the same location group. Although stands within one group are spatially aggregated, vehicles may still need to move between different VOP positions inside that group. These movements are therefore not treated as zero-distance trips. Instead, internal VOP-to-VOP distances are examined for each group. The maximum internal distance is used as an upper reference point, while the final within-group distance is selected as a representative approximation based on the spread of internal distances. This approach avoids underestimating short local movements without assuming that every within-group trip equals the longest possible internal movement.

Travel times are calculated from the distance matrix and vehicle-specific driving speeds. The speed parameter represents an average operational driving speed rather than a maximum technical speed. This allows the model to approximate realistic apron driving behaviour while keeping the spatial representation tractable. For a movement from location group i to location group j, the travel time is calculated as a function of distance and vehicle speed, with a minimum travel time enforced for all non-zero movements. This prevents unrealistically short travel times between nearby location groups.

Infrastructure is modelled as a set of capacity-constrained service locations. These include charging locations, refill points, dump points, depots, and other vehicle-specific support locations where relevant. Each infrastructure location has a predefined capacity, meaning that only a limited number of vehicles can use it at the same time. When the capacity of a required resource is occupied, vehicles wait in a queue before the corresponding activity can start.

This structure allows the model to capture infrastructure-related vehicle unavailability without representing the full physical layout of each service area in microscopic detail. Queueing at chargers, refill points, or dump points can delay vehicle availability and can therefore affect task assignment, service-window performance, and fleet-size requirements.

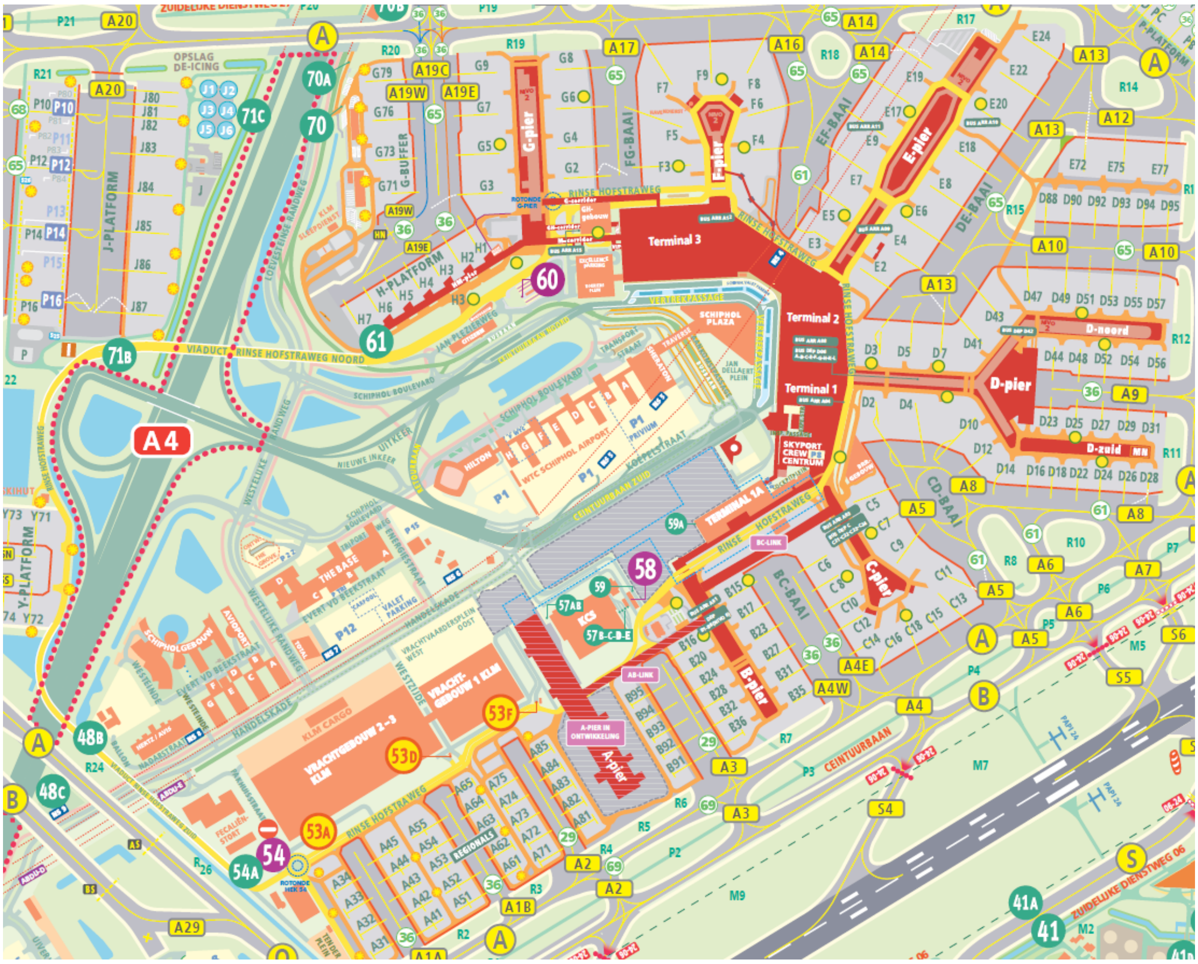


Figure 5.4: *Map Schiphol*

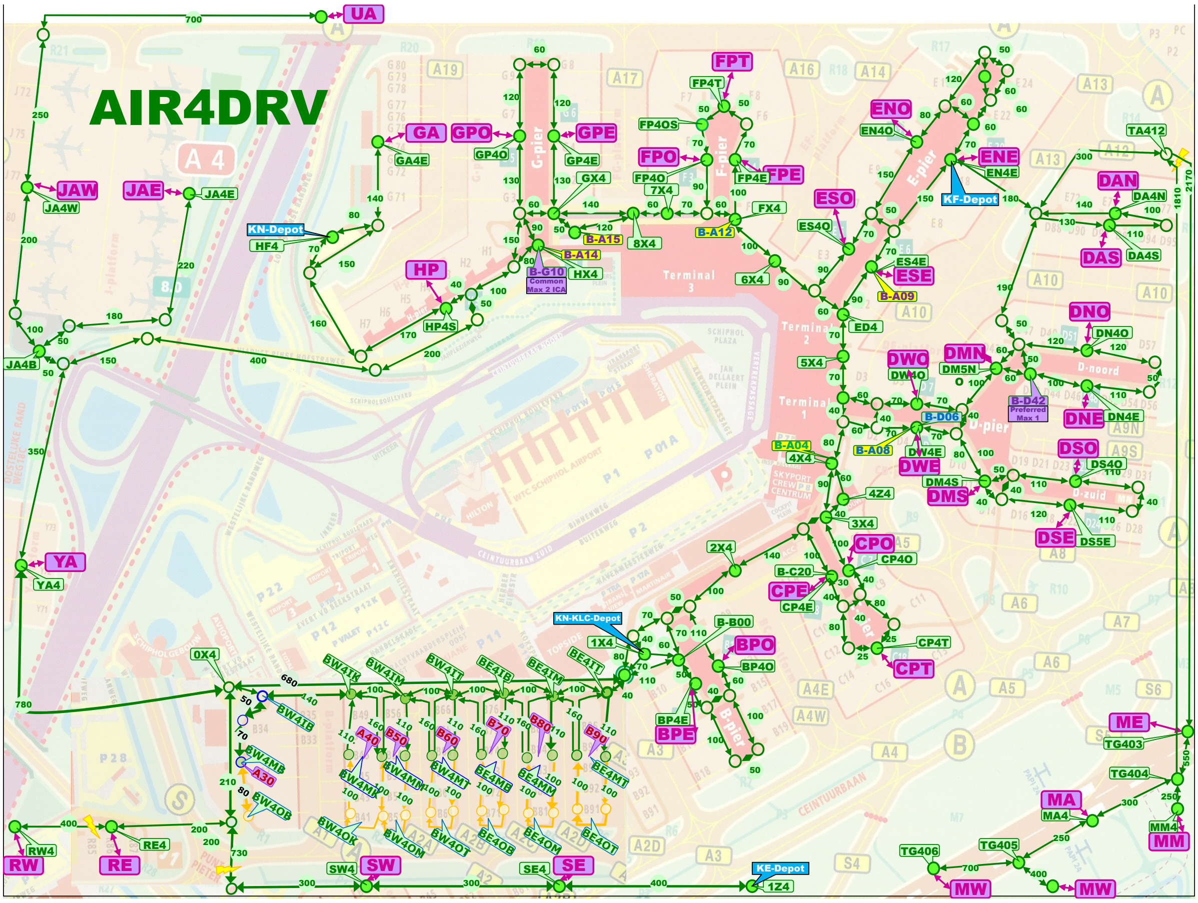


Figure 5.5: *Netpoints Schiphol*

5.4.4. Operational Policies and Control Parameters

Vehicle behaviour in the case-study model is governed by a set of operational policies and control parameters. These parameters define how tasks are released, how vehicles are dispatched, when vehicles must charge or visit service infrastructure, and when vehicles stop accepting new assignments. Together, they translate the generic simulation logic into the operational setting used for the KLM Schiphol case.

Tasks are released into the system before their planned service-window opening. This early-release offset allows vehicles to be dispatched before the service window starts, so that travel time can be accounted for without allowing service to begin too early. Service execution remains constrained by the task-specific service window.

Charging, refilling, and waste dumping are controlled through threshold-based rules. Vehicles can be sent to charge when their SOC falls below the configured operational threshold. Vehicles located at their depot or charging location can also charge when idle and below the configured depot-charging threshold. Water vehicles must visit a refill point when their onboard potable-water level falls below the predefined threshold. Toilet vehicles must visit a dump location when the collected waste volume exceeds the configured share of tank capacity. These rules create additional operational movements and can make vehicle availability dependent on infrastructure access.

Task allocation is governed by the rule-based dispatcher described in Chapter 4. Tasks are prioritised based on urgency, using service-window timing, service duration, and minimum travel-time considerations. Vehicle assignment is updated dynamically as vehicles become available, tasks are released, and operational constraints change. The dispatcher also includes limited lookahead and reassignment-related control parameters, which affect whether a task is assigned immediately or deferred when a more suitable vehicle is expected to become available shortly.

The model further includes stop-cutoff logic to represent the operational end of the working period for a vehicle group. After the cutoff becomes active, vehicles may stop accepting new regular assignments and may instead complete required end-of-shift actions, such as returning to depot, refilling, dumping, or charging. These rules prevent vehicles from being assigned beyond the modelled operating period and ensure that end-state requirements are handled consistently.

Uncertainty settings are configured separately for task-timing uncertainty and travel-time uncertainty. Both mechanisms can be enabled or disabled independently, allowing deterministic baseline runs, task-timing uncertainty runs, travel-time uncertainty runs, and combined uncertainty runs. Task-timing uncertainty is applied during preprocessing and changes the temporal demand pattern before simulation execution. Travel-time uncertainty is applied dynamically during simulation execution and affects individual vehicle movements. Fixed random seeds are used to ensure reproducibility across stochastic replications.

In addition to stochastic uncertainty, deterministic demand-generation rules can be applied during preprocessing. For selected aircraft subtypes, a deterministic demand-reduction rule removes every second water or toilet task to approximate observed servicing frequency. This rule is not treated as stochastic cancellation, but as a reproducible preprocessing assumption for aircraft types that do not require servicing after every turnaround.

Together, these operational policies and control parameters define the case-specific operating logic used in the Schiphol simulation experiments. They determine not only how tasks are assigned, but also how charging, onboard resources, infrastructure access, shift constraints, and uncertainty affect vehicle availability during the simulated day.

5.5. Performance Metrics

This section defines the performance metrics used to evaluate the simulation experiments. The metrics assess service reliability, fleet utilisation, vehicle movement, and charging-infrastructure use. Together, they provide the basis for comparing model configurations and for determining the fleet sizes required under the selected service-level interpretations.

The primary service-performance metric is the on-time task completion rate. This metric measures the percentage of completed tasks that are finished within their assigned service window. It is the main indicator used to assess whether a fleet can meet the operational service requirements. In this thesis, on-time completion refers to completion within internal service windows. It should therefore not be interpreted directly as aircraft departure-delay performance.

A second service metric is the unserved task rate. This metric measures the percentage of tasks that remain incomplete under the implemented stopping rules. Its interpretation differs between vehicle groups. For water and toilet vehicles, a task can become unserved when it can no longer be assigned after the relevant stop-cutoff times, which represent the end of the operator service period. For loaders, no equivalent service-specific cutoff is applied; loader tasks are therefore classified as unserved only if they remain incomplete after the 24-hour simulation horizon. The unserved task rate should therefore be interpreted as unresolved operational demand under the implemented stopping rules, not as a measure of service-window lateness.

To capture the severity of service-window violations, the average lateness of late tasks is included. This metric measures by how many minutes late tasks exceed their service-window closing time. It provides additional information beyond the binary distinction between on-time and late completion.

Fleet utilisation and vehicle-activity metrics are used to assess how vehicle capacity is used during the simulated day. Overall vehicle utilisation expresses the share of available vehicle time spent in active operational states. In addition, time-share metrics distinguish between service time, travel time, charging time, depot idle time, and waiting time. These indicators help identify whether performance is limited by service workload, travel requirements, charging, infrastructure waiting, or idle capacity.

Distance-related metrics are included to evaluate vehicle movement. Total travel distance measures the total number of kilometres driven by the fleet during the simulation period. Dispatch-to-task travel distance measures the distance travelled specifically between dispatch and task arrival. These metrics indicate how much movement is required to deliver the simulated service level and help compare alternative operational-control policies.

Charging performance is evaluated through charging-session, occupancy, and queueing metrics. The immediate charging start rate measures the share of charging attempts that can start without waiting. Average and peak charger occupancy indicate how intensively the charging infrastructure is used. The number of charging sessions shows how often vehicles charge during the day. Queue-related metrics, including queue frequency, queue length, and charging waiting time, indicate whether charging infrastructure creates operational delays or vehicle unavailability.

The performance indicators used in the experiment analysis are summarised in Table 5.3.

Table 5.3: *Performance indicators used in the experiment analysis*

Category	**Performance indicators**
Service performance	On-time task completion rate (%)
	Unserved task rate (%)
	Average delay of late tasks (min)
Vehicle operations	Overall vehicle utilisation rate (%)
	Operational time shares: service, travel, charging, depot idle, waiting (%)
	Total travel distance (km)
	Dispatch-to-task travel distance (km)
Charging performance	Immediate charging start rate (%)
	Average charger occupancy (%)
	Peak charger occupancy (%)
	Number of charging sessions
	Share of time with a charging queue (%)
	Average charging queue length
	Maximum charging queue length
	Average waiting time for charging (min)
	Maximum waiting time for charging (min)

A central outcome derived from these metrics is the required fleet size. This is defined as the smallest tested fleet size that satisfies the selected service-level interpretation. Because different interpretations can be used, such as deterministic sufficiency, strict stochastic robustness, or a pragmatic lower bound, the required fleet size is not a single universal value. It depends on the performance target used to translate simulation outcomes into a fleet-capacity requirement.

5.6. Validation

This section assesses whether the simulation model is valid for its intended purpose: comparing fleet-size, charging, service-resource, and uncertainty settings for KLM water, toilet, and loader operations at Schiphol. The validation does not aim to reproduce every individual vehicle movement, charging decision, or operational intervention exactly. Instead, it evaluates whether the model produces plausible system-level behaviour that is sufficiently credible for scenario analysis and fleet-capacity assessment. The validation should therefore be interpreted as a fitness-for-purpose assessment rather than an exact calibration of the complete ground-handling system.

The validation focuses on five aspects. Validation Case A evaluates baseline fleet and service performance. Validation Case B assesses whether the generated task demand represents a plausible operational demand structure. Validation Case C evaluates energy use, SOC development, and charging behaviour. Validation Case D assesses water-filling and toilet-dumping behaviour for the vehicle groups with onboard resource constraints. Validation Case E evaluates the representation of operational uncertainty. Together, these cases examine whether the model inputs, process logic, and simulated outputs are consistent with the available operational data, model assumptions, and expert judgement.

Unless stated otherwise, the validation results are based on the deterministic baseline configuration. This means that task-timing and travel-time uncertainty are disabled in the baseline validation cases. The Key Performance Indicator (KPI) window covers a 24-hour period on 14 June 2026, starting with the first relevant task event in the simulation and ending 24 hours later. This fixed evaluation window is used to compare task completion, time-window performance, vehicle use, charging behaviour, and service-resource behaviour consistently across the three vehicle groups.

5.6.1. Validation Case A – Baseline Fleet and Service Performance

Validation Case A assesses whether the model produces plausible service performance under the current fleet configurations. Its purpose is to evaluate whether the model produces credible results for service performance, vehicle use, travel behaviour, and charging-infrastructure use under baseline conditions.

Water Vehicles

The water-service baseline was run with 11 vehicles. The KPI window covers the period from 14 June 2026 05:47 to 15 June 2026 05:47. Table 5.4 summarises the resulting KPI performance.

The water-service baseline completes the full deterministic demand pattern, but it is not entirely unconstrained. The main observation is that all tasks are served, while a small number of tasks still finish outside their time window. These late completions are short and caused by local pressure during concentrated demand periods.

The time-share pattern shows both operational slack and peak-period pressure. The depot idle share is substantial, which is expected for a practical fleet that must be available during demand peaks rather than continuously occupied throughout the day. However, the late completions show that this slack is not always available at the right moment or location. For water vehicles, the baseline therefore represents a useful operational reference: the fleet can complete all tasks, but local timing constraints can still become binding.

The workload is distributed across the full vehicle group, with a relatively narrow task range between vehicles. This suggests that the dispatching logic uses the fleet broadly rather than relying on only a small subset of vehicles. The total travel distance is also relevant because it includes more than direct dispatch-to-task travel. For water vehicles, return movements, charging-related movements, and filling-related movements contribute to the difference between dispatch-to-task distance and total driven distance.

The simulated travel distance was compared with track and trace data from the same operational week. The empirical dataset is incomplete and does not exactly match the simulated baseline configuration: it contains seven vehicles, missing observations for some vehicles, and several partial-day operations. After correcting for incomplete availability, the observed distance corresponds to an estimated 380.19 km per day for an 11-vehicle fleet. This is of the same order of magnitude as the simulated 425.4 km. However, the applied corrections are themselves approximate and may introduce additional uncertainty, as assumptions are required to compensate for missing and partial observations. These uncertainties can accumulate and propagate through the adjustment process.

Table 5.4: *Validation Case A: baseline fleet and KPI performance*

Category	**KPI**	**Result**
Fleet	Water vehicles	11
Service performance	Completed tasks	304 / 304
	Tasks started within time window	99.67%
	Tasks completed within time window	97.70%
	Unserved task rate	0.00%
	Late tasks	7
	Average delay of late tasks	2.9 min
	Median delay of late tasks	2.0 min
	Tasks started outside time window	1
	Average start delay outside window	3.0 min
Time shares	Service time share	26.58%
	Travel time share	10.60%
	Depot idle time share	54.26%
	Waiting time share	8.57%
	Charging time share	13.75%
Vehicle use	Overall vehicle utilisation rate	45.74%
	Tasks per vehicle	23–31
	Average tasks per vehicle	27.6
	Dispatch-to-task travel distance	298.8 km
	Total travel distance	425.4 km
Charging	Immediate charging start rate	100.00%
	Average charger occupancy	13.75%
	Median charger occupancy	9.09%
	Peak charger occupancy	63.64%
	Number of charging sessions	48
	Time with charging queue	0.00%
	Average charging queue length	0.00
	Maximum charging queue length	0
	Average charging waiting time	0.0 min
	Maximum charging waiting time	0.0 min

The comparison is therefore not an exact calibration, but it provides an indicative check that the simulated driving distance is operationally credible.

Charging does not constrain the water-service baseline. All charging sessions start immediately and no charging queue occurs. The late completions should therefore mainly be interpreted as the result of service timing, spatial assignment, and vehicle availability during peaks, rather than as a charging-capacity bottleneck.

Toilet Vehicles

The toilet-service baseline was run with 11 vehicles. The KPI window covers the period from 14 June 2026 05:28 to 15 June 2026 05:28. Table 5.5 summarises the resulting KPI performance.

Table 5.5: *Validation Case A: baseline fleet and KPI performance for toilet service*

Category	KPI	Result
Fleet	Toilet-service vehicles	11
Service performance	Completed tasks	301 / 301
	Tasks started within time window	100.00%
	Tasks completed within time window	100.00%
	Unserved task rate	0.00%
	Late tasks	0
	Average delay of late tasks	0.0 min
	Tasks started outside time window	0
	Average start delay outside window	0.0 min
Time shares	Service time share	18.28%
	Travel time share	12.05%
	Depot idle time share	63.04%
	Waiting time share	6.64%
	Charging time share	15.53%
Vehicle use	Overall vehicle utilisation rate	36.96%
	Tasks per vehicle	24–35
	Average tasks per vehicle	27.4
	Dispatch-to-task travel distance	273.3 km
	Total travel distance	489.3 km
Charging	Immediate charging start rate	100.00%
	Average charger occupancy	18.98%
	Median charger occupancy	11.11%
	Peak charger occupancy	77.78%
	Number of charging sessions	48
	Time with charging queue	0.00%
	Average charging queue length	0.00
	Maximum charging queue length	0
	Average charging waiting time	0.0 min
	Maximum charging waiting time	0.0 min

The toilet-service baseline achieves full compliance with the service windows. Compared with the water-service case, the key difference is that the available slack is sufficient to absorb deterministic demand peaks without causing late task completions. However, this does not imply that the fleet size is minimal to achieve the 100% on-time completion rate. It only shows that the current baseline configuration is sufficient under deterministic conditions.

The time-share pattern shows a high depot idle share and a lower utilisation rate than the water-service fleet. This is consistent with a baseline fleet that is dimensioned for peak availability rather than average workload. The task distribution also shows that all vehicles are used, although some variation remains because assignments depend on location, timing, charging state, and vehicle availability.

The total travel distance is higher than the dispatch-to-task distance, which indicates that toilet-service operations also generate relevant non-task travel. For this vehicle group, this is expected because the operational process can include depot returns, charging movements, and waste-disposal-related movements in addition to direct travel to aircraft stands.

Charging remains non-binding in the deterministic toilet-service baseline. Peak charger occupancy is relatively high, but the absence of queues and charging waiting time indicates that the configured charging infrastructure is still sufficient for the simulated baseline operation.

Lower-Deck Loaders

The loader baseline was run with 25 loaders. The KPI window covers the period from 14 June 2026 05:11 to 15 June 2026 05:11. Table 5.6 summarises the resulting KPI performance.

Table 5.6: *Validation Case A: baseline fleet and KPI performance for loaders*

Category	**KPI**	**Result**
Fleet	Loaders	25
Service performance	Completed tasks	147 / 147
	Tasks started within time window	100.00%
	Tasks completed within time window	100.00%
	Unserved task rate	0.00%
	Late tasks	0
	Average delay of late tasks	0.0 min
	Tasks started outside time window	0
	Average start delay outside window	0.0 min
Time shares	Service time share	16.18%
	Travel time share	1.20%
	Depot idle time share	80.51%
	Waiting time share	2.11%
	Charging time share	11.23%
Vehicle use	Overall vehicle utilisation rate	19.49%
	Tasks per vehicle	3–8
	Average tasks per vehicle	5.9
	Dispatch-to-task travel distance	25.2 km
	Total travel distance	58.3 km
Charging	Immediate charging start rate	100.00%
	Average charger occupancy	10.02%
	Median charger occupancy	0.00%
	Peak charger occupancy	75.00%
	Number of charging sessions	32
	Time with charging queue	0.00%
	Average charging queue length	0.00
	Maximum charging queue length	0
	Average charging waiting time	0.0 min
	Maximum charging waiting time	0.0 min

The loader baseline also achieves full service-window compliance, but the operational pattern differs strongly from the water and toilet-service fleets. The most notable feature is the high idle share and the low number of tasks per vehicle. This reflects the combination of a relatively large baseline loader fleet and a lower number of active loader-service tasks.

The low travel time share is also important. Loader vehicles typically remain close to or at the VOPs while on standby and do not require process-specific movements such as water filling or toilet dumping. As a result, loader availability is driven mainly by task timing, service duration, and local positioning rather than by long travel or replenishment movements.

The charging indicators show that charging does not constrain loader operations in the deterministic baseline. Peak charger occupancy reach a maximum of 75% at specific moments and the absence of charging queues means that charging infrastructure availability does not affect task completion in this run.

Overall interpretation

Across the three vehicle groups, Validation Case A shows that the model can reproduce plausible deterministic baseline behaviour. All active tasks are completed for water vehicles, toilet vehicles, and loaders, and no unserved tasks occur. The main difference between the vehicle groups is the degree to which the baseline fleet is locally constrained. Water vehicles show small time-window violations during peak periods, whereas toilet vehicles and loaders complete all tasks within their service windows. The loader fleet shows the largest amount of operational slack, reflected in its high idle share and low average number of tasks per vehicle.

The results were also discussed with an experienced KLM operations specialist, who confirmed that the simulated service performance and vehicle-use patterns are operationally reasonable. Charging behaviour is assessed separately in the energy and charging validation case. Overall, the baseline runs provide a plausible reference point for the subsequent scenario analysis, with results that are consistent with the model assumptions and with each other.

5.6.2. Validation Case B – Operational Task Demand

Validation Case B evaluates whether the active task demand used in the baseline simulation represents a plausible operational demand pattern for the selected Schiphol–KLM case-study day. In contrast to the verification of the preprocessing logic, this validation case does not focus on whether flight records are technically converted correctly into simulation tasks. Instead, it assesses whether the resulting active task sets have a realistic operational structure in terms of demand level, operation type, temporal distribution, aircraft mix, and service-duration composition.

Water Vehicles

The baseline task set contains 304 active water-service tasks. Table 5.7 summarises the operational structure of this demand set.

The water-service task set is dominated by turnaround-related work, while still including a smaller number of isolated arrival and departure tasks. This structure is consistent with the hub-operation setting of the case study, where many services are linked to aircraft turnarounds rather than to isolated movements.

The time-of-day distribution shows that water-service demand is concentrated in the morning and afternoon. This is relevant for fleet sizing because vehicle requirements are driven mainly by clustered peak demand rather than by average demand over the full operational day. The corresponding time-window pattern is shown in Figure 5.6.

The aircraft-class distribution is mixed, with tasks spread across commuter, narrow-body, and wide-body aircraft. The service-duration structure also contains both short routine services and a smaller group of longer tasks. The longer water-service tasks are especially relevant for capacity analysis because they occupy vehicles for more time and can therefore contribute to peak-period pressure.

Table 5.7: *Validation Case B: operational task demand and temporal structure for water service*

Category	Class	Tasks	Share
Total demand	Active water-service tasks	304	100.0%
Operation type	Arrival only (A)	24	7.9%
	Departure only (D)	40	13.2%
	Arrival–departure (AD)	240	78.9%
Time bucket	Morning, before 11:00	129	42.4%
	Afternoon, 11:00–17:00	105	34.5%
	Evening, after 17:00	70	23.0%
Aircraft class	Commuter	94	30.9%
	Narrow-body	120	39.5%
	Wide-body	90	29.6%
Service duration	5 min	130	42.8%
	10 min	84	27.6%
	15–19 min	40	13.2%
	28–35 min	50	16.4%

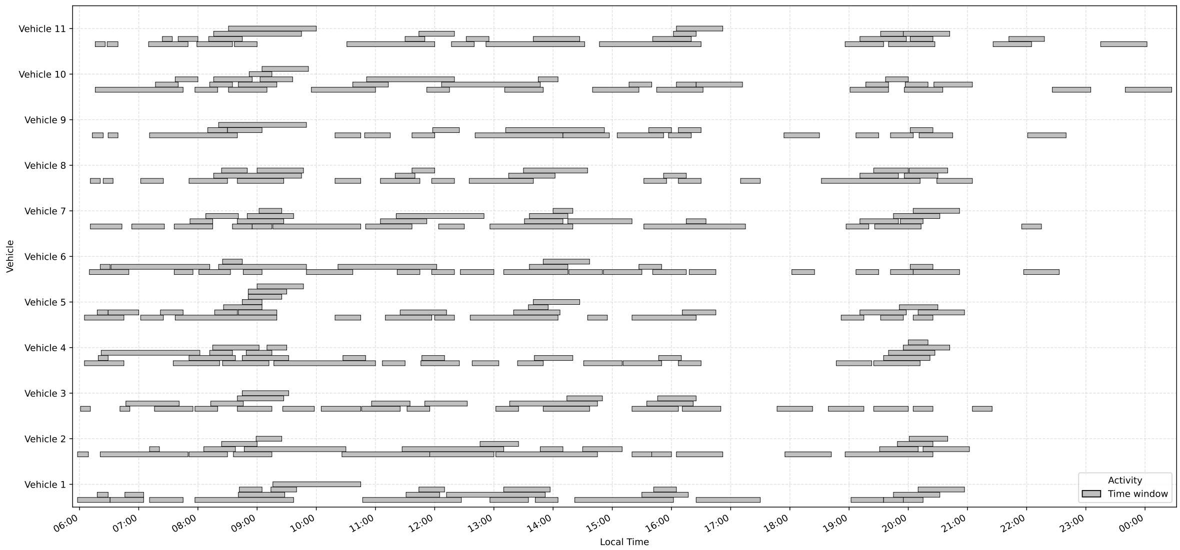


Figure 5.6: *Time-Windows of water services on 14-06-2026*

Toilet Vehicles

The baseline task set contains 301 active toilet-service tasks. Table 5.8 summarises the operational structure of this demand set.

The toilet-service task set is similar to the water-service task set in its overall structure. Turnaround-related AD tasks are again dominant, which is expected because toilet service is typically connected to the aircraft turnaround process. Compared with water service, the shares of isolated arrivals and departures differ slightly due to different service requirements, with departures more often requiring water service and arrivals more often requiring toilet service. Multistretch aircraft also introduce alternating water and toilet service needs for some subtypes. However, the overall demand structure remains comparable.

The temporal distribution is again peak-based rather than uniform. Morning and afternoon together account for most of the active demand, while the evening still contains a relevant number

Table 5.8: *Validation Case B: operational task demand and temporal structure for toilet service*

Category	Class	Tasks	Share
Total demand	Active toilet-service tasks	301	100.0%
Operation type	Arrival only (A)	37	12.3%
	Departure only (D)	24	8.0%
	Arrival–departure (AD)	240	79.7%
Time bucket	Morning, before 11:00	117	38.9%
	Afternoon, 11:00–17:00	103	34.2%
	Evening, after 17:00	81	26.9%
Aircraft class	Commuter	94	31.2%
	Narrow-body	117	38.9%
	Wide-body	90	29.9%
Service duration	5 min	127	42.2%
	10 min	84	27.9%
	11–13 min	90	29.9%

of tasks. This pattern is visible in Figure 5.7 and confirms that the selected baseline day contains demand clustering that is relevant for fleet-sizing analysis.

The aircraft mix is distributed across the same three aircraft classes as water service. The main difference is found in the service-duration structure. Toilet-service tasks are generally shorter and less dispersed than the water-service tasks, with no very long service-duration group. Nevertheless, the 11–13 minute tasks remain operationally relevant because they still keep vehicles occupied during peak periods.

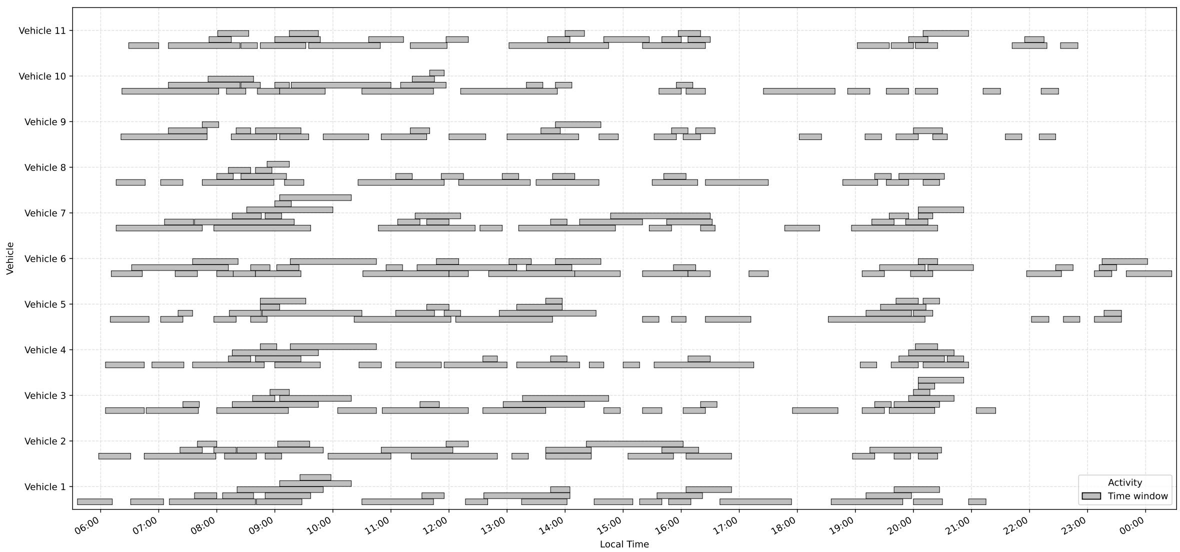


Figure 5.7: *Time-Windows of toilet services on 14-06-2026*

Lower-Deck Loaders

The baseline task set contains 147 active loader-service tasks. Table 5.9 summarises the operational structure of this demand set.

The loader task set differs clearly from the water- and toilet-service task sets. Loader demand is almost evenly split between arrival and departure tasks, because the model treats unloading and load-

Table 5.9: *Validation Case B: operational task demand and temporal structure for loaders*

Category	Class	Tasks	Share
Total demand	Active loader-service tasks	147	100.0%
Operation type	Arrival task (A)	75	51.0%
	Departure task (D)	72	49.0%
Time bucket	Morning, before 11:00	76	51.7%
	Afternoon, 11:00–17:00	54	36.7%
	Evening, after 17:00	17	11.6%
Aircraft class	Commuter	0	0.0%
	Narrow-body	0	0.0%
	Wide-body	147	100.0%
Service duration	29–36 min	63	42.9%
	40–41 min	35	23.8%
	46–49 min	38	25.9%
	60 min	11	7.5%

ing as separate operational processes. This is appropriate for loader operations, which are directly linked to either the arrival-side or departure-side handling process rather than to one combined replenishment task.

The temporal distribution is more concentrated than for water and toilet service. More than half of the active loader tasks occur in the morning period, while evening demand is limited. This pattern, also shown in Figure 5.8, means that loader fleet requirements are particularly sensitive to the first part of the operational day.

The aircraft-class distribution is specialised: all active loader tasks are associated with wide-body aircraft. This differs from the mixed demand structure of water and toilet service, but is consistent with the modelling scope for loaders in this case. The service-duration structure is also substantially longer than for water and toilet service. Even though the total number of loader tasks is lower, the longer task durations can create capacity pressure because loaders remain occupied for extended periods.

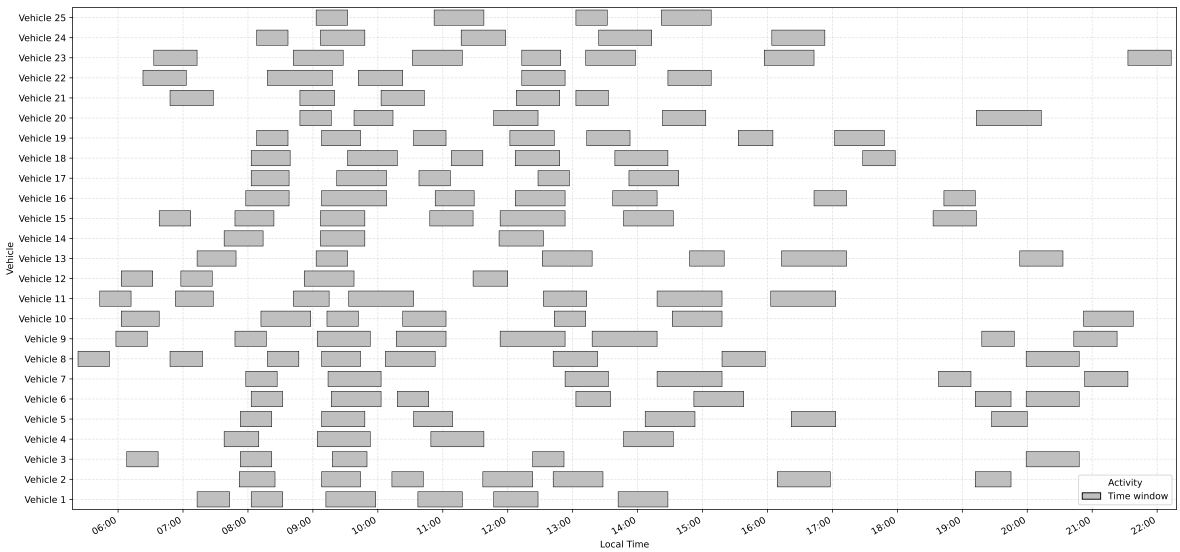


Figure 5.8: *Time-Windows of loader services on 14-06-2026*

Overall interpretation

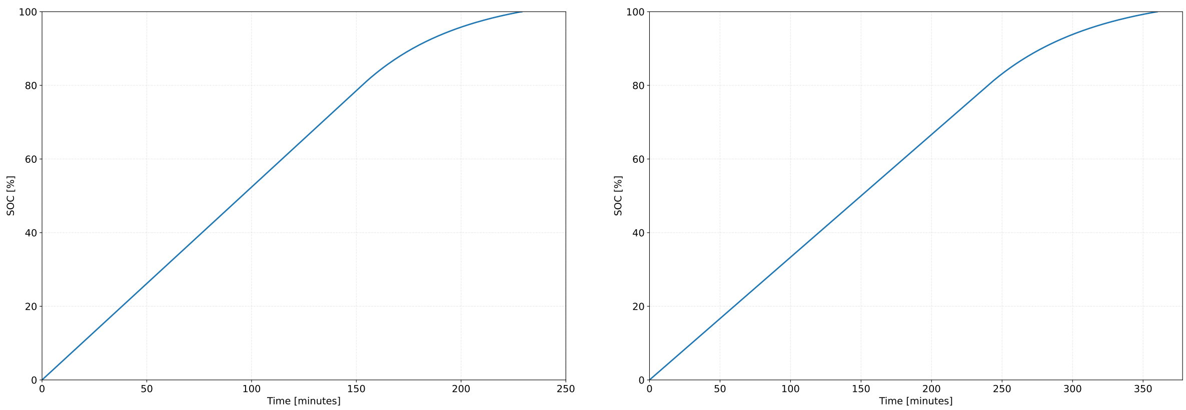
Across the three vehicle groups, Validation Case B shows that the baseline task sets capture distinct but operationally explainable demand structures. Water and toilet service are dominated by combined arrival–departure tasks, have mixed aircraft-class demand, and show a strong but not exclusive concentration in the morning and afternoon periods. Loader demand is structurally different: it is split between arrival and departure tasks, concentrated more strongly in the morning, limited to wide-body aircraft in this model, and characterised by substantially longer service durations.

The structure of the model input was also reviewed with operational experts. They confirmed that the total demand level, operation types, time buckets, aircraft-class distinction, and service categories are consistent with practical ground handling operations. The demand structures provide a credible input basis for the subsequent performance, fleet-sizing, and uncertainty analyses.

5.6.3. Validation Case C – Energy and Charging Behaviour

Validation Case C evaluates whether the model produces plausible energy and charging behaviour for the three case-study vehicle groups: water vehicles, toilet vehicles, and loaders. The checks include SOC development, charging-session timing, charging duration, cumulative charged energy, and consistency with available operational measurements or expert judgement.

The charging curve differs by vehicle group. For water and toilet vehicles, the model uses an assumed charger power of 22 kW, while loaders use an assumed charger power of 16 kW, as shown in Figure 5.9a and Figure 5.9b. A charging efficiency of 0.90 is applied. The charging curves define how much energy and SOC can be recovered during a charging session. They therefore provide the basis for translating charger power and charging time into vehicle-level battery recovery.



(a) *Water and toilet vehicles* **(b)** *Loaders*

Figure 5.9: *Charging curves of the simulated vehicle groups, showing state of charge as a function of charging time*

Water Vehicles

For water vehicles, the energy-consumption inputs are based on operational measurements collected at the beginning of July. The measured driving-energy consumption is 1.04 kWh/km, while service-energy consumption is set to 0.10 kWh per service [40]. The service-energy value represents a weighted average across commuter, narrow-body, and wide-body aircraft. A relatively large operational dataset was available, including SOC and charger-use data. The water-vehicle energy inputs therefore have the strongest empirical basis within this validation case. The simulated energy use, SOC development, and charging behaviour are compared with these operational measurements.

The simulated SOC profile in Figure 5.10 shows a high-SOC operating regime. Vehicles generally remain well above 60% SOC, with only one vehicle briefly reaching a minimum of approximately 55.4%. The time-weighted average SOC across the fleet is approximately 80.7%, which is close to the operational daily average of approximately 75–80%. In the operational data, vehicles were usually connected to a charger at relatively high SOC values, with an average charging start SOC of approximately 64–66%. This indicates that charging in practice is also primarily used for top-up charging rather than deep battery recovery. The relevant observation is therefore not that individual SOC trajectories match exactly, but that the model reproduces the operational pattern of repeated top-up charging instead of deep discharge.

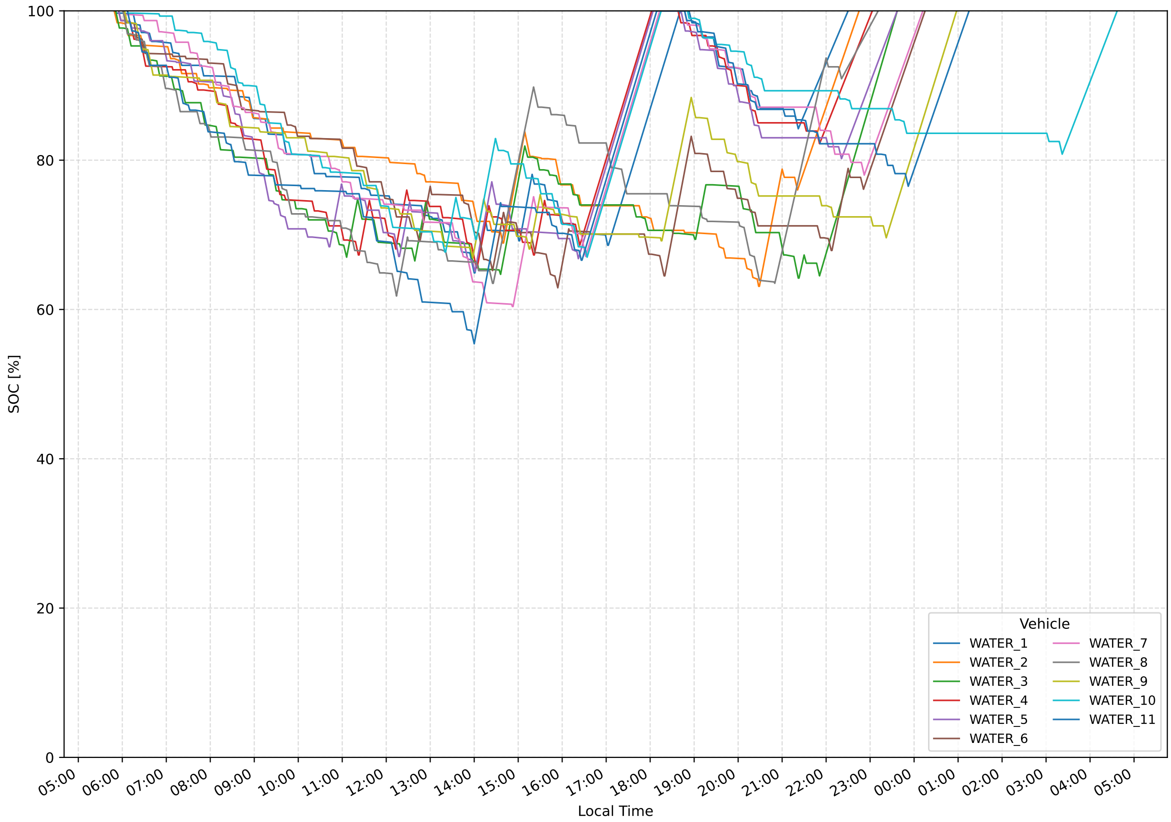


Figure 5.10: *State of charge over time per simulated water-service vehicle*

The operational dashboard in Figure 5.11 links service demand, vehicle deployment, and charging behaviour. Charger use is limited during periods with high vehicle deployment and increases when vehicles become available between service peaks. This pattern explains why SOC declines during service execution and recovers during operational gaps.

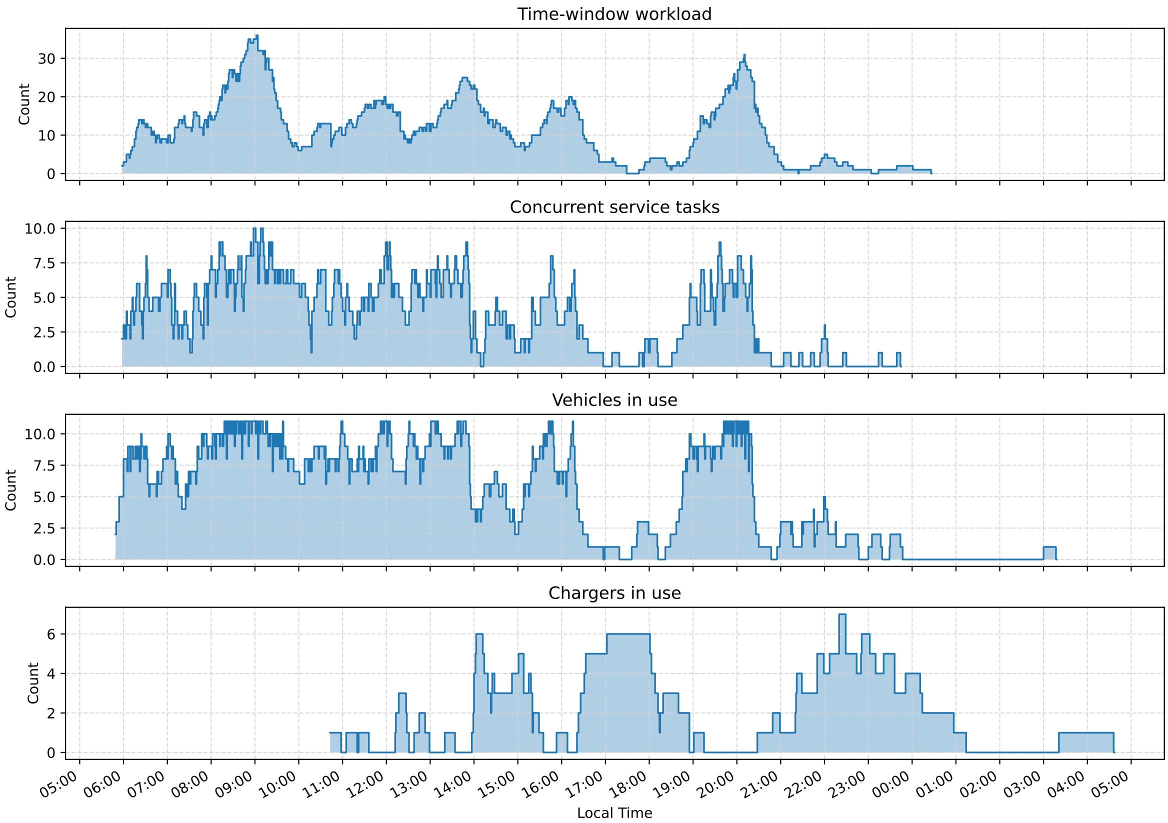


Figure 5.11: *Operational dashboard showing workload, concurrent service tasks, vehicles in use, and chargers in use for the simulated water-service vehicles*

The charger-use pattern in Figure 5.12 is broadly consistent with the operational reference data. In practice, simultaneous charging can reach up to seven vehicles, although most periods contain only one or two vehicles charging at the same time. The simulation reproduces the same maximum of seven simultaneous charging sessions. No persistent charging queue develops, meaning that charging infrastructure does not act as a binding constraint in the deterministic baseline.

Operational charging sessions most often started around 09:00, 12:00–13:00, 16:00, and 20:00–21:00. In the simulation, charging starts occur especially around 14:00–15:00, 16:00–18:00, and 21:00–23:00, with one additional late charging session after 03:00. The exact timing therefore differs, but the underlying operating logic is comparable: vehicles mainly charge between demand peaks rather than continuously throughout the day.

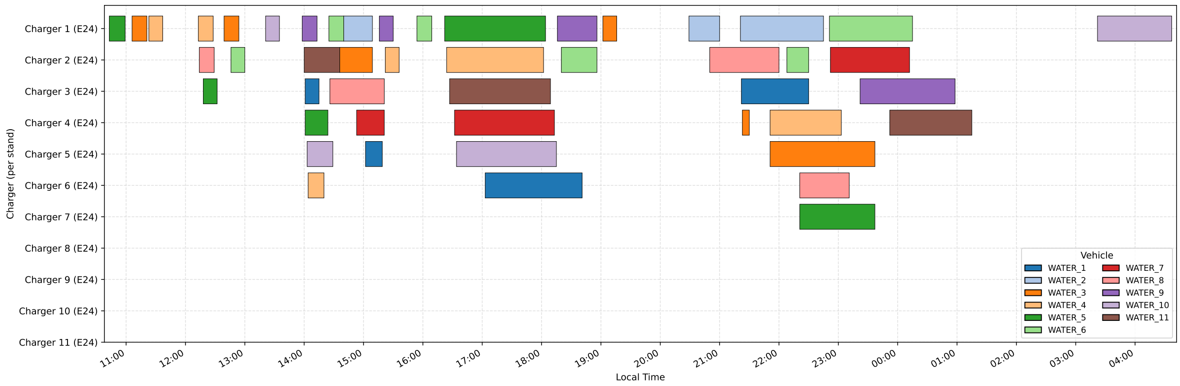


Figure 5.12: *Gantt chart of charger usage by simulated water-service vehicles*

Charging durations remain within the operationally observed range. In practice, most charging sessions last between 20 and 100 minutes, with a median duration of 50 minutes, although longer sessions can occur. In the simulation, 48 charging sessions occur, ranging from approximately 7 to 106 minutes. The average duration is approximately 45.4 minutes and the median duration is 30.0 minutes. The model therefore captures the broad duration range, but produces more short top-up sessions than the operational data. This follows from the implemented charging logic, in which vehicles can charge as soon as they are available at the depot and their SOC conditions allow charging.

The SOC increase per charging session is mostly within the commonly observed operational range. In the operational data, SOC increases are mostly below 40 percentage points, with a median of 22 percentage points and occasional larger increases up to approximately 80 percentage points. In the simulation, the average SOC increase is approximately 16.5 percentage points, the median is 14.8 percentage points, and the maximum is 35.5 percentage points. The model therefore captures regular top-up charging, but not the largest real-world charging increments. This is expected because real vehicles are not always plugged in immediately after returning to the depot, which can lead to lower starting SOC values and larger subsequent charging increments.

A semi-quantitative comparison was made using available charger-energy data. The observed charging energy in a March reference week was 423.29 kWh per day. Since the reference data includes other charger users, uses eight electric vehicles instead of the eleven simulated water-service vehicles, and was collected in March rather than June, it is not directly comparable to the model output. After correction to an estimated average June day, the operational reference value is 450.49 kWh per day. In Figure 5.13, the simulated cumulative charged energy reaches approximately 498.3 kWh by the end of the simulated day. This is 47.81 kWh higher than the adjusted operational reference value, corresponding to a difference of approximately 10.61%.

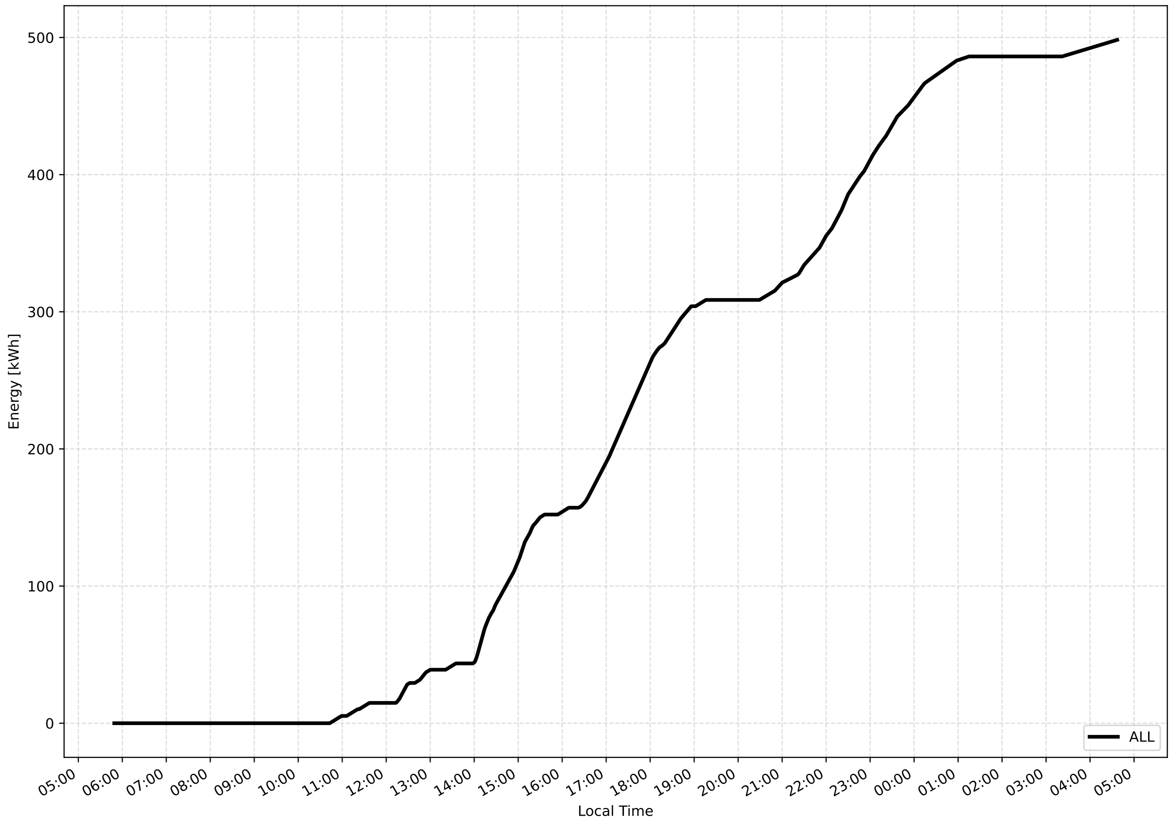


Figure 5.13: *Total cumulative charged energy for the simulated water-service vehicles*

The cumulative energy comparison should be interpreted as a semi-quantitative check. The simulated value is of the same order of magnitude as the corrected operational reference. Although corrections were made for differences in fleet size, vehicle selection, season, and charger-user composition, these corrections are only approximate and some deviations may remain. The remaining difference is therefore considered reasonable given the uncertainties involved.

Toilet Vehicles

For toilet vehicles, no separate operational energy-consumption measurements were available. The same energy-consumption parameters as for water vehicles were therefore used as a proxy: 1.04 kWh/km for driving and 0.10 kWh per service. This assumption is defensible because both vehicle groups are built by the same manufacturer and share the same electric chassis, resulting in similar vehicle characteristics. However, the toilet-vehicle energy representation is not independently calibrated with toilet-specific consumption data.

The simulated SOC profile in Figure 5.14 shows the same high-SOC operating regime as the water-vehicle case. Vehicles generally remain above 60% SOC, with only one vehicle briefly reaching a minimum of approximately 56.8%. The time-weighted average SOC across the simulated toilet fleet is approximately 81.9%, which is close to the operational daily average of approximately 75–80%. The main point is that the model does not create unrealistic deep-discharge behaviour; it represents toilet vehicles as repeatedly charged during operational gaps.

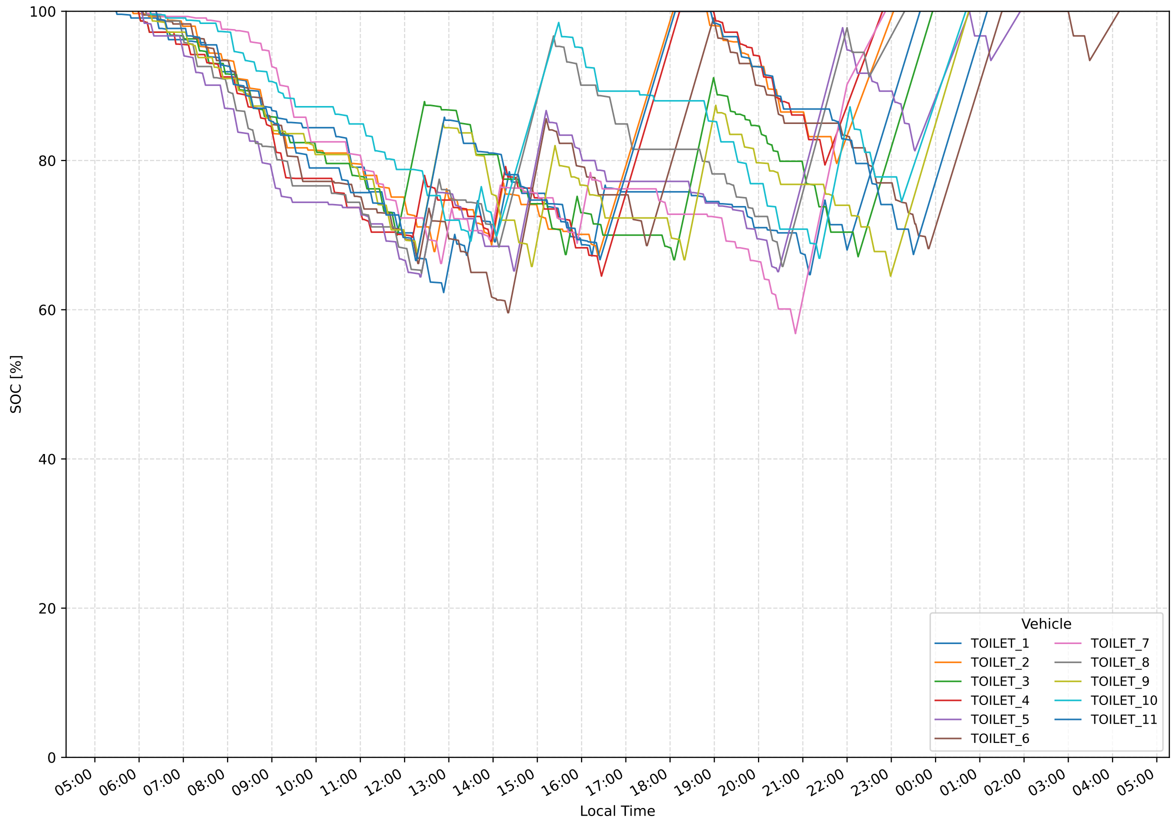


Figure 5.14: *State of charge over time per simulated toilet-service vehicle*

No separate operational charger-use data were available for toilet vehicles. The charger Gantt chart in Figure 5.15 is therefore used to assess internal consistency rather than direct empirical fit. Charging sessions are distributed across the available chargers, the maximum number of simultaneous charging sessions reaches seven vehicles, and no persistent charging queue develops. This means that charging infrastructure does not constrain the deterministic toilet-service baseline.

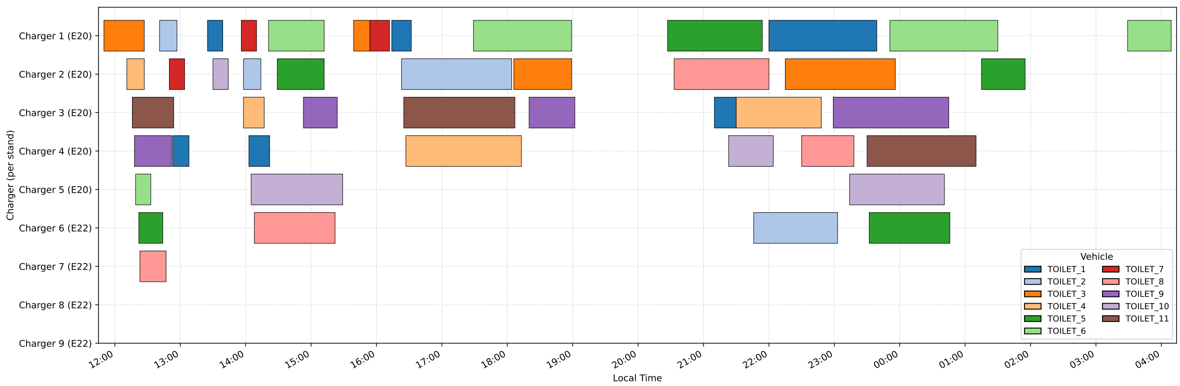


Figure 5.15: *Gantt chart of charger usage by simulated toilet-service vehicles*

The operational dashboard in Figure 5.16 shows that charger use increases mainly after vehicles become available between service peaks. This supports the SOC pattern in Figure 5.14: energy is consumed during service execution and recovered during periods when vehicles return to the depot.

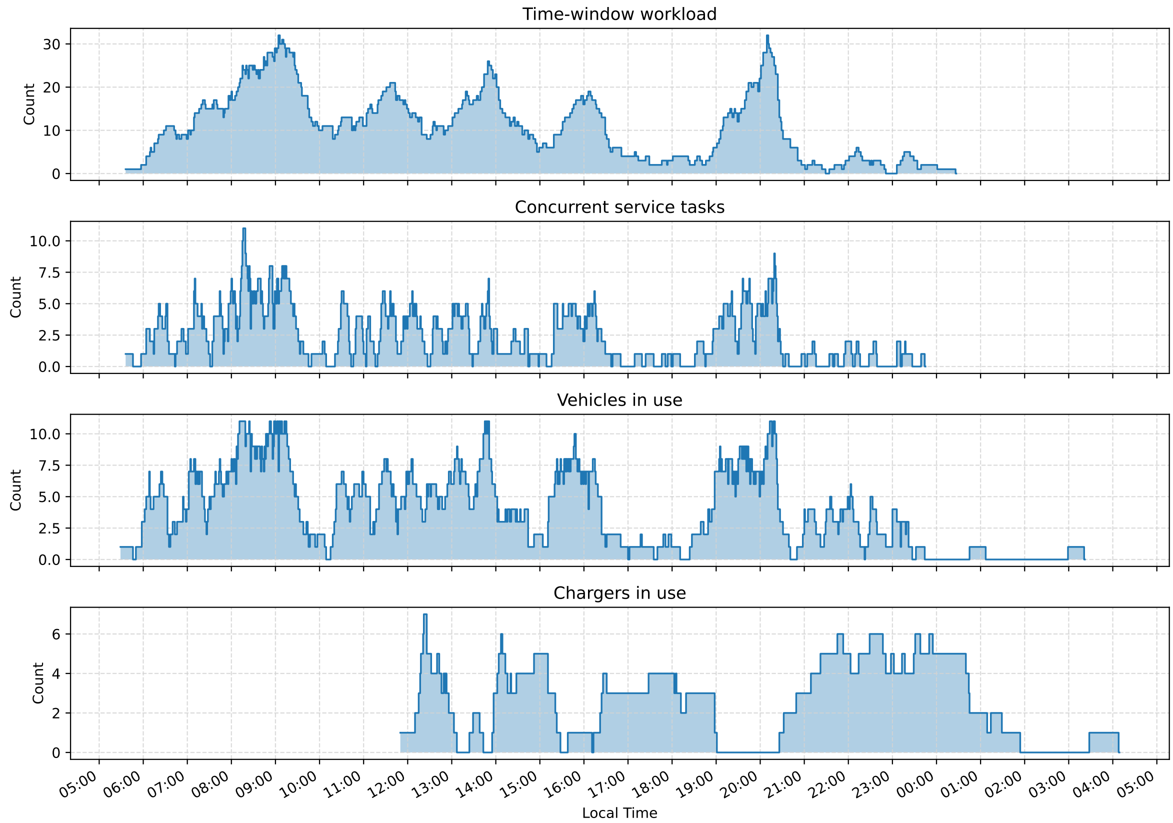


Figure 5.16: *Operational dashboard showing workload, concurrent service tasks, vehicles in use, and chargers in use for the simulated toilet-service vehicles*

Charging durations remain within a plausible range. In the simulation, 45 charging sessions occur, ranging from approximately 14 to 106 minutes. The average duration is approximately 51.4 minutes and the median duration is 41.0 minutes. The model therefore produces short to medium top-up sessions rather than long full charging cycles, which is consistent with the high-SOC pattern.

The SOC increase per charging event is also consistent with top-up charging. The average SOC increase is approximately 19.2 percentage points, the median is 18.7 percentage points, and the maximum is 43.2 percentage points. Most events therefore remain within a moderate SOC-increase range. The absence of very large SOC increases follows from the model logic, because vehicles are charged when they are available at the depot and their SOC falls below the charging threshold.

No operational charger-energy data were available for toilet vehicles. The cumulative charged energy in Figure 5.17 should therefore be interpreted as a simulated energy-output indicator rather than as a directly validated empirical result. In the deterministic baseline, cumulative charged energy reaches approximately 568.4 kWh by the end of the simulated day. The profile increases most strongly during periods with several simultaneous charging sessions, particularly in the afternoon and evening.

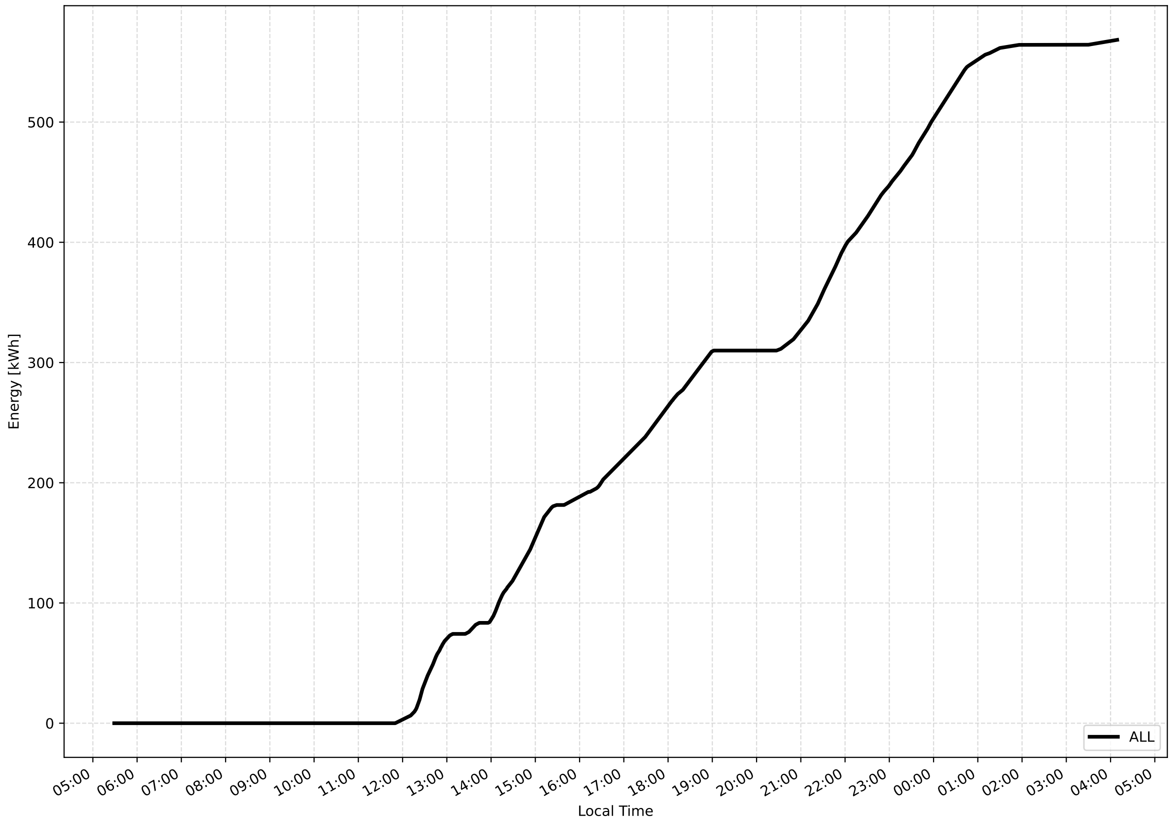


Figure 5.17: *Total cumulative charged energy for the simulated toilet-service vehicles*

As an additional check, the toilet-vehicle energy result was compared with the water-vehicle result. Both vehicle groups serve a similar number of tasks at approximately the same aircraft locations and use the same energy-consumption parameters. The toilet vehicles nevertheless consume about 10% more energy. This difference is mainly explained by the larger simulated travel distance. Toilet tasks have shorter service durations, so vehicles become available again more quickly. The dispatcher then spreads work across more vehicles than necessary, generating additional depot-to-task, return, dumping, and other support movements. In this run, toilet vehicles travel approximately 65 km more than water vehicles. The higher energy consumption is therefore linked to task sequencing and intermediate movements, not to higher service demand or different task locations.

The simulated charger use and energy consumption were discussed with an operational expert. The expert indicated that the resulting charging behaviour and energy demand could be realistic for electric toilet vehicles. However, this could not be fully verified with current operational data, because only a limited number of electric toilet vehicles are currently in use. The toilet validation is therefore mainly based on proxy input data, internal consistency of results, comparison with water vehicles, and expert judgement.

Loaders

For loaders, the energy-consumption input is based on operational measurements and calculations. Driving energy consumption is set to 1.6 kWh/km, while service energy consumption is set to an average of 3.2 kWh per task. These values are higher than those used for water and toilet vehicles, which is expected because loaders perform heavier operational tasks and consume more energy during service execution.

The simulated SOC profile in Figure 5.18 shows that loaders also operate in a high-SOC regime throughout the deterministic baseline. The vehicles start at 100% SOC and gradually discharge during the main operational period. The minimum SOC is approximately 63.7%, while the time-weighted average SOC across the loader fleet is approximately 82.9%. The important observation is that no unrealistic deep-discharge events occur. Energy is consumed during service execution and recovered later when operational demand decreases.

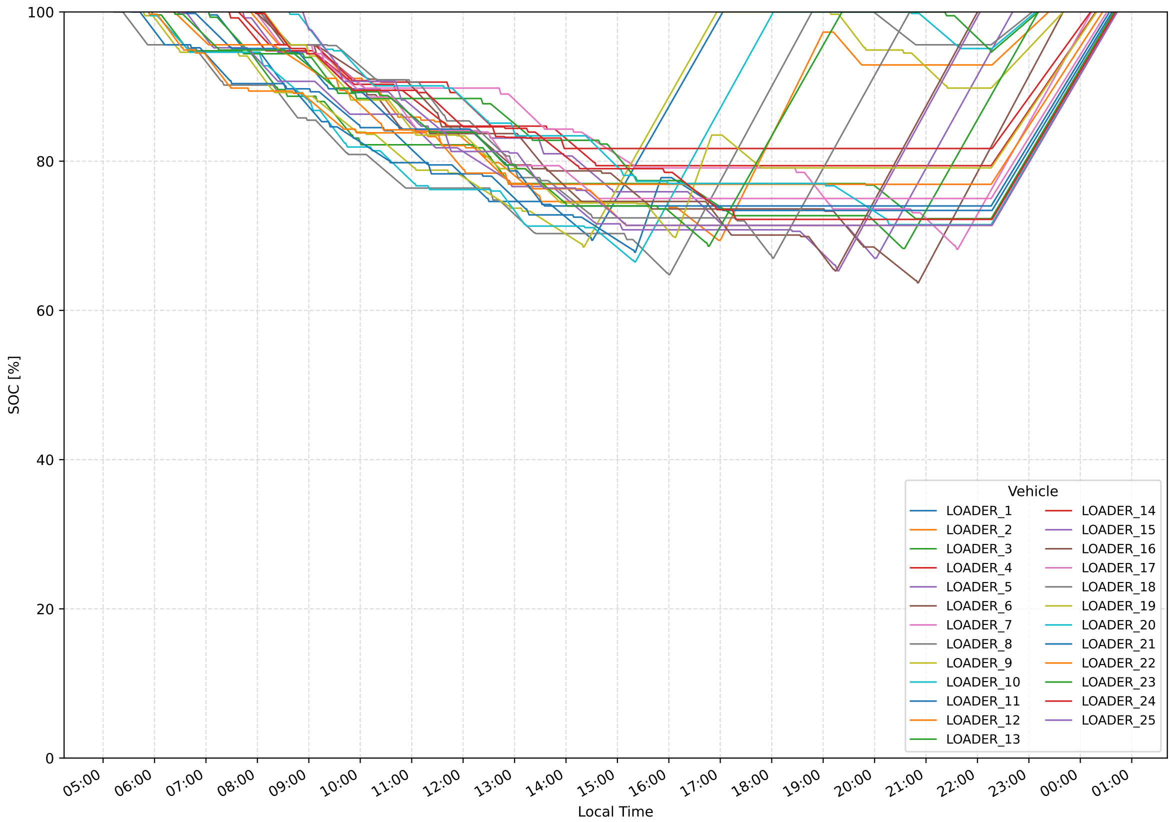


Figure 5.18: *State of charge over time per simulated loader*

The charger Gantt chart in Figure 5.19 shows a more concentrated charging pattern than for water and toilet vehicles. In the deterministic baseline, 32 charging sessions occur. The maximum number of simultaneous charging sessions reaches 21 chargers, mainly during the late-evening charging period. This shows that loader operations create a concentrated charging demand after the main service period. No persistent charging queue is visible in the baseline, but charger use is intensive during the evening peak.

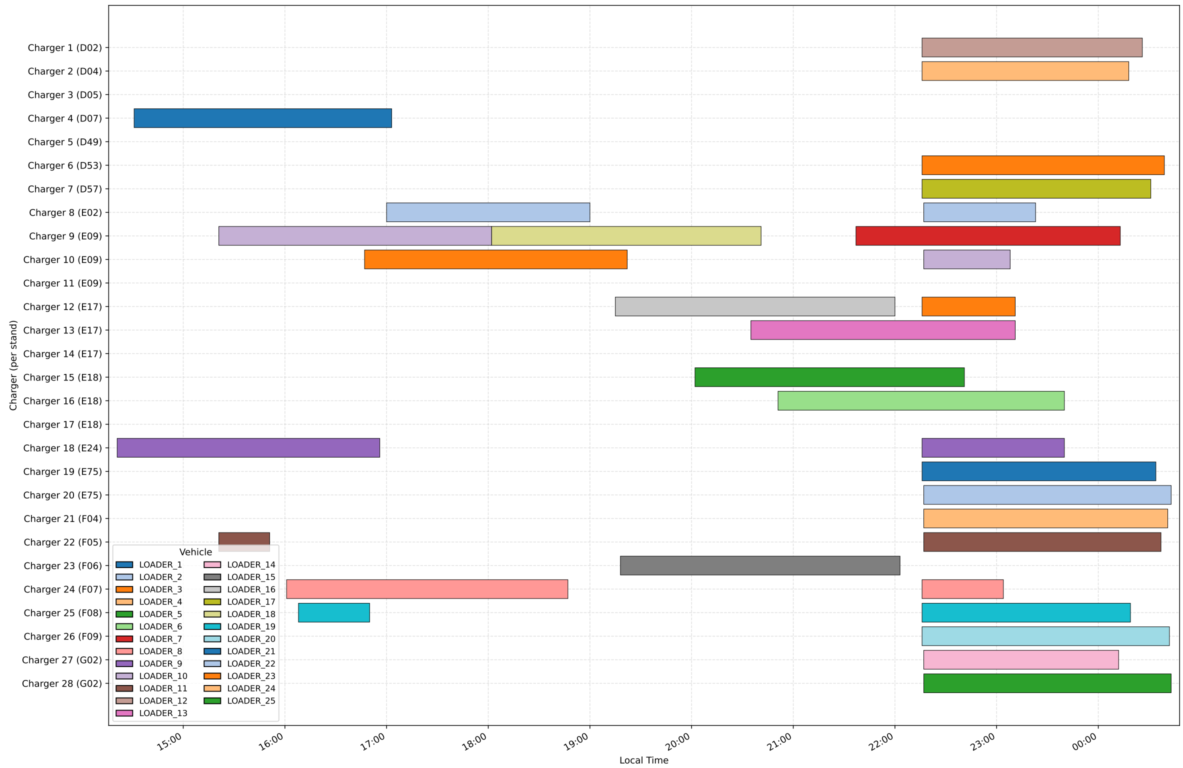


Figure 5.19: *Gantt chart of charger usage by simulated loaders*

The operational dashboard in Figure 5.20 links this charging pattern to the loader workload. Loader workload and concurrent service tasks are highest during the morning and early midday, with a peak of slightly above 20 simultaneous service tasks. Charger use remains relatively low during the main service period and rises later in the day. This confirms that loader energy is mainly consumed during operational peaks and recovered afterwards through concentrated depot charging.

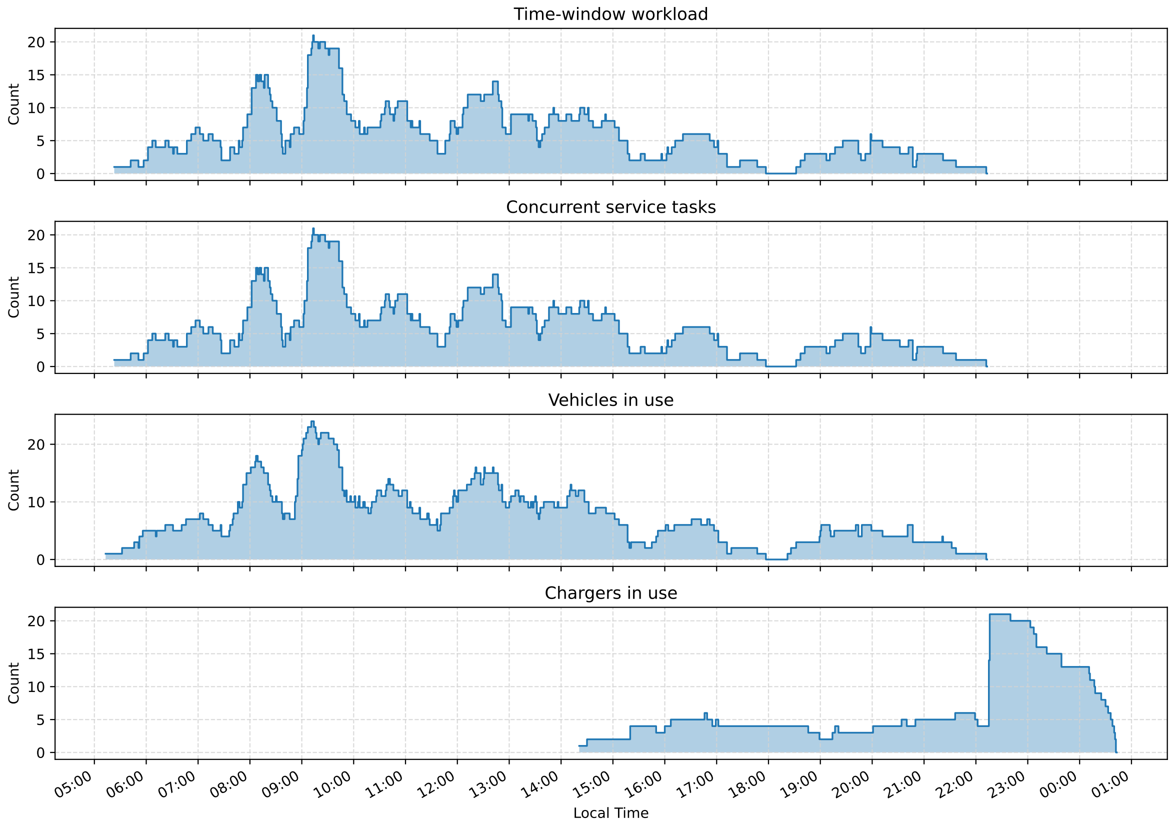


Figure 5.20: *Operational dashboard showing workload, concurrent service tasks, vehicles in use, and chargers in use for the simulated loaders*

Loader charging sessions are longer than those of water and toilet vehicles. The 32 charging sessions range from approximately 30 to 169 minutes. The average charging duration is approximately 126.3 minutes, while the median duration is 143.5 minutes. This follows directly from the higher energy-consumption assumptions for loaders. Several chargers are therefore occupied continuously for multiple hours, as visible in Figure 5.19.

The SOC increase per charging session remains moderate. The average SOC increase is approximately 24.5 percentage points, the median is 27.9 percentage points, and the maximum is 36.3 percentage points. The model therefore represents repeated top-up charging rather than full deep-cycle charging. This is consistent with the SOC profile, where vehicles generally remain above 60% SOC throughout the simulated day.

The cumulative charged energy in Figure 5.21 reaches approximately 563.5 kWh by the end of the simulated day. The cumulative profile increases gradually during the afternoon and rises more strongly during the late-evening charging period. This matches the charger-use pattern in Figure 5.19 and Figure 5.20, where the highest simultaneous charger use occurs after the main service demand has decreased.

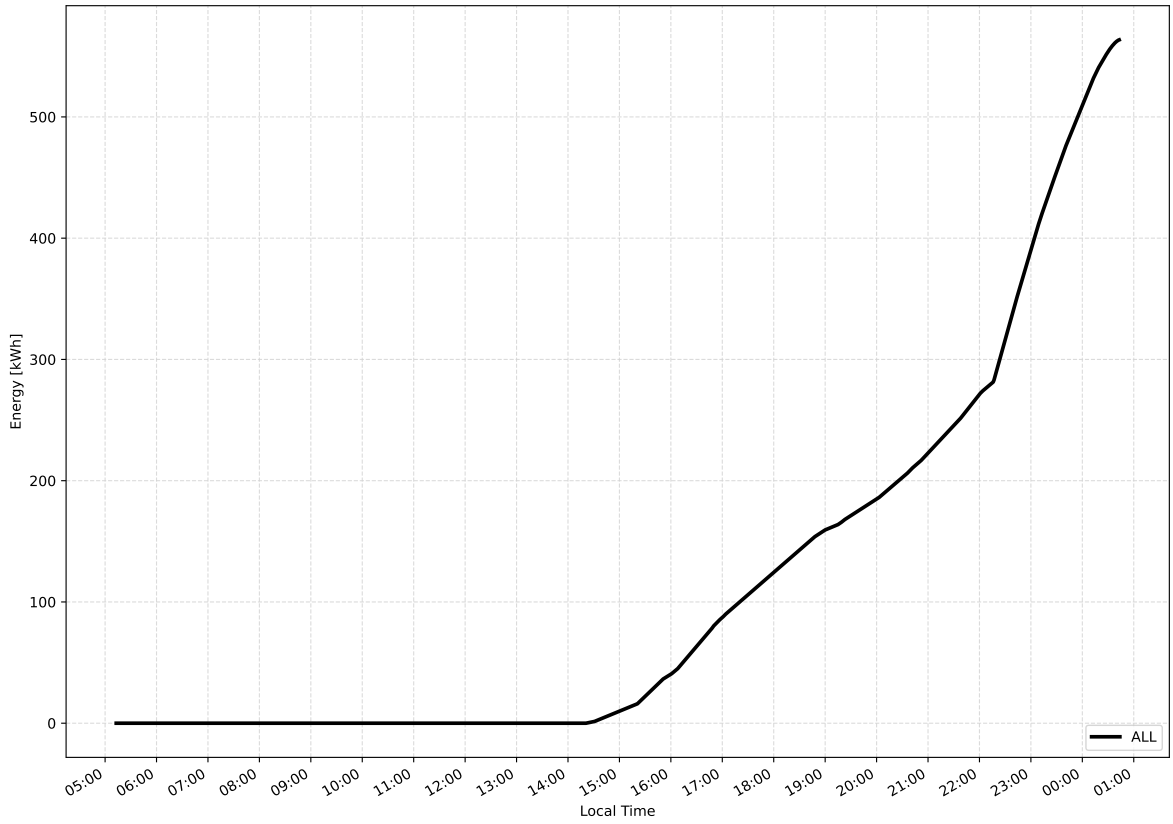


Figure 5.21: *Total cumulative charged energy for the simulated loaders*

For loaders, the validation is mainly based on measured and calculated energy-consumption input values, simulated SOC behaviour, charging-session structure, and expert judgement. The expert indicated that the simulated energy and charging behaviour should be interpreted with caution. In practice, not all loaders necessarily start the morning 100% SOC. In addition, loader SOC can sometimes decrease further during the day when vehicles are not charged immediately or when charging opportunities are missed. Charger use is also more spread across the operational day in practice, whereas the simulation produces a more concentrated charging pattern after the main service period.

Overall interpretation

Across the three vehicle groups, Validation Case C supports the plausibility of the model's energy and charging representation, but the strength and interpretation of the validation evidence differ by vehicle group. For water vehicles, the empirical basis is strongest because measured energy-consumption input values, SOC behaviour, charger-use patterns, charging-session durations, and cumulative charged energy can be compared with available operational reference data. For toilet vehicles, the simulated SOC behaviour follows the same high-SOC, repeated top-up pattern as observed in the available water/toilet reference data. However, the validation is more limited and relies mainly on proxy input data, comparison with water vehicles, internal consistency, and expert judgement.

For loaders, the baseline operation and operational task demand were confirmed by expert judgement as representative. However, the energy and charging behaviour is only partially validated. In practice, loaders do not necessarily start the day fully charged, SOC may decrease further when vehicles are not charged during the day, and charger use is normally more distributed across the operational day.

Overall, the validation supports the use of the energy and charging model for scenario analysis, charger-utilisation assessment, and fleet-sizing experiments. It does not prove that the model exactly reproduces observed charging energy for each vehicle group.

5.6.4. Validation Case D – Water Demand and Filling and Dumping Behaviour

Validation Case D evaluates whether the model produces plausible water-service demand, toilet-service demand, filling behaviour, and dumping behaviour in the deterministic baseline case. The purpose of this validation case is not to calibrate exact water or waste volumes, but to assess whether the generated task demand, assigned quantities, and intermediate resource activities follow operationally explainable patterns.

The filling and dumping tables distinguish between threshold-based and end-of-day activities. Threshold-based filling or dumping occurs during the operational day when the onboard level of a vehicle reaches the predefined operational threshold. End-of-day filling or dumping occurs after the main service demand has decreased and is used to return vehicles to an operationally ready state.

Water Vehicles

The water-service process is parameterised using operational rules for service frequency, onboard water capacity, and filling behaviour. Each water truck has a tank capacity of 4000 L and is sent to the fill point when the remaining onboard water level falls below 1000 L. Filling takes place at stand F04, where two fill-point resources are available in the model. Each filling task has a fixed duration of 14 minutes, which is close to the observed average filling duration of 13.6 minutes in the reference dataset. These settings determine when a vehicle temporarily becomes unavailable for service due to refilling.

The model also applies deterministic service rules to avoid overestimating water demand. For selected aircraft subtypes, water service is not required after every movement; instead, every second potential service opportunity is removed from the active task set. This multi-stretch rule is applied to the following subtypes: E70, E7W, E75, E90, 295, 735, 73H, 73J, and 73W. In addition, some candidate tasks are removed because water service is not required according to the service schedule, for example for latest-arrival movements that do not need to be prepared for a subsequent departure.

Table 5.10: *Validation Case D: water-service demand by aircraft class*

Aircraft class	Potential	Active	Active share	Avg. water/task [L]	Total water [L]
Commuter	186	94	50.5%	44.1	4,145.4
Narrow-body	170	120	70.6%	75.8	9,096.0
Wide-body	91	90	98.9%	485.8	43,723.3
Total	447	304	68.0%	187.4	56,964.7

Table 5.10 shows that the model generates water-service tasks selectively rather than for every potential opportunity. This is important because aircraft do not always require replenishment after each movement. The active share increases from commuter to narrow-body and wide-body aircraft, which is consistent with the larger operational role and higher replenishment likelihood of larger aircraft.

The assigned water volumes also scale logically with aircraft class. Wide-body aircraft account for fewer than one-third of the active water-service tasks, but they dominate the total simulated water demand because their average water quantity per task is much higher. This is operationally plausible, since larger aircraft carry more passengers and require larger onboard water volumes. The total simulated daily water demand is 56,964.7 L, and the total filled volume at a filling point is 56,967.5

L. The small difference is caused by rounding in the logged quantities and indicates that the model produces a consistent daily water balance.

Filling behaviour is governed by the onboard water level of each vehicle. A water truck is sent to the fill point when its remaining water level falls below the refill threshold of 1000 L. Filling takes 14 minutes per vehicle, and two vehicles can be filled simultaneously at F04. This allows the model to capture both refill movements and potential waiting effects when multiple vehicles require filling at similar times.

Table 5.11: *Validation Case D: filling behaviour of water-service vehicles*

Metric	Threshold filling	End-of-day filling	Total
Number of filling sessions	12	11	23
Total filled volume [L]	37,986.0	18,981.5	56,967.5
Average filled volume [L]	3,165.5	1,725.6	2,476.8
Average water before filling [L]	834.5	2,274.4	1,523.2
Minimum water before filling [L]	521.9	1,169.9	–
Average filling wait [min]	0.7	0.0	0.3
Maximum filling wait [min]	8.0	0.0	8.0
Maximum fill queue length	1	0	1

Table 5.11 shows a clear distinction between operational filling during the day and procedural filling at the end of operations. Threshold-based filling occurs when onboard water levels have fallen below the operational threshold, while end-of-day filling occurs at substantially higher remaining water levels. This confirms that the model does not treat all filling events as identical, but distinguishes between demand-driven refilling and preparation for the next operating period.

The filling process does not create a significant bottleneck in the deterministic baseline. Only limited waiting occurs, with a maximum fill queue length of one vehicle and a maximum waiting time of 8.0 minutes. Therefore, the late water-service task completions observed in Validation Case A should mainly be interpreted as a service scheduling and fleet-availability effect during concentrated demand periods, rather than as a filling-capacity problem.

Toilet Vehicles

The toilet-service process is parameterised using operational rules for service frequency, onboard waste capacity, and dumping behaviour. Each toilet truck collects wastewater during service tasks and is sent to the dump point when the onboard waste level reaches the dumping threshold. In the baseline run, threshold-based dumping occurs when vehicles carry approximately 2,600–3,000 L of waste. Dumping takes place at the dump point, where one dump-point resource is available in the model. Each dumping operation has a fixed duration of 15 minutes. These settings determine when a vehicle temporarily becomes unavailable for service due to dumping.

The same deterministic multi-stretch logic is applied to toilet service as to water service. For selected aircraft subtypes, toilet service is not required after every movement; instead, every second potential service opportunity is removed from the active task set. This rule is applied to the following subtypes: E70, E7W, E75, E90, 295, 735, 73H, 73J, and 73W. In addition, not every candidate task requires toilet service according to the service schedule. These rules prevent the model from overestimating toilet-service demand for aircraft that are not serviced after every turnaround or for movements where toilet service is not operationally required.

Table 5.12 shows a similar demand-filtering pattern to the water-service case. Not every potential service opportunity becomes an active task, and the active share increases for larger aircraft classes. This is operationally explainable because wide-body aircraft are more likely to require toilet servicing than commuter or narrow-body aircraft.

Table 5.12: *Validation Case D: toilet-service demand by aircraft class*

Aircraft class	Potential	Active	Active share	Avg. waste/task [L]	Total waste [L]
Commuter	356	211	59.3%	70.0	14,770.0
Narrow-body	91	36	39.6%	250.0	9,000.0
Wide-body	55	54	98.2%	400.0	21,600.0
Total	502	301	60.0%	150.7	45,370.0

The assigned waste volumes scale with aircraft class. Commuter aircraft generate an average of 70.0 L per task, compared with 250.0 L for narrow-body aircraft and 400.0 L for wide-body aircraft. Although wide-body aircraft account for only 54 of the 301 active toilet-service tasks, they generate 21,600.0 L, corresponding to approximately 47.6% of the total simulated waste volume. Commuter aircraft account for the largest number of active tasks, with 211 services, but contribute 14,770.0 L, or approximately 32.6% of the total volume. The total simulated daily toilet-service volume is 45,370.0 L, and the total dumped volume is also 45,370.0 L. This confirms that the model maintains a consistent daily waste balance.

Dumping behaviour is governed by the onboard waste level of each vehicle. Threshold-based dumping occurs during the operational day when accumulated service demand brings the onboard waste level close to the dumping threshold. End-of-day dumping is used to return vehicles to an empty or operationally ready state. Since only one dump-point resource is available, the model can also represent limited waiting when multiple vehicles require dumping close together.

Table 5.13: *Validation Case D: dumping behaviour of toilet-service vehicles*

Metric	Threshold dumping	End-of-day dumping	Total
Number of dumping sessions	12	10	22
Total dumped volume [L]	33,000.0	12,370.0	45,370.0
Average dumped volume [L]	2,750.0	1,237.0	2,062.3
Average waste before dumping [L]	2,750.0	1,237.0	2,062.3
Minimum waste before dumping [L]	2,640.0	350.0	–
Average dumping wait [min]	2.9	4.6	3.7
Maximum dumping wait [min]	16.0	15.0	16.0
Maximum dump queue length	2	1	2

Table 5.13 shows that threshold-based dumping occurs at high onboard waste volumes, while end-of-day dumping occurs at lower average volumes. This distinction is operationally plausible: during the day, dumping is triggered by accumulated waste demand, whereas end-of-day dumping is a procedural reset to prepare vehicles for subsequent operations.

Dumping introduces more waiting than water filling, which is expected because only one dump-point resource is available. However, the waiting remains limited in relation to the full operating day. The maximum dump queue length is two vehicles and the maximum waiting time is 16.0 minutes. Since the deterministic baseline still completes all toilet-service tasks on time, dumping does not act as a major service-performance bottleneck in this case.

Overall interpretation

Across both water and toilet service, Validation Case D shows that the model represents intermediate resource processes in an operationally plausible way. Service tasks are generated selectively rather than for every potential flight movement, and the active task shares differ logically between aircraft classes. The assigned water and waste volumes scale with aircraft class, with wide-body aircraft accounting for a large share of total daily volume. The model also produces internally consistent daily balances: water delivered during service is replenished through filling, and waste collected during service is discharged through dumping.

The filling and dumping results show that intermediate resource use follows accumulated operational demand rather than random activity patterns. Water filling does not create a significant bottleneck, while toilet dumping introduces limited waiting because only one dump-point resource is available. However, this waiting does not constrain deterministic baseline service performance significantly.

Direct measured water-meter and waste-volume data were not available for one-to-one comparison. The validation therefore supports the use of the water-filling and toilet-dumping logic for scenario analysis, while recognising that exact volume calibration would require more detailed operational measurement data.

5.6.5. Validation Case E – Uncertainty Representation

Validation Case E assesses whether the uncertainty representation used in the case-study model is plausible for the Schiphol–KLM operational context. This validation case applies to all three vehicle groups: water vehicles, toilet vehicles, and loaders. The technical implementation of the stochastic mechanisms was verified in Chapter 4. Therefore, this validation case does not examine seed reproducibility or whether the stochastic code functions as intended. Instead, it evaluates whether the uncertainty inputs are sufficiently grounded in empirical data and whether the resulting stochastic behaviour represents plausible operational variability.

The uncertainty representation consists of two components. The first component is arrival-time uncertainty, which shifts the timing of aircraft-dependent service tasks. For each flight, the model samples whether the aircraft arrives late, early, or on time. When a late or early arrival is sampled, the corresponding delay magnitude is drawn from an empirical distribution and applied as a signed shift to the planned task time. The second component is travel-time uncertainty, which increases the duration of selected vehicle movements. This component represents small airside disturbances, such as apron congestion, interactions with other traffic, short waiting periods, or inefficient routing behaviour that is not explicitly captured by the aggregated distance matrix.

The arrival-time uncertainty profiles are derived from historical operational flight data and segmented by time of day. Separate empirical uncertainty profiles are therefore used for the morning, afternoon, and evening periods instead of applying one generic delay distribution to the full day. Both late and early arrivals are included because deviations in either direction can shift the timing of ground-service tasks. The empirical data show that late-arrival probabilities differ between the time buckets and range from approximately 35.0% to 60.7%. This daypart-specific structure is relevant because operational variability at Schiphol is influenced by the hub-wave structure, accumulated upstream delays, and changes in traffic intensity throughout the operating day. The realised late- and early-arrival uncertainty probabilities over the 30 replications are presented in Figure 5.22.

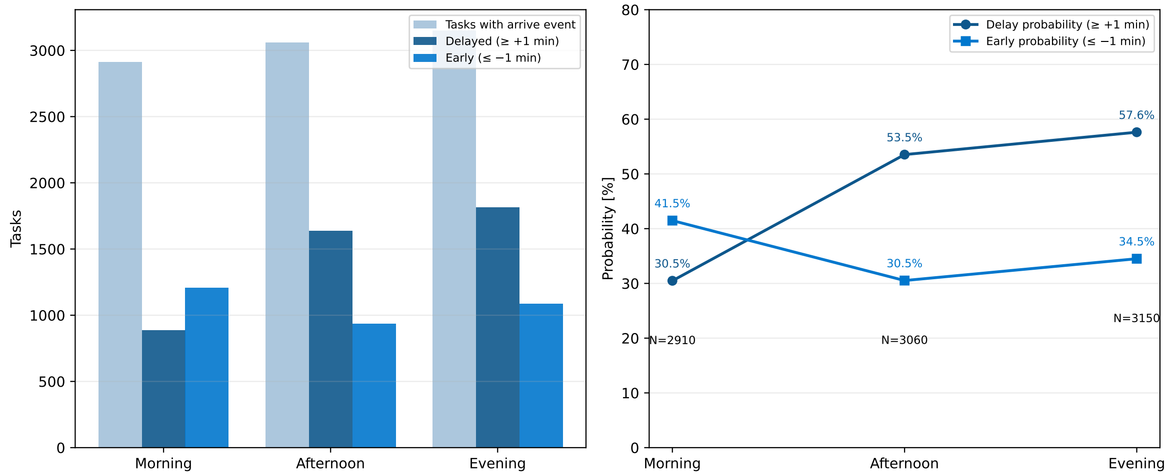


Figure 5.22: *Probability of late- and early-arrival uncertainty over 30 simulation runs*

Arrival-delay magnitudes are sampled through empirical resampling rather than by fitting a parametric distribution. The simulation therefore does not assume that arrival delays follow a predefined theoretical distribution, such as a normal or lognormal distribution. Instead, delay magnitudes are sampled directly from the historical values in the dataset. This approach preserves the observed shape, spread, and irregularity of the delay data. Separate empirical magnitude arrays are constructed for late and early arrivals within each daypart. Before being exported to the simulation model, the values in these arrays are capped at the 99th percentile. This approach retains the right-skewed structure of the historical delay data while limiting the influence of extreme observations and potential timestamp artefacts. The resulting uncertainty mechanism therefore combines empirical grounding with a bounded representation of operational variability.

The sampled delay behaviour was compared with the empirical profiles used to parameterise the model. Figure 5.23 and Figure 5.24 show the applied late- and early-arrival uncertainty for one simulation run. A single run is not expected to reproduce the empirical profiles exactly because it represents only one stochastic realisation. Individual runs may therefore contain different delay frequencies and magnitudes, provided that the sampled outcomes remain consistent with the empirical probabilities and capped magnitude arrays.

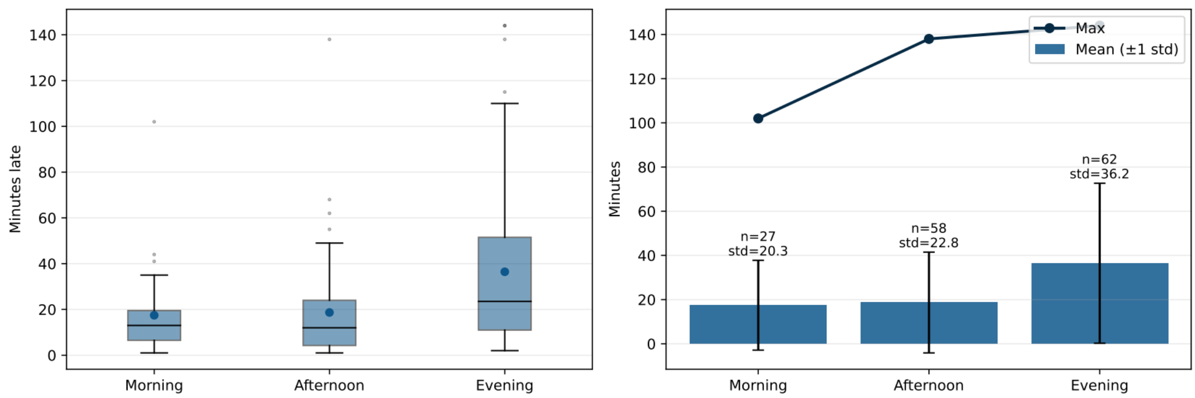


Figure 5.23: *Application of late-arrival uncertainty in one simulation run*

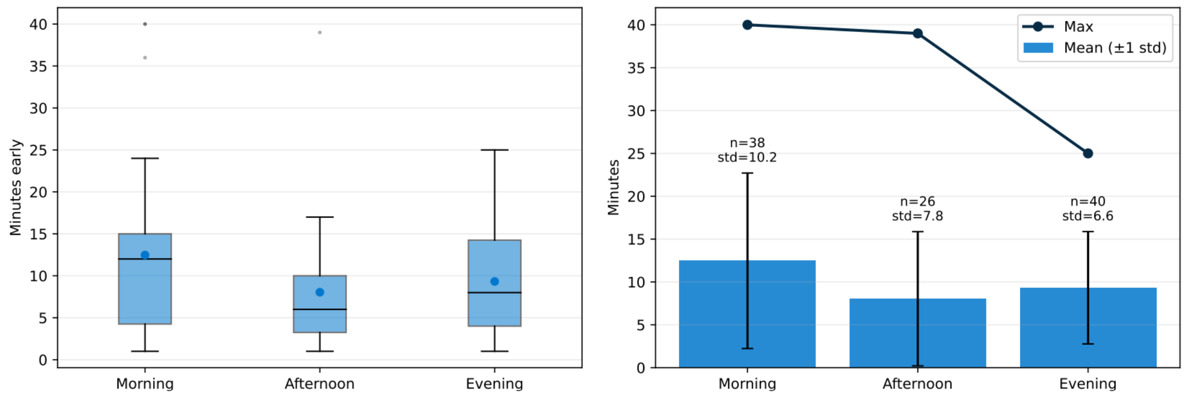


Figure 5.24: *Application of early-arrival uncertainty in one simulation run*

The aggregated results over 30 simulation runs provide a stronger plausibility check. Figure 5.25 and Figure 5.26 show that repeated sampling produces more stable late- and early-arrival patterns than a single stochastic run. The realised probabilities were compared with the empirical input probabilities for each time bucket. The results show that the simulated shares of late and early arrivals are consistent with the corresponding input probabilities for the morning, afternoon, and evening periods, as presented in Figure 4.7. Small deviations are expected because the uncertainty is sampled stochastically. However, across multiple runs, the realised probabilities converge towards the empirical input probabilities.

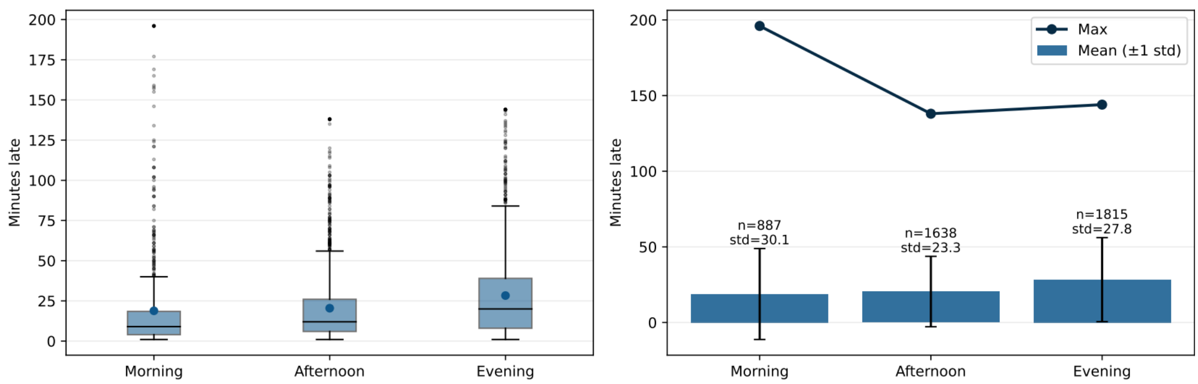


Figure 5.25: *Application of late-arrival uncertainty over 30 simulation runs*

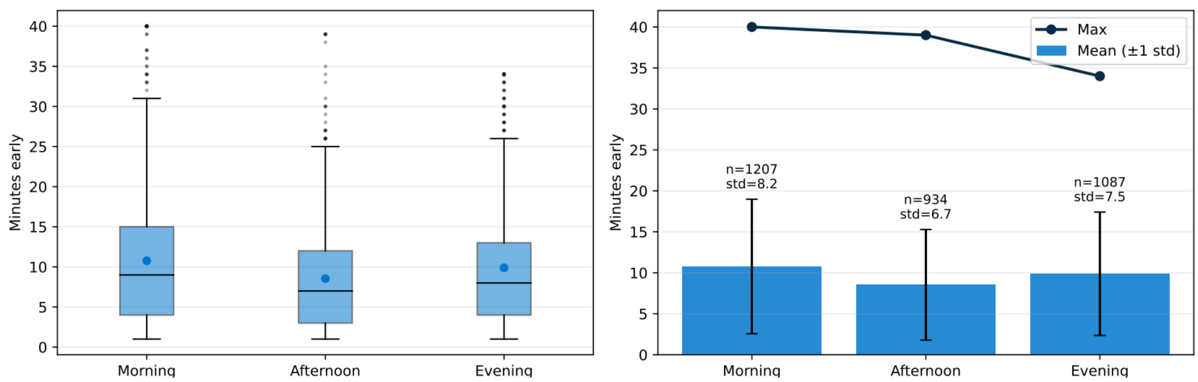


Figure 5.26: *Application of early-arrival uncertainty over 30 simulation runs*

A comparison with the empirical input profiles further shows that the stochastic sampling produces late- and early-arrival patterns with a similar structure and order of magnitude to the historical data from which the uncertainty profiles were derived. The empirical input profiles are shown in Figure 4.9 and Figure 4.11. The uncertainty representation therefore captures both the direction and the approximate scale of the observed arrival-time variability.

These results indicate that the 30 replications provide a sufficiently representative sample of the implemented task-timing uncertainty for the purpose of the simulation experiments. However, this representativeness applies only to the uncertainty mechanisms included in this study. The replications represent repeated samples from the empirical arrival-time profiles and the implemented travel-time disturbances. They do not cover all possible operational disruptions, such as weather conditions, vehicle breakdowns, variation in service durations, or alternative demand days. The results therefore support the validity of the implemented stochastic sampling but do not demonstrate robustness against all real-world sources of uncertainty.

Travel-time uncertainty is included as a secondary source of uncertainty. Unlike arrival-time uncertainty, this component is not calibrated using detailed vehicle-level travel-time observations. It should therefore be interpreted as an operational modelling assumption rather than as an empirically calibrated distribution. In the current configuration, each travel leg has a 30% probability of receiving a 15% increase in travel time. Because the deterministic travel times between most aggregated airport locations are relatively short, this increase often corresponds to approximately one additional minute.

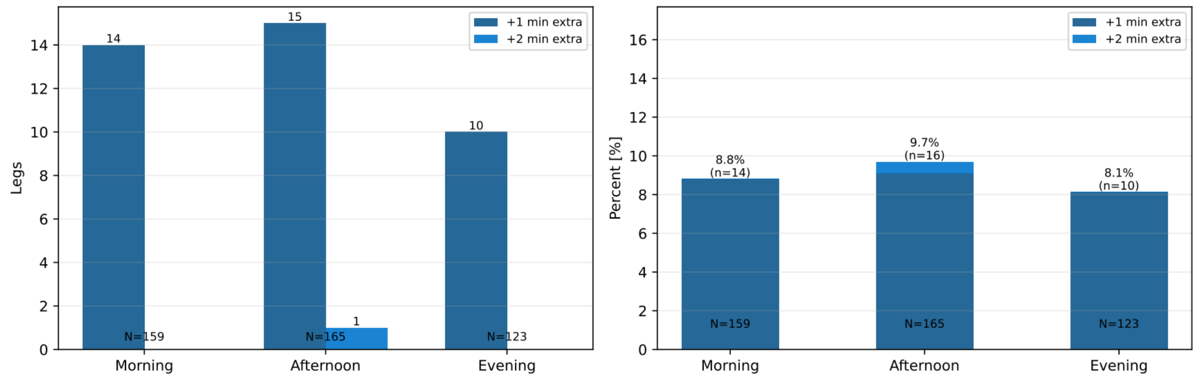


Figure 5.27: *Application of travel-time uncertainty in one simulation run*

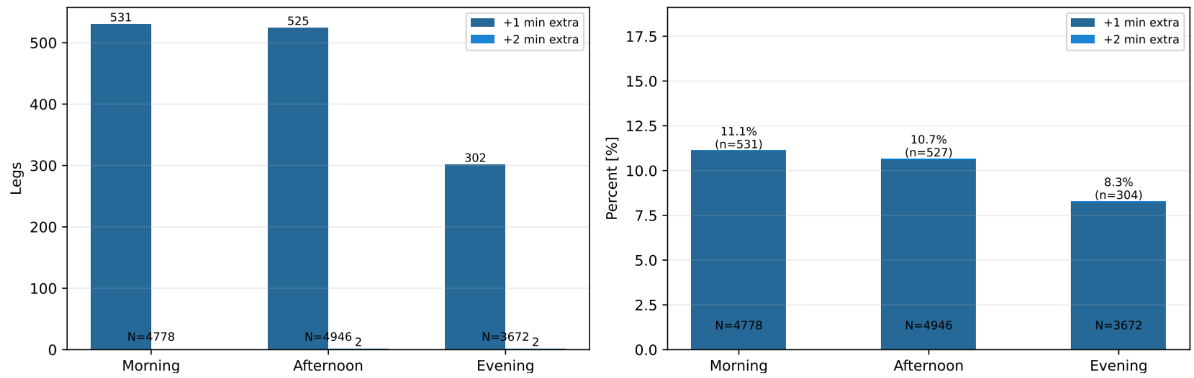


Figure 5.28: *Application of travel-time uncertainty over 30 simulation runs*

The results in Figure 5.27 and Figure 5.28 show that the mechanism introduces bounded variation in vehicle travel times without generating extreme delays. This is appropriate for the case-study model because the aggregated distance matrix does not explicitly represent local airside disturbances. The travel-time uncertainty settings were also reviewed with a Manager GSE, who considered the simulated variation in travel durations plausible for airside ground-handling operations.

Overall, Validation Case E supports the use of the uncertainty representation across the three vehicle groups. Arrival-time uncertainty is grounded in historical flight data, segmented by time of day, and implemented through capped empirical resampling. Travel-time uncertainty has a weaker empirical basis but provides a transparent and bounded representation of small disturbances in apron movements. The uncertainty model should therefore be interpreted as an empirically informed experimental mechanism rather than as an exact prediction of daily disruptions. It is suitable for assessing the sensitivity of fleet-size and charging conclusions to plausible operational variability. The main validation checks for each vehicle group are summarised in Table 5.14.

Table 5.14: *Summary of validation checks by vehicle type*

Validation check	Observed evidence	Water	Toilet	Loaders
Baseline service performance	All vehicle groups were evaluated on completed tasks, on-time task completion, and unserved tasks under deterministic baseline conditions.	✓	✓	✓
Operational task demand	Generated task volumes and daily demand patterns were checked against the expected hub-based structure, including morning, afternoon, and evening demand peaks.	✓	✓	✓
Energy and charging behaviour	Energy use, charging sessions, charger occupancy, and charging timing were checked for plausibility and order-of-magnitude consistency.	✓	✓	✓
Service-specific resource logic	Onboard resource constraints and intermediate service-point behaviour were checked where relevant, including filling for water vehicles and dumping for toilet vehicles.	✓	✓	–
Uncertainty representation	Arrival-delay profiles were derived from empirical flight data, while travel-time uncertainty was implemented as a stochastic disturbance based on expert judgement.	✓	✓	✓
Expert validation	Operational assumptions, service performance, vehicle use, travel uncertainty, and model behaviour were reviewed with KLM expert input.	✓	✓	✓

5.6.6. Scope and Limitations of Validation

The validation in this section assesses whether the model is sufficiently realistic for its intended purpose: comparing fleet-size, charging, service-resource, and uncertainty settings for water vehicles, toilet vehicles, and loaders in the KLM Schiphol case. The aim is not to reproduce every individual vehicle movement, charging decision, or operational intervention exactly. Instead, the validation examines whether the model produces plausible system-level behaviour for scenario analysis and fleet-capacity assessment. No validation claim is made for the other GSE types included in the broader potential scope of the model.

The validation combines three approaches. First, deterministic baseline performance is compared with the current operational reference configuration. Second, selected model outputs are compared with available operational data. These outputs include service performance, task-demand structure, energy use, charging behaviour, SOC development, and, where applicable, filling and dumping behaviour. Third, model assumptions and outputs that could not be fully validated using data were reviewed with KLM operational experts. These expert reviews covered baseline fleet use, task-demand structure, service logic, charging behaviour, water and waste-resource assumptions, travel-time uncertainty, and the plausibility of vehicle utilisation.

Several limitations remain. The available operational data do not provide a complete one-to-one reference for the simulated system. Some datasets include multiple vehicle types or charger users, whereas the simulation evaluates water vehicles, toilet vehicles, and loaders separately. In addition, some reference data were collected during a different period from the simulated case-study day. The validation should therefore be interpreted mainly as an order-of-magnitude and plausibility assessment rather than as an exact calibration of all model parameters.

The strength of the validation differs across vehicle groups and processes. Baseline service performance and task-demand patterns are assessed for all three vehicle groups. Energy use and charging behaviour are also evaluated for all three groups, although the available SOC and charger-use data do not always provide a vehicle-specific empirical reference. The empirical basis is strongest for water vehicles, more limited for toilet vehicles, and partly qualitative for loaders. Expert feedback indicated that the simulated loader baseline and task demand are representative, but that the simulated charging pattern does not fully reflect current operational practice.

Service-resource validation applies only to water and toilet vehicles because these vehicles have onboard resource constraints and require intermediate visits to service points. For water vehicles, the validation concerns water demand and filling behaviour. For toilet vehicles, it concerns waste accumulation and dumping behaviour. In both cases, direct measurements of service volumes per task were unavailable. The validation therefore remains an operational plausibility assessment rather than an exact calibration. This type of validation is not applicable to loaders because they do not have onboard water or waste-capacity constraints.

The strength of the uncertainty validation also differs by uncertainty source. Arrival-time uncertainty has the strongest empirical basis because the delay profiles are derived from historical flight data and differentiated by time of day. Travel-time uncertainty is less strongly validated because detailed vehicle-level travel-time observations were unavailable. It is therefore included as a bounded modelling assumption that represents small movement disturbances, such as local apron congestion, routing inefficiencies, and short waiting interactions. The uncertainty model should consequently be interpreted as an empirically informed experimental mechanism rather than as an exact prediction of daily disruptions.

The model scope further limits the interpretation of the validation results. The simulation represents selected GSE operations and their interaction with vehicle availability, charging, and service-specific resource constraints. It does not represent the complete aircraft turnaround process, detailed airside traffic interactions, operator behaviour, or real-time human decision-making. These simplifications are considered acceptable for the intended use of the model in fleet-sizing and scenario analysis.

Overall, the validation supports the model's fitness for purpose as a decision-support tool for eGSE fleet-sizing analysis in the KLM Schiphol case. The model produces plausible results for baseline service performance, task-demand structure, charging behaviour, service-resource behaviour where applicable, and the effects of the implemented uncertainty mechanisms. The results should therefore be interpreted at the system level and within the defined case-study scope. Within these boundaries, the validation supports the use of the model for experiment comparison and fleet-capacity assessment.

6
Simulation Experiments and Results

This chapter uses the validated KLM Schiphol case-study model to evaluate the effects of fleet size, operational uncertainty, demand conditions, operational-control rules, and parameter assumptions on eGSE service performance. The objective is to translate the simulation framework from Chapter 4 and the case-study validation from Chapter 5 into fleet-sizing outcomes and operational insights.

The experiments are organised as a connected sequence. First, the baseline case is evaluated to establish the reference performance of the current model configuration. Second, deterministic fleet-sizing experiments identify the minimum fleet size required under fixed operating conditions. Third, task-timing uncertainty, travel-time uncertainty, and combined uncertainty are introduced to assess how stochastic operating conditions affect service reliability. Fourth, a demand-case robustness check, an idle forward-staging policy, and a feasibility-aware task-selection variant are tested to examine how demand conditions and operational-control assumptions influence the results. Finally, sensitivity analyses evaluate how selected operational and energy-related parameters affect punctuality, vehicle use, charging behaviour, and infrastructure utilisation.

In this chapter, an uncertainty setting refers to the stochastic operating condition applied to the simulation. A demand case refers to a complete operational input day. An operational-control variant modifies vehicle positioning or dispatching behaviour, while a task-selection variant modifies the ordering of waiting tasks. A sensitivity case changes a single parameter while keeping the baseline model structure unchanged. These distinctions separate external operational variability from internal operational-control choices.

6.1. Experiment Design

The experiment design supports robust fleet-sizing decisions. The main reference case is based on a peak-demand day because fleet capacity for operationally critical eGSE should remain sufficient under demanding workload conditions and operational uncertainty. This provides a conservative basis for estimating the fleet size required to maintain the selected service level. To assess whether this peak-day basis is overly conservative, an additional average-day experiment is performed for water vehicles.

Each experiment group corresponds to a specific part of the research logic. The deterministic baseline experiments establish nominal operational performance under fixed input conditions. The fleet-size experiments examine how service performance changes when vehicle capacity is varied. The uncertainty experiments assess whether deterministic fleet-size conclusions remain valid under stochastic task timing and travel-time variation. The operational-control variants examine whether part of the fleet-capacity requirement is caused by controllable positioning or dispatching rules. The sensitivity experiments evaluate whether the conclusions depend mainly on operational assumptions or on energy- and charging-related parameters. Together, these experiments provide the empirical basis for answering the central RQ in Section 8.2.

The deterministic baseline cases serve as the reference for the analysis. A selected day from the summer peak period is used as the baseline day. This day represents a stress-test case based on operational demand pressure, baseline service performance, and the maximum number of concurrent tasks among days with an equal on-time completion rate. Separate deterministic baseline cases are defined for the three vehicle categories considered in the case study: water vehicles, toilet vehicles, and loaders. The corresponding reference fleet configurations are 11 water vehicles, 11 toilet vehicles, and 25 loaders. Task-timing and travel-time uncertainty are disabled in these cases. This isolates the core operational behaviour of the model and shows whether the assigned fleets can complete the scheduled tasks within their service windows under fixed input conditions.

Fleet-size experiments investigate the relationship between vehicle capacity and operational performance. For each vehicle category, the number of vehicles is varied systematically while all other model settings remain constant. The main performance criterion is the on-time task completion rate because it directly measures whether tasks are completed before the end of their required service windows. Additional indicators include the number of late and unserved tasks, vehicle utilisation, driven distance, and charging behaviour. Together, these indicators show how performance develops across fleet sizes and reveal the marginal benefit of adding an extra vehicle.

The uncertainty experiments assess how operational variability affects system performance. Deterministic runs are compared with stochastic runs in which task-timing uncertainty, travel-time uncertainty, or both uncertainty sources are activated. For each stochastic setting, 30 simulation replications are performed using different random seeds. This reduces dependence on a single stochastic realisation and allows both average performance and variability across runs to be evaluated. The uncertainty experiments show whether fleet-size conclusions from the deterministic analysis remain valid under more realistic operating conditions.

The sensitivity analyses examine how strongly the model outcomes depend on selected operational and energy-related assumptions. Unlike the fleet-size experiments, fleet size is not re-optimised. Instead, the sensitivity analysis is performed around fixed reference configurations of 11 water vehicles, 11 toilet vehicles, and 25 loaders. This isolates the effect of individual parameter changes on task punctuality, operational time allocation, charging demand, and charger utilisation. The sensitivity analysis therefore provides a local robustness assessment around the selected reference configurations rather than a separate fleet-sizing method.

Finally, two exploratory operational variants are included. These variants are not part of the final fleet-sizing procedure but indicate possible directions for improving daily execution. The first variant is idle forward staging, in which idle vehicles remain at their current VOP when their battery state is sufficient and further demand is expected. This examines whether unnecessary depot-centred deadheading can be reduced. The second variant is feasibility-aware triage dispatching, in which waiting tasks that can still meet their service-window deadline are prioritised over tasks that are already infeasible. This evaluates whether scarce vehicle capacity can be allocated more effectively under below-threshold fleet conditions.

Together, the experiment groups provide a structured basis for evaluating fleet requirements and operational robustness. The deterministic and fleet-size experiments show how fleet capacity affects performance under fixed conditions. The uncertainty settings test whether these conclusions remain valid when operational variability is introduced. The sensitivity analyses identify which assumptions have the strongest influence on the model outcomes and whether performance is mainly constrained by operational factors or by energy- and charging-related factors. The exploratory operational variants assess whether changes in staging and dispatching logic can improve performance while remaining separate from the final fleet-size recommendations.

6.2. Parameter Variations

The experiment groups described above are implemented through explicit parameter variations. Unless stated otherwise, all parameters that are not part of a specific experiment remain identical to the deterministic baseline configuration. This ensures that differences between experiments can be attributed to the parameter or rule being tested.

For the fleet-size experiments, the number of vehicles is varied separately for each vehicle category. The tested ranges include both undercapacity and overcapacity conditions. Fleet sizes are evaluated up to the point at which the on-time task completion rate reaches 100.00%, with additional fleet sizes included where needed to assess the marginal value of extra vehicles. A step size of one vehicle is used so that the effect of adding or removing a single vehicle can be observed. The final interpretation combines deterministic performance, stochastic performance, marginal improvement, unserved tasks, and charging behaviour.

For the uncertainty settings, two sources of operational uncertainty are considered: task-timing uncertainty and travel-time uncertainty. The analysis includes a deterministic case in which both sources are disabled. It also includes cases with only task-timing uncertainty, only travel-time uncertainty, and both uncertainty sources combined. For each stochastic setting, 30 simulation replications are performed using different random seeds. This provides insight into both average system performance and the variability between stochastic runs.

For the exploratory operational variants, only one operational rule is changed at a time. In the idle forward-staging variant, the idle-positioning rule is modified while demand, task definitions, service windows, service durations, fleet size, dispatching logic, charging assumptions, and refill or dump rules remain unchanged. In the feasibility-aware triage variant, only the ordering of waiting tasks is modified. Fleet size, demand, service windows, service durations, uncertainty settings, charging assumptions, refill and dump rules, and vehicle execution remain unchanged. These variants are therefore interpreted as limited operational comparisons rather than as additional fleet-sizing experiments.

For the sensitivity analyses, the parameter variations are divided into operational and energy-related sensitivities. All sensitivity cases use the fixed reference fleet configurations of 11 water vehicles, 11 toilet vehicles, and 25 loaders. The evaluated outputs include task punctuality, charging sessions, charger occupancy, peak charger use, and charging queues. The sensitivity results support the interpretation of the fleet-size and uncertainty experiments but do not directly determine the final recommended fleet size.

The operational sensitivities include travel speed and service time. Travel speed is varied by (-20%), (-10%), (0%), (+10%), and (+20%) relative to the baseline value. This examines how assumptions about vehicle movement across the airport affect response times, vehicle availability, and task punctuality. Service time is varied using the same five levels to evaluate the effect of shorter or longer handling durations.

The service-time sensitivity requires careful interpretation. The baseline service times are fixed operational input values provided by KLM and are expected to fit within the corresponding service windows. Increasing service time is therefore not treated as a realistic change in the standard handling process. If the service duration exceeds the available time within the service window, the task cannot be completed on time by definition, regardless of vehicle availability or dispatching performance. The service-time sensitivity should therefore be interpreted as a stress test of the available planning slack and of the vulnerability of the schedule to longer-than-expected handling durations.

The energy-related sensitivities include charger power, energy consumption, and a combined battery bundle. Charger power is varied by comparing the baseline level with increases of (+100%), (+200%), (+300%), and (+400%). This tests whether faster charging improves operational performance or whether charging is not a binding constraint in the reference configuration. Energy consumption is tested by comparing the baseline value with a (+100%) case. This represents a stress test of the charging system under substantially higher energy demand.

The combined battery bundle varies several energy-related assumptions simultaneously. In the negative bundle cases, usable battery capacity and charger power are reduced while energy consumption is increased. In the positive bundle cases, usable battery capacity and charger power are increased while energy consumption is reduced. Five bundle levels are considered: (-20%), (-10%), (0%), (+10%), and (+20%). Within each bundle level, all included parameters are adjusted by the same percentage. For example, the (-20%) case combines a 20% reduction in usable battery capacity and charger power with a 20% increase in energy consumption. This test represents a more adverse or more favourable battery and charging configuration.

Table 6.1 summarises the fleet-size ranges and uncertainty settings used across the experiment groups. For the deterministic fleet-sizing experiments, the tested ranges start at low fleet sizes and extend to 13 vehicles for water and toilet vehicles and 25 vehicles for loaders. The upper bounds are based on the baseline configurations and are extended where necessary to capture the transition from infeasible operation to 100% on-time task completion under fixed conditions. For the stochastic fleet-sizing experiments, the ranges extend to 16 water vehicles, 15 toilet vehicles, and 28 loaders. These wider ranges allow the full performance transition to be observed under combined task-timing and travel-time uncertainty.

The average-day check focuses on the fleet-size range around the transition to 100% on-time performance because its objective is to assess whether the main fleet-sizing conclusions remain consistent under a lower-demand day. The idle forward-staging variant is tested around the higher and most relevant fleet-size ranges, where changes in vehicle positioning can influence the transition towards full feasibility. The feasibility-aware triage variant is mainly tested at lower fleet sizes because this dispatching logic is not expected to reduce the first fleet size at which 100% on-time completion is reached. Its value is instead assessed below the strict feasibility threshold, where prioritisation decisions have the greatest effect on lateness and the allocation of scarce vehicle capacity.

Table 6.1: *Overview of experiment settings*

Experiment	Vehicle type			Uncertainty		Remarks
	Water	Toilet	Loader	Task	Travel	
Baseline	11	11	25			Reference fleet configuration.
Deterministic fleet sizing	1–13	1–13	1–25			Full fleet-size curves; detailed tables report selected near-threshold values.
Uncertainty impact	11	11	25	✓	✓	Tested at the baseline fleet configuration.
Stochastic fleet sizing	1–16	1–15	1–28	✓	✓	Fleet sizing under combined task-timing and travel-time uncertainty.
Average-day check	D: 6–11 S: 10–13			✓	✓	Water-only robustness check on an average-demand day.
Idle forward-staging	D: 10–12 S: 10–14	D: 9–11 S: 7–10	D: 22–23 S: 22–25	✓	✓	Operational positioning variant; deterministic and combined stochastic settings tested.
Feasibility-aware triage	4–11	3–10	15–22			Deterministic task-selection variant.
Sensitivity	11	11	25			Travel speed: −20%, −10%, 0%, +10%, +20% Service time: −20%, −10%, 0%, +10%, +20% Charger power: 0%, +100%, +200%, +300%, +400% Energy consumption: 0%, +100% Battery bundle: −20%, −10%, 0%, +10%, +20%

Note. D = deterministic and S = stochastic. Task refers to task-timing uncertainty and travel refers to travel-time uncertainty. Deterministic settings are simulated once per configuration, while stochastic settings are simulated with 30 replications per configuration.

6.3. Results

This section presents the simulation results for the experiment groups defined in Section 6.1. The results are reported for the three case-study vehicle groups: water vehicles, toilet vehicles, and loaders. The analysis examines how fleet size, operational uncertainty, charging behaviour, and service-specific resource constraints affect system performance in the KLM Schiphol case.

The results are organised into nine parts. First, the fleet-size performance criteria are defined to explain how the required fleet size is determined. Second, the deterministic baseline case is analysed to establish the reference performance of the current fleet configurations. Third, deterministic fleet-sizing experiments identify the minimum fleet size required under fixed operating conditions. Fourth, the effects of task-timing uncertainty, travel-time uncertainty, and combined uncertainty are evaluated. Fifth, fleet sizing is repeated under combined uncertainty to assess the additional capacity required for robust performance.

The sixth and seventh parts evaluate two exploratory operational variants. These variants examine whether idle vehicle positioning and task-ordering logic can improve operational performance without changing demand, service windows, service durations, fleet size, charging assumptions, or service-resource rules. They are not used to determine the final recommended fleet sizes but illustrate possible directions for operational improvement. Eighth, sensitivity analyses assess how selected operational and charging parameters influence service reliability, vehicle use, and charging demand. Finally, the result synthesis combines the main findings and explains how the fleet-sizing outcomes should be interpreted.

The results are evaluated using the performance indicators introduced in Section 5.5. Particular attention is given to the on-time task completion rate, unserved tasks, lateness severity, vehicle utilisation, travel distance, charging sessions, charger occupancy, peak charger use, and the minimum required fleet size. For water and toilet vehicles, service-specific resource behaviour, including filling and dumping, is also considered where relevant.

6.3.1. Fleet Size Performance Criteria

The fleet-sizing results are evaluated primarily using the on-time task completion rate. This metric represents the percentage of tasks completed before the end of their required service window. It is used as the main performance criterion because it indicates whether the available fleet can support the operational schedule without causing service-window violations. A fleet configuration that completes every task but completes some tasks after their deadlines does not fully satisfy the strict operational requirement.

A distinction must be made between internal service-window performance and aircraft departure punctuality. The service windows used in this study represent operational task windows within the turnaround process. A task completed shortly after the end of its internal service window does not necessarily delay the aircraft if sufficient turnaround slack remains. The on-time task completion rate should therefore be interpreted as a system-level reliability indicator for GSE operations rather than as a direct prediction of D0 performance. Consequently, a strict requirement of 100% on-time task completion is conservative.

In principle, the required operational fleet size is defined as the smallest fleet that achieves 100% on-time task completion. However, the final fleet-size decision should not be based on this metric alone. The marginal benefit of adding one vehicle should also be considered. When a smaller fleet already achieves an on-time completion rate close to 100%, an additional vehicle may only remove a small number of short service-window violations. In such cases, the number of late tasks, their average lateness, and the operational criticality of the vehicle group should also be considered.

This is particularly relevant because the case study is based on the highest-demand week of the year and the busiest day within that week. The selected day can therefore be interpreted as a stress-test day rather than as an average operating day. Sizing the fleet only for this single day may not be efficient when an additional vehicle provides only a small improvement in performance. A limited number of short internal service-window violations may therefore be operationally acceptable, provided that they are rare and do not affect turnaround completion.

The acceptable performance level may also differ between vehicle groups because not all GSE tasks have the same operational criticality. Water and toilet tasks are generally required until approximately 10 minutes before departure and therefore usually retain more turnaround buffer. Small internal service-window violations may be acceptable when they are infrequent, short, and do not delay the turnaround. Loader tasks, by contrast, may continue until approximately 2 minutes before departure and are more directly linked to baggage and cargo handling. Loader delays are therefore interpreted more strictly, and near-threshold loader fleet sizes should be evaluated more cautiously than comparable water or toilet fleet sizes. For all vehicle groups, unserved tasks or repeated large delays indicate insufficient operational capacity.

The fleet-size results are therefore interpreted from two perspectives. The strict fleet size is the smallest fleet that achieves 100% on-time task completion. The pragmatic lower-bound fleet size is a smaller fleet that achieves a very high service level but still results in a small number of short internal service-window violations. These lower-bound fleets are not equivalent to the strict fleet sizes, but they help evaluate the trade-off between operational reliability and fleet efficiency.

For stochastic experiments, replication-level on-time completion rates are used to calculate sample means, standard deviations, and 95% confidence intervals. These confidence intervals indicate the precision of the estimated mean on-time completion rate. Throughout this chapter, the terms 'strict robust threshold' and 'stochastic strict' refer to the smallest tested fleet size that achieves 100% on-time completion in all 30 sampled replications for the relevant uncertainty setting. This definition does not imply a formal probabilistic guarantee across all possible operational realisations, unmodelled disruptions, or operating days outside the tested experiment set.

Supporting indicators are used to interpret the fleet-size outcomes. Late tasks are distinguished from unserved tasks. Late tasks are completed after the end of their service window, whereas unserved tasks are not completed within the applicable operational horizon. For water and toilet vehicles, this horizon is limited by the stop-cutoff logic associated with operator service periods. For loaders, the horizon is the 24-hour simulation period. Other indicators include the average lateness of late tasks, vehicle utilisation, charging queues, charger occupancy, and idle time. Together, these indicators help determine whether performance losses are caused by a structural shortage of vehicle capacity, local demand peaks, charging constraints, or incidental operating conditions.

After determining the operational fleet size, a technical reserve should be added separately. This reserve accounts for planned and preventive maintenance, unexpected failures, battery degradation, charging problems, and other causes of temporary vehicle unavailability. A common rule of thumb is to add a reserve of approximately 10% of the operational fleet, rounded upward. However, this rule should not be applied mechanically. The required reserve depends on the criticality of the operation, the size of the fleet, and the availability of backup vehicles. For small or operationally critical fleets, at least one reserve vehicle should be included, while a second reserve vehicle may be justified in some cases. Determining the technical reserve falls outside the active simulation scope of this study and should be performed after the operational fleet requirement has been established.

6.3.2. Baseline Case

This section presents the results for the baseline case for the water, toilet, and loader operations. The baseline case is selected as the most operationally constrained day in the dataset. First, the day with the lowest number of tasks completed on time is identified. If multiple days have the same number of tasks completed on time, the tie is resolved by selecting the day with the highest number of concurrent tasks. This day is used as the baseline because it represents the strongest bottleneck in the operation and therefore provides a suitable reference case for evaluating the performance of the eGSE system. Figure 6.1 shows the total number of arrivals and departures on 14 June. This provides an overview of all flights on the selected day. However, not every flight requires service from each vehicle type. The required services differ for water vehicles, toilet vehicles, and loaders. Therefore, the vehicle-specific sections only show the arrivals and departures that require service from that specific vehicle type.

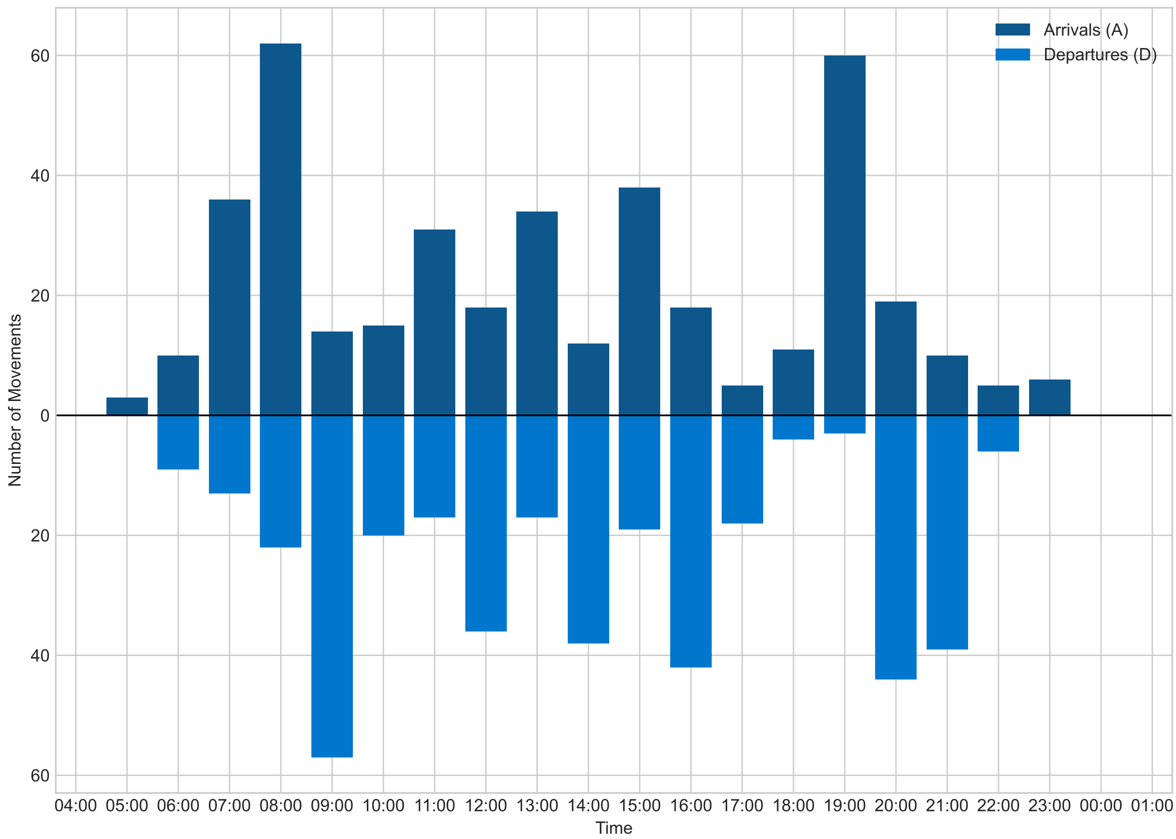


Figure 6.1: *Arrivals vs Departures on 14-06-2026 for KLM GS*

Water Vehicles

For the water operation, day 14-06-2026 is selected as the baseline day. On this day, 97.7% of the water tasks are completed on time. The baseline fleet consists of 11 water vehicles. Each vehicle has a battery capacity of 70 kWh, a driving speed of 18 km/h, and can charge with a maximum charging power of 22 kW. These settings are used to evaluate whether the electric water fleet can complete the required operational demand under the most constraining observed conditions.

Figure 6.2 shows the operational structure of the selected baseline day, including the distribution of arrival tasks, departure tasks, and turnaround tasks. This figure provides the operational context for the water results by showing how the task demand is distributed over the day and when workload peaks occur. The Table 5.4 shows the performance of this case. The baseline task-execution pattern for water vehicles is illustrated in Figure 6.3.

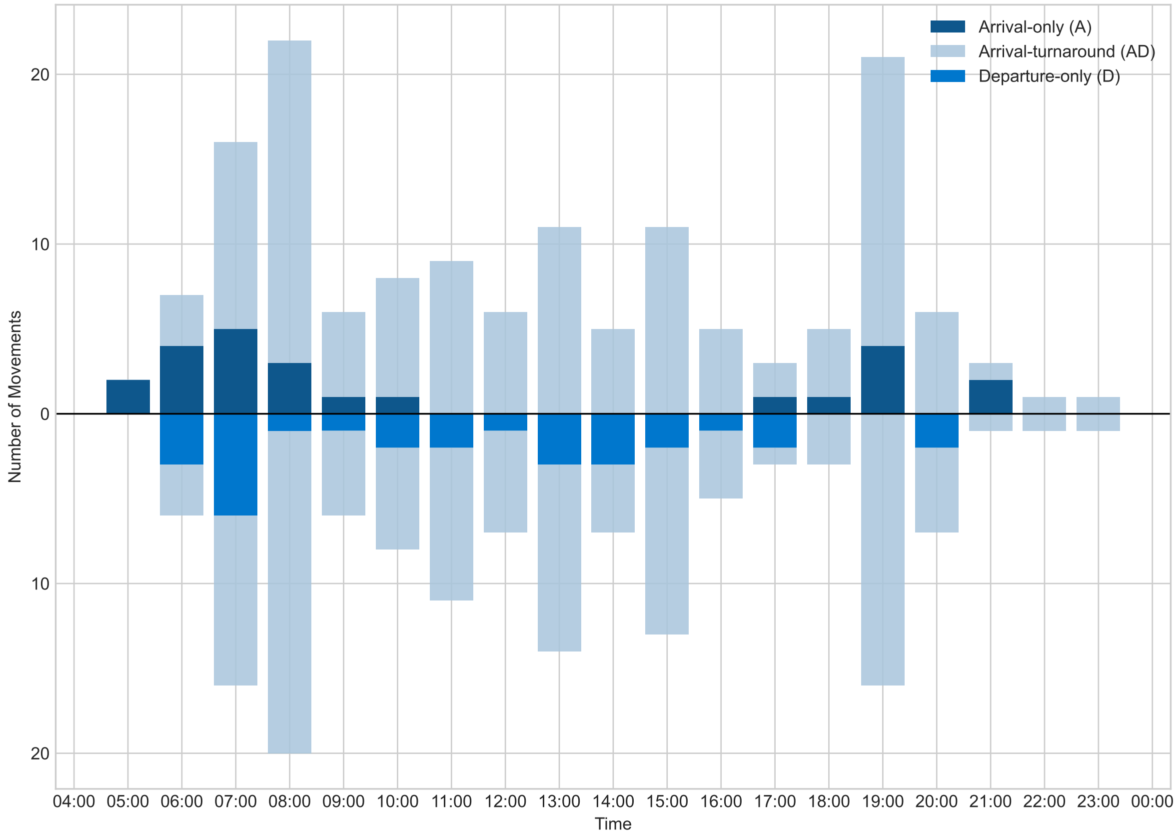


Figure 6.2: *Arrivals vs Departures which need service from a water vehicle on 14-06-2026*

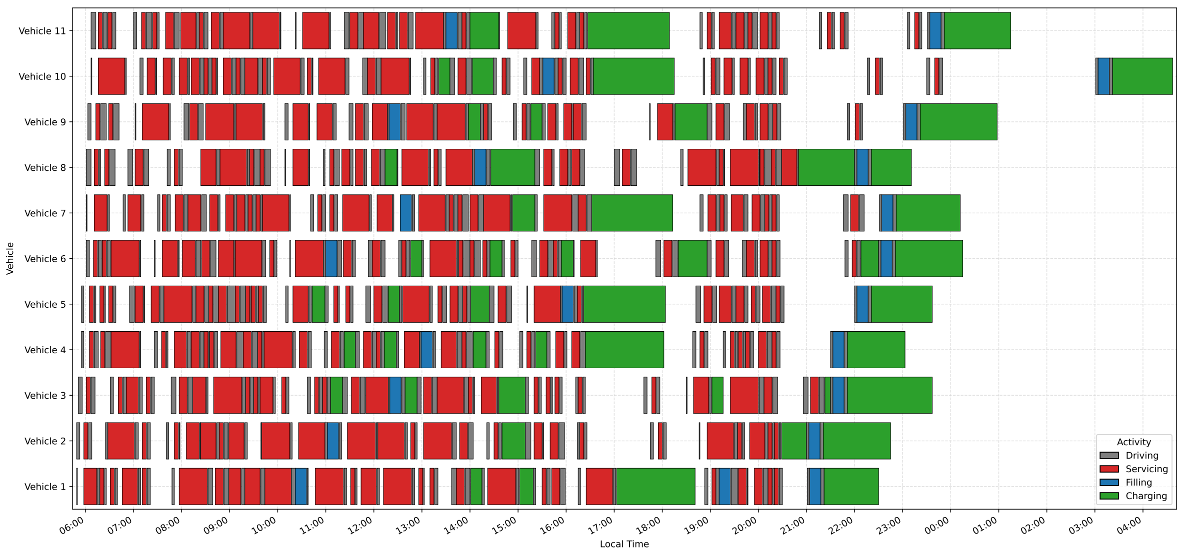


Figure 6.3: *Gantt chart of the baseline of water vehicles*

Toilet Vehicles

For the toilet operation, day 14-06-2026 is selected as the baseline day. On this day, 100% of the toilet tasks are completed on time. This day also contains the maximum number of concurrent toilet tasks, with 11 tasks occurring at the same time. The baseline fleet consists of 11 toilet vehicles. Each vehicle has a battery capacity of 70 kWh, a driving speed of 18 km/h, and can charge with a maxi-

mum charging power of 22 kW. These settings are used to evaluate whether the electric toilet fleet can complete the required operational demand under the most constraining observed conditions.
Figure 6.4 shows the operational structure of the selected baseline day, including the distribution of arrival tasks, departure tasks, and turnaround tasks. This figure provides the operational context for the toilet results by showing how the task demand is distributed over the day and when workload peaks occur. Table 5.5 shows the performance of this case. The baseline task-execution pattern for toilet vehicles is illustrated in Figure 6.5.

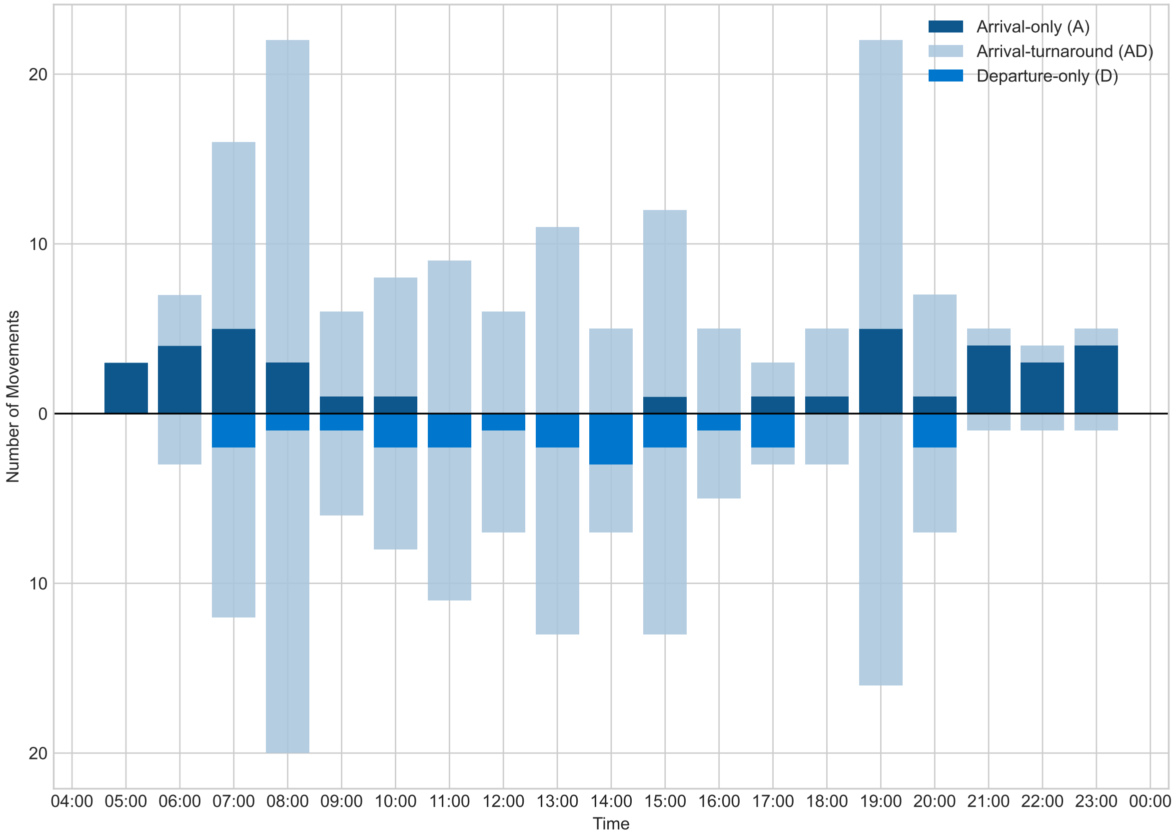


Figure 6.4: *Arrivals vs Departures which need service from a toilet vehicle on 14-06-2026*

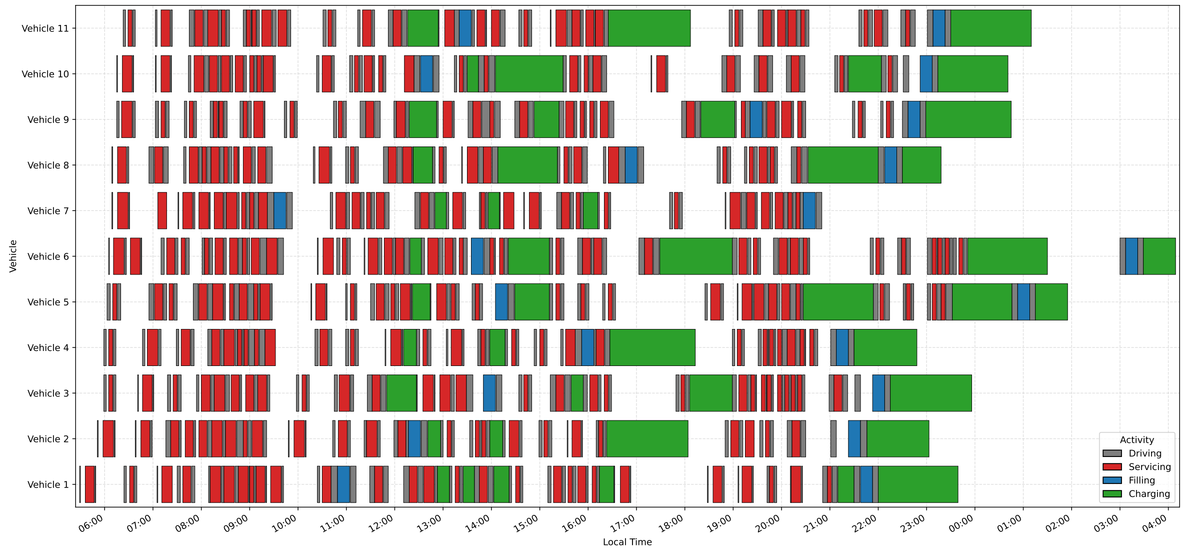


Figure 6.5: *Gantt chart of the baseline of toilet vehicles*

Loaders

For the loader operation, day 14-06-2026 is selected as the baseline day. On this day, 100% of the loader tasks are completed on time. This day also contains the maximum number of concurrent loader tasks, with 21 tasks occurring at the same time. The baseline fleet consists of 25 loader vehicles. Each vehicle has a battery capacity of 80 kWh, a driving speed of 11 km/h, and can charge with a maximum charging power of 16 kW. These settings are used to evaluate whether the electric loader fleet can complete the required operational demand under the most constraining observed conditions.

Figure 6.6 shows the operational structure of the selected baseline day, including the distribution of arrival tasks, departure tasks, and turnaround tasks. This figure provides the operational context for the loader results by showing how the task demand is distributed over the day and when workload peaks occur. Table 5.6 shows the performance of this case. The baseline task-execution pattern for loaders is illustrated in Figure 6.7.

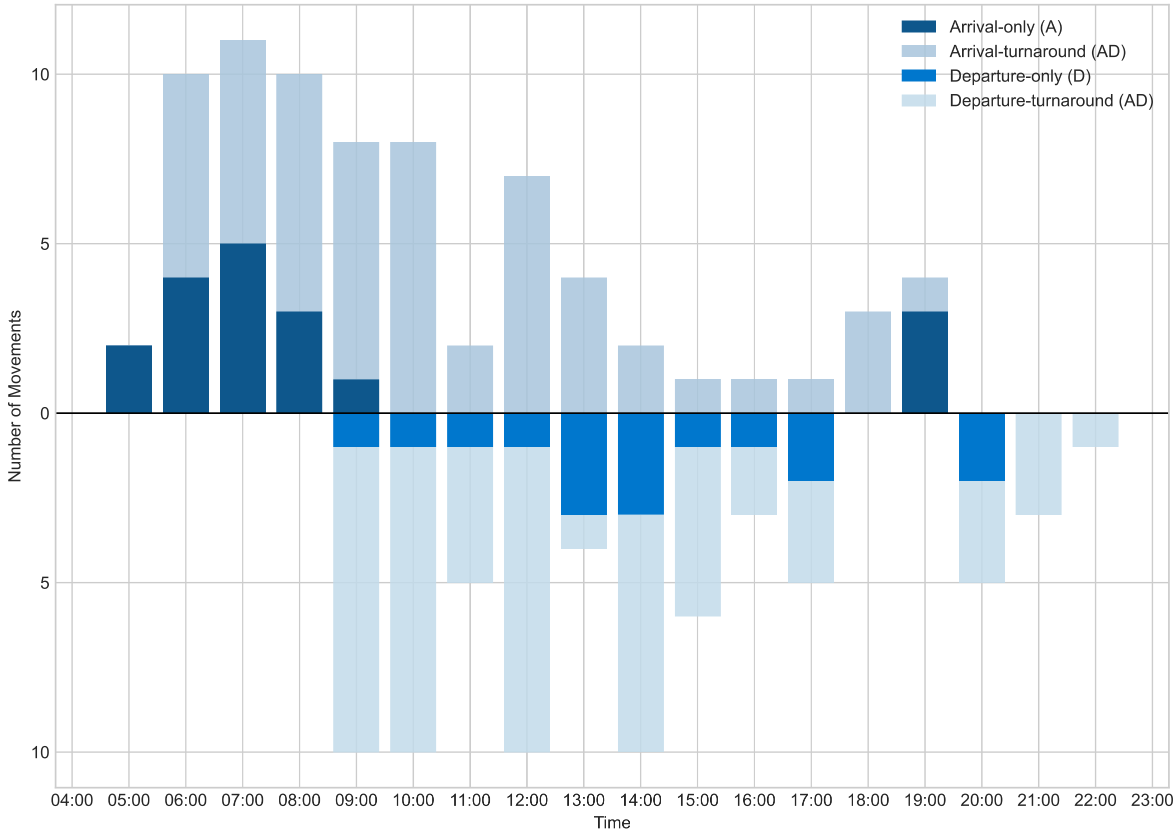


Figure 6.6: *Arrivals vs Departures which need service from a loader on 14-06-2026*

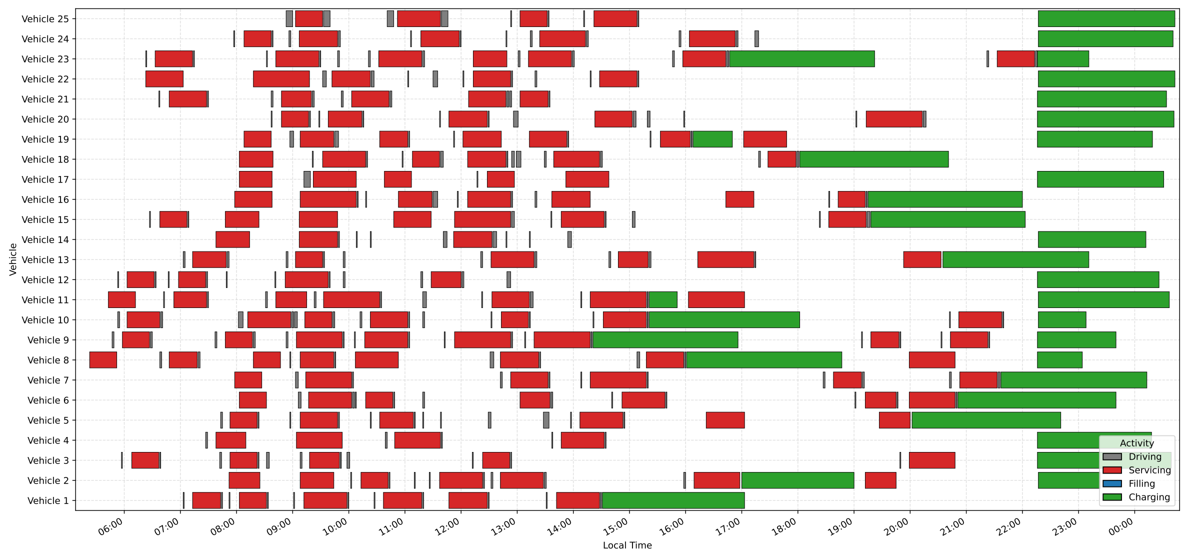


Figure 6.7: *Gantt chart of the baseline of loader vehicles*

6.3.3. Deterministic Fleet Sizing Results

The deterministic fleet-sizing experiments evaluate the minimum number of vehicles required under nominal operating conditions. All stochastic uncertainty sources are disabled, meaning that task-timing uncertainty and travel-time uncertainty are not included. For each vehicle group, fleet size is varied and performance is assessed using completed tasks, unserved tasks, on-time task completion, late tasks, average lateness, and total distance. A fleet is considered sufficient under the strict deterministic criterion when all tasks are completed and 100% of tasks are completed within their service windows. Across the three vehicle groups, the same interpretation applies: small fleets may eventually complete most or all tasks, but the operational bottleneck is whether tasks can be completed on time during demand peaks.

The deterministic fleet-sizing results should not be interpreted as the final recommended fleet sizes under real operating conditions, because uncertainty and robustness are evaluated separately in the subsequent experiments.

Water Vehicles

Table 6.2 and Figure 6.8 show the deterministic fleet-sizing results for water vehicles. Fleet sizes below seven vehicles are unable to complete all 304 active tasks. With six vehicles, 298 tasks are completed and 1.97% of tasks remain unserved. From seven vehicles onwards, all tasks are completed, but this does not mean that the service level is sufficient. With seven vehicles, only 42.11% of tasks are completed within their service windows, showing that the system is still strongly time-constrained during demand peaks.

Table 6.2: *Fleet-sizing results for the deterministic water-service experiment*

Vehicles	Completed tasks	Unserved rate (%)	On-time rate (%)	Late tasks	Avg. lateness (min)	Distance (km)
6	298/304	1.97	18.09	243	81.08	381.0
7	304/304	0.00	42.11	176	20.38	340.9
8	304/304	0.00	80.26	60	14.58	362.0
9	304/304	0.00	94.08	18	11.17	378.0
10	304/304	0.00	95.07	15	8.07	399.1
11	304/304	0.00	97.70	7	2.86	425.4
12	304/304	0.00	100.00	0	0.00	441.6
13	304/304	0.00	100.00	0	0.00	457.8

The strongest performance improvement occurs when the fleet increases from seven to eight vehicles. The on-time completion rate rises from 42.11% with seven vehicles to 80.26% with eight vehicles and 94.08% with nine vehicles. Beyond this point, the improvement becomes smaller. Increasing the fleet from nine to ten vehicles only increases on-time completion from 94.08% to 95.07%. With 11 vehicles, the system reaches 97.70% on-time completion, with seven late tasks and an average lateness of 2.86 minutes. This indicates that 11 vehicles are close to the strict deterministic service requirement, although they do not formally meet the 100% criterion.

The first fleet size that completes all 304 tasks within their service windows is 12 vehicles. This fleet size therefore represents the minimum deterministic fleet size under the strict performance criterion used in this study. Adding a thirteenth vehicle does not improve on-time completion, because the service level is already 100.00% with 12 vehicles. The additional vehicle mainly increases operational slack and total driven distance, from 441.6 km to 457.8 km, without providing a measurable service-performance benefit in the deterministic experiment.

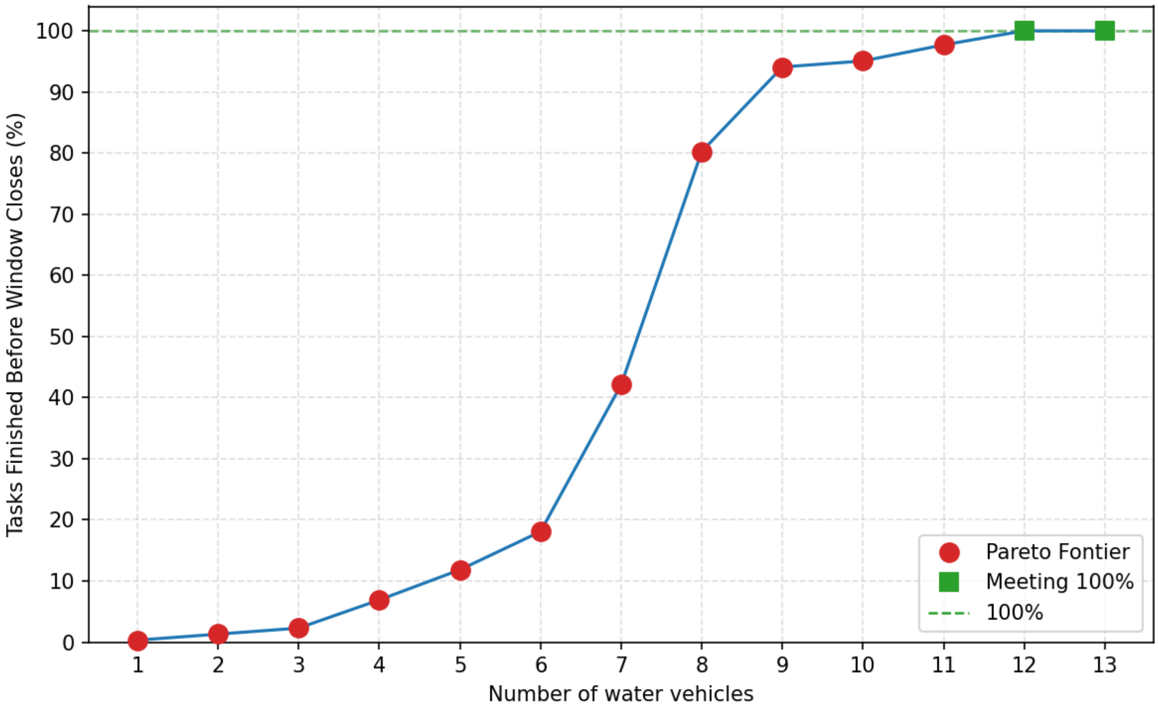


Figure 6.8: *Running different amounts of vehicles*

Toilet Vehicles

Table 6.3 and Figure 6.9 show the deterministic fleet-sizing results for toilet vehicles. Fleet sizes below six vehicles are unable to complete all 301 active tasks. With five vehicles, 265 tasks are completed and 11.96% of tasks remain unserved. From six vehicles onwards, all tasks are completed, but this does not mean that the service level is sufficient. With six vehicles, 81.40% of tasks are completed within their service windows, showing that the system is still time-constrained during peak demand periods.

Table 6.3: *Fleet-sizing results for the deterministic toilet-service experiment*

Vehicles	Completed tasks	Unserved rate (%)	On-time rate (%)	Late tasks	Avg. lateness (min)	Distance (km)
5	265/301	11.96	17.28	213	43.23	389.4
6	301/301	0.00	81.40	56	11.48	383.7
7	301/301	0.00	96.68	10	7.60	415.7
8	301/301	0.00	99.34	2	2.00	423.7
9	301/301	0.00	99.34	2	2.50	451.9
10	301/301	0.00	99.67	1	1.00	466.6
11	301/301	0.00	100.00	0	0.00	489.3
12	301/301	0.00	100.00	0	0.00	515.8
13	301/301	0.00	100.00	0	0.00	535.5

The strongest performance improvement occurs when the fleet increases from five to six vehicles. The on-time completion rate rises from 81.40% with six vehicles to 96.68% with seven vehicles and 99.34% with eight vehicles. Beyond this point, the improvement becomes small. Increasing the fleet from eight to nine vehicles does not improve the on-time completion rate, which remains 99.34%. With ten vehicles, the system reaches 99.67% on-time completion, with only one late task and an average lateness of 1.00 minute. This indicates that 10 vehicles are very close to the strict deterministic service requirement, although they do not formally meet the 100% criterion.

The first fleet size that completes all 301 tasks within their service windows is 11 vehicles. This fleet size therefore represents the minimum deterministic fleet size under the strict performance criterion used in this study. Adding a twelfth or thirteenth vehicle does not improve on-time completion, because the service level is already 100.00% with 11 vehicles. The additional vehicles mainly increase operational slack and total driven distance, from 489.3 km with 11 vehicles to 515.8 km with 12 vehicles and 535.5 km with 13 vehicles, without providing a measurable service-performance benefit in the deterministic experiment.

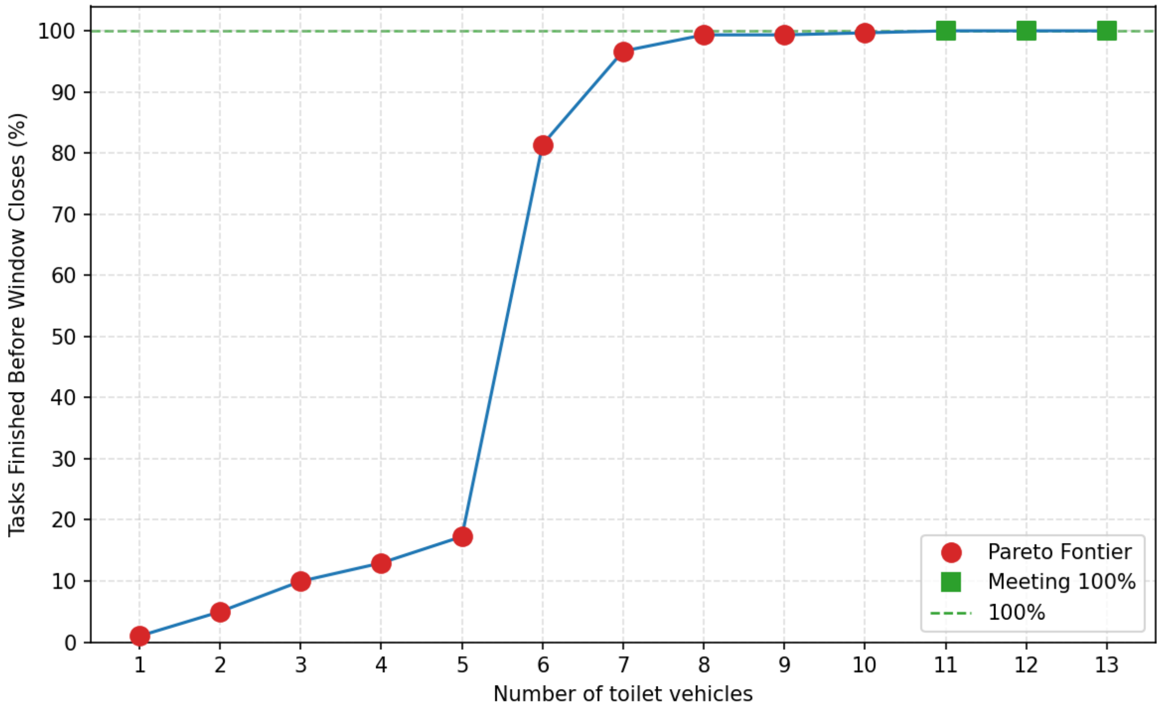


Figure 6.9: *Running different amounts of vehicles*

Loaders

Table 6.4 and Figure 6.10 show the deterministic fleet-sizing results for loaders. From six vehicles onwards, all 147 active tasks are completed. However, this does not mean that the service level is sufficient. With six vehicles, only 12.24% of tasks are completed within their service windows, with 129 late tasks and an average lateness of 248.37 minutes. This shows that the loader operation is mainly constrained by time-window feasibility rather than by the ability to eventually complete the total task volume.

Table 6.4: *Fleet-sizing results for the deterministic loader experiment*

Vehicles	Completed tasks	Unserved rate (%)	On-time rate (%)	Late tasks	Avg. lateness (min)	Distance (km)
6	147/147	0.00	12.24	129	248.37	101.2
10	147/147	0.00	37.41	92	36.55	88.4
15	147/147	0.00	89.12	16	15.69	71.8
18	147/147	0.00	93.20	10	10.00	62.3
20	147/147	0.00	94.56	8	3.50	55.3
21	147/147	0.00	97.96	3	3.33	54.0
22	147/147	0.00	98.64	2	3.00	54.9
23	147/147	0.00	100.00	0	0.00	58.6
24	147/147	0.00	100.00	0	0.00	57.3
25	147/147	0.00	100.00	0	0.00	58.3

The strongest performance improvement occurs when the fleet increases from 10 to 15 vehicles. The on-time completion rate rises from 37.41% with 10 vehicles to 89.12% with 15 vehicles. Further increases continue to improve service-window compliance, but the marginal improvement becomes smaller. With 20 vehicles, 94.56% of tasks are completed on time. With 21 vehicles, this increases to 97.96%, leaving three late tasks with an average lateness of 3.33 minutes. With 22 vehicles, the system reaches 98.64% on-time completion, leaving only two late tasks with an average lateness of 3.00 minutes. This indicates that 22 vehicles are close to the strict deterministic service requirement, although they do not formally meet the 100% criterion.

The first fleet size that completes all 147 tasks within their service windows is 23 vehicles. This fleet size therefore represents the minimum deterministic fleet size under the strict performance criterion used in this study. Adding a twenty-fourth or twenty-fifth vehicle does not improve on-time completion, because the service level is already 100.00% with 23 vehicles. The small changes in total driven distance for larger fleets reflect changes in task assignment and routing rather than a measurable improvement in service performance.

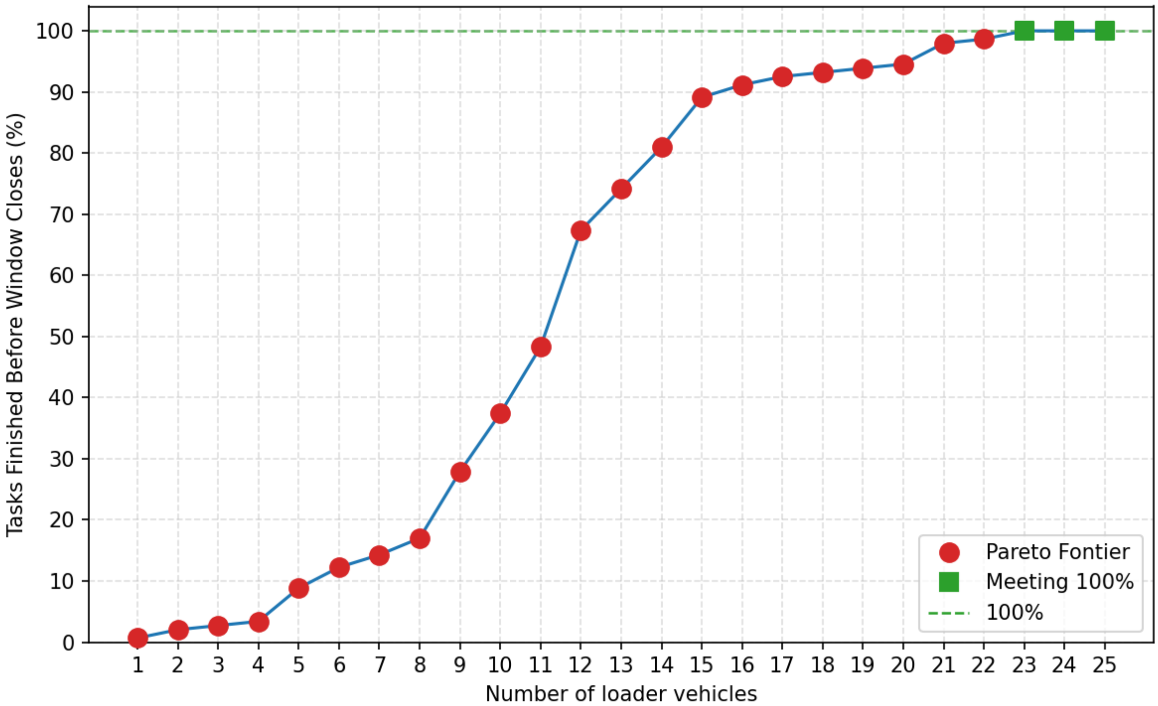


Figure 6.10: *Running different amounts of vehicles*

6.3.4. Impact of Operational Uncertainty at Baseline Case

This section evaluates how operational uncertainty affects the baseline configuration of each vehicle group. For water vehicles, toilet vehicles, and loaders, the deterministic baseline is compared with three stochastic settings: task-timing uncertainty only, travel-time uncertainty only, and combined task-timing and travel-time uncertainty. Each stochastic setting is evaluated using 30 simulation runs.

Task-timing uncertainty changes the temporal distribution of demand by shifting task times. In this model, the corresponding service windows are shifted together with the task timing. Therefore, task-timing uncertainty mainly evaluates whether the fleet can handle a changed demand pattern, rather than whether tasks are still completed within the original deterministic schedule. Depending on how the sampled shifts affect peak overlap, this can either improve or worsen measured service-window performance. Travel-time uncertainty has a different effect: task timing and service windows remain fixed, while selected vehicle movements take longer during simulation execution. This directly reduces the available operational margin between tasks and may delay subsequent vehicle availability.

Water Vehicles

For water vehicles, the deterministic baseline consists of 11 vehicles operating on day 14-06. In this baseline, all 304 water-service tasks are completed and no tasks remain unserved. However, 7 tasks are completed after their service-window close, resulting in an on-time task completion rate of 97.70%. The average lateness of these late tasks is 2.86 minutes. Baseline vehicle utilisation is 45.74%, and the total driven distance is 425.4 km.

Table 6.5 compares the deterministic baseline with the mean results of the three uncertainty experiments. In all stochastic experiments, all 304 tasks are completed and no unserved tasks occur. This indicates that the baseline fleet has sufficient capacity to complete the full daily water-service demand under the tested uncertainty settings. The main effect of uncertainty is therefore observed in time-window compliance, vehicle utilisation, and total driven distance.

Table 6.5: *Comparison of the deterministic baseline and uncertainty settings for the water operation*

Experiment	Completed tasks	On-time rate (%)	Unserved tasks	Late tasks	Avg. lateness (min)	Utilisation (%)	Distance (km)
Baseline	304	97.70	0.00	7.00	2.86	45.74	425.4
Task uncertainty	304	99.91	0.00	0.27	1.80	46.62	434.0
Travel uncertainty	304	97.63	0.00	7.20	3.76	45.81	421.7
Combined uncertainty	304	99.91	0.00	0.27	0.50	46.74	434.3

Under task-timing uncertainty, the mean on-time task completion rate increases from 97.70% in the deterministic baseline to 99.91%. The average number of late tasks decreases from 7.00 to 0.27 tasks per run. Across the 30 runs, the on-time completion rate ranges from 99.34% to 100.00%, and 24 runs achieve full on-time completion. This improvement should not be interpreted as uncertainty being generally beneficial. Instead, it indicates that the sampled task shifts reduce some of the peak overlap present in the deterministic baseline, while the service windows move with the shifted tasks.

Travel-time uncertainty has a small negative effect. Because task timing and service windows remain fixed, longer travel times directly reduce the available margin between tasks. The mean on-time task completion rate decreases slightly from 97.70% to 97.63%. The average number of late tasks increases from 7.00 to 7.20, and the average lateness of late tasks increases from 2.86 to 3.76 minutes. Across the 30 runs, no run achieves full on-time completion, with on-time rates ranging from 96.71% to 98.03%. Travel-time uncertainty therefore does not create unserved tasks, but it slightly worsens time-window compliance.

The combined uncertainty setting produces an average on-time completion rate of 99.91%, with 0.27 late tasks per run. This is almost identical to the task-timing uncertainty setting and higher than the deterministic baseline. Across the 30 runs, the on-time completion rate ranges from 99.34% to 100.00%, and 23 runs achieve full on-time completion. This indicates that, under the tested settings, the shifted task pattern dominates the measured time-window performance, while the additional travel-time uncertainty does not reduce punctuality significantly.

Vehicle activity changes only slightly under uncertainty. Utilisation rises from 45.74% in the baseline to 46.62% under task-timing uncertainty and 46.74% under combined uncertainty. Total driven distance increases from 425.4 km in the baseline to 434.0 km and 434.3 km, respectively. By contrast, travel-time uncertainty alone has only a minor effect on utilisation and results in a slightly lower total distance of 421.7 km. This is consistent with the model logic: travel-time uncertainty changes travel duration rather than physical distance, while task-timing uncertainty can change assignment timing and routing patterns.

Overall, the 11-vehicle water fleet remains able to complete the full daily task demand under the tested stochastic settings. However, it does not satisfy the strict 100% time-window criterion in the deterministic baseline, and travel-time uncertainty alone slightly worsens punctuality. The results show that measured water-service performance is sensitive to local peak overlap and to the modelling choice that task-timing uncertainty shifts both task timing and service windows.

Toilet Vehicles

For toilet vehicles, the deterministic baseline consists of 11 vehicles operating on day 14-06. In this baseline, all 301 toilet-service tasks are completed and no tasks remain unserved. All tasks are completed within their service windows, resulting in an on-time task completion rate of 100.00%. Baseline vehicle utilisation is 36.96%, and the total driven distance is 489.3 km.

Table 6.6 compares the deterministic baseline with the mean results of the three uncertainty experiments. In all stochastic experiments, all 301 tasks are completed and no unserved tasks occur. This indicates that the baseline fleet has sufficient capacity to complete the full daily toilet-service demand under the tested uncertainty settings. The main effect of uncertainty is therefore observed in time-window compliance, vehicle utilisation, and total driven distance.

Table 6.6: *Comparison of the deterministic baseline and uncertainty settings for the toilet operation*

Experiment	Completed tasks	On-time rate (%)	Unserved tasks	Late tasks	Avg. lateness (min)	Utilisation (%)	Distance (km)
Baseline	301	100.00	0.00	0.00	0.00	36.96	489.3
Task uncertainty	301	99.96	0.00	0.13	1.07	37.77	508.8
Travel uncertainty	301	99.99	0.00	0.03	0.03	37.59	501.0
Combined uncertainty	301	99.88	0.00	0.37	0.93	38.02	506.2

The toilet operation remains highly stable under the tested stochastic settings. Task-timing uncertainty reduces the mean on-time completion rate only slightly, from 100.00% to 99.96%, with an average of 0.13 late tasks per run. Across the 30 runs, 26 runs still achieve full on-time completion, and the worst run contains only one late task. This shows that the shifted task pattern has a limited effect on the baseline toilet configuration.

Travel-time uncertainty has an even smaller effect. The mean on-time completion rate decreases marginally to 99.99%, with an average of 0.03 late tasks per run. In 29 out of 30 runs, all tasks are completed within their service windows. This indicates that longer travel times slightly reduce operational slack, but do not create a meaningful capacity bottleneck for the tested 11-vehicle fleet.

The combined uncertainty setting produces the largest reduction in punctuality, but the effect remains small. The mean on-time completion rate is 99.88%, with an average of 0.37 late tasks per run. Across the 30 runs, 20 runs achieve full on-time completion, and the worst run contains two late tasks. This suggests that unfavourable combinations of task timing and travel-time variation can create small local bottlenecks, but these bottlenecks remain limited under the tested settings.

Vehicle activity increases slightly when uncertainty is included. Utilisation rises from 36.96% in the baseline to 37.77% under task-timing uncertainty, 37.59% under travel-time uncertainty, and 38.02% under combined uncertainty. Total driven distance also increases in all stochastic experiments, from 489.3 km in the baseline to between 501.0 km and 508.8 km. This suggests that uncertainty changes the dispatching pattern and leads to slightly more vehicle movement across the airport.

Overall, the 11-vehicle toilet fleet remains sufficient to complete the full daily task demand under the tested stochastic settings. The deterministic baseline achieves 100.00% on-time completion, and the stochastic experiments only introduce a very small number of late tasks. Compared with the water operation, the toilet baseline contains more operational slack, making it less sensitive to the tested uncertainty settings.

Loaders

For loaders, the deterministic baseline consists of 25 vehicles operating on day 14-06. In this baseline, all 147 loader tasks are completed and no tasks remain unserved. All tasks are completed within their service windows, resulting in an on-time task completion rate of 100.00%. Baseline vehicle utilisation is 19.49%, and the total driven distance is 58.3 km.

Table 6.7 compares the deterministic baseline with the mean results of the three uncertainty experiments. In all stochastic experiments, all 147 tasks are completed and no unserved tasks occur. This indicates that the baseline fleet has sufficient capacity to complete the full daily loader demand under the tested uncertainty settings. The main effect of uncertainty is therefore limited to small changes in time-window compliance, vehicle utilisation, and total driven distance.

Table 6.7: *Comparison of the deterministic baseline and uncertainty settings for the loader operation*

Experiment	**Completed tasks**	**On-time rate (%)**	**Unserved tasks**	**Late tasks**	**Avg. lateness (min)**	**Utilisation (%)**	**Distance (km)**
Baseline	147	100.00	0.00	0.00	0.00	19.49	58.3
Task uncertainty	147	99.91	0.00	0.13	0.50	19.45	56.1
Travel uncertainty	147	100.00	0.00	0.00	0.00	19.52	58.3
Combined uncertainty	147	99.91	0.00	0.13	0.50	19.46	55.6

The loader operation remains highly stable under the tested stochastic settings. Task-timing uncertainty reduces the mean on-time completion rate only slightly, from 100.00% to 99.91%, with an average of 0.13 late tasks per run. Across the 30 runs, 26 runs still achieve full on-time completion. In the affected runs, only one task is completed late, with a maximum reported lateness of 7 minutes. This shows that the shifted task pattern can create small local bottlenecks, but does not create a structural service problem for the tested loader configuration.

Travel-time uncertainty has no measurable effect on time-window compliance in this configuration. The mean on-time completion rate remains 100.00%, and no late or unserved tasks occur in any of the 30 runs. This indicates that the 25-loader baseline contains sufficient travel-time margin for the tested travel uncertainty assumptions.

The combined uncertainty setting remains close to full time-window compliance. The average on-time completion rate is 99.91%, with 0.13 late tasks per run. Across the 30 runs, 26 runs achieve full on-time completion, and the worst runs contain one late task. The combined experiment therefore produces almost the same service-level outcome as the task-timing uncertainty setting, indicating that travel-time uncertainty does not amplify the small punctuality losses caused by shifted task timing significantly.

Vehicle activity changes only slightly across the uncertainty settings. Utilisation remains close to the deterministic baseline of 19.49%, with mean values between 19.45% and 19.52%. Total driven distance is also similar, although it is slightly lower when task-timing uncertainty is included. These differences mainly reflect changes in task assignment and dispatching rather than a substantial change in operational effort.

Overall, the 25-loader fleet remains sufficient to complete the full daily task demand under the tested stochastic settings. No unserved tasks occur, and the on-time completion rate remains at or very close to 100.00%. Compared with the water and toilet operations, the loader baseline shows the smallest effect of uncertainty, mainly because the tested 25-vehicle configuration already contains sufficient operational slack for the selected baseline day.

6.3.5. Fleet Sizing under Operational Uncertainty

The stochastic fleet-sizing experiments repeat the fleet-size analysis under combined task-timing and travel-time uncertainty. For each tested fleet size, 30 simulation runs are performed using the same set of random seeds. The interpretation is the same for all vehicle groups. At very small fleet sizes, part of the demand may remain unserved. Once all tasks can be completed, the main performance limitation becomes time-window compliance rather than task completion. Near the highest-performing fleet sizes, additional vehicles mainly add operational buffer and reduce the probability of rare late tasks.

The fleet-size choice is interpreted using two perspectives. Under the strict robust criterion, the required fleet size is the smallest tested fleet that achieves 100% on-time completion in every stochastic run. The pragmatic lower bound is interpreted using the mean on-time completion rate across replications, supported by the confidence-interval analysis reported in Subsection 6.3.10. Slightly smaller fleets may also be relevant if the number of late tasks is very small and the lateness is limited. The vehicle-specific subsections therefore report both the strict robust fleet size and the near-threshold alternatives.

The use of 30 replications is also supported by the uncertainty-validation results in Subsection 5.6.5, where the realised late- and early-arrival probabilities over 30 runs were shown to be consistent with the empirical FlightLegs-based uncertainty profiles. The confidence intervals in this section therefore assess the precision of the resulting on-time completion rates, while the uncertainty-validation results support the representativeness of the sampled input uncertainty within the modelled uncertainty scope.

Water Vehicles

Figure 6.11 shows the effect of water-service fleet size under combined task-timing and travel-time uncertainty. Each run contains 304 water-service tasks, meaning that each full fleet-size evaluation is based on 9,120 task instances.

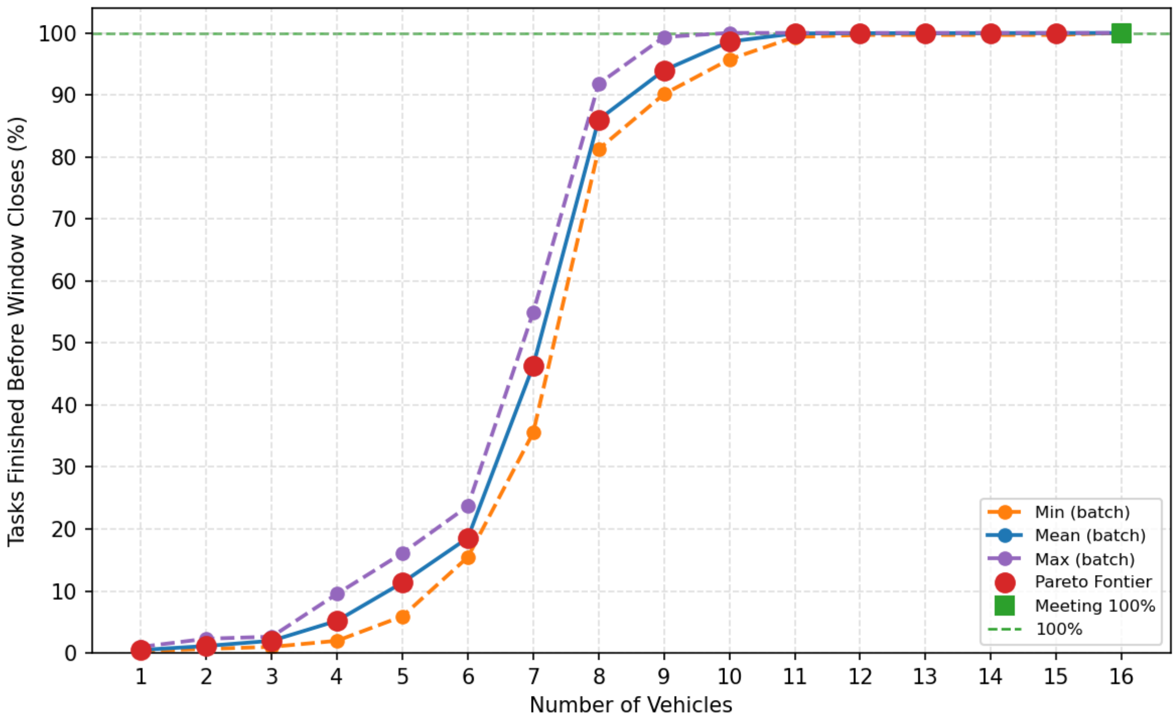


Figure 6.11: *Running different amounts of water vehicles under combined task-timing and travel-time uncertainty*

The results show a clear transition from structural undersizing to near-complete service reliability. With one to six vehicles, part of the demand remains unserved. From seven vehicles onward, all 9,120 task instances are completed, but many tasks are still completed after their delay-adjusted service-window closure. This indicates that the main bottleneck shifts from task completion to time-window compliance.

The strongest performance improvement occurs between six and ten vehicles. The mean on-time completion rate increases from 18.56% with six vehicles to 98.60% with ten vehicles. However, ten vehicles still result in 128 late tasks over 9,120 task instances, and 25 out of 30 runs contain at least one late task. This means that ten vehicles provide a high average service level, but not a robust solution under the tested stochastic conditions.

From eleven vehicles onward, the operation is close to full reliability. With eleven vehicles, only 8 tasks are late over all 9,120 task instances, resulting in an on-time completion rate of 99.91%. In this case, 23 out of 30 runs achieve full on-time completion, and the remaining late tasks have an average delay of 2.13 minutes. With twelve and thirteen vehicles, the number of late tasks decreases to 4, while fourteen and fifteen vehicles reduce this further to 2 late tasks. These results indicate that the system is already near-perfect from eleven vehicles onward, although none of these fleet sizes satisfies the strict 100% criterion across all stochastic runs. The near-threshold stochastic fleet-size results for water vehicles are summarised in Table 6.8.

Table 6.8: *Water-service performance for fleet sizes close to full on-time completion*

Vehicles	On-time rate (%)	Late tasks (total)	Runs with late tasks	Avg. late tasks/run	Avg. late-task delay (min)
8	86.02	1275/9120	30/30	42.50	11.20
9	93.97	550/9120	30/30	18.33	6.88
10	98.60	128/9120	25/30	4.27	3.70
11	99.91	8/9120	7/30	0.27	2.13
12	99.96	4/9120	4/30	0.13	1.75
13	99.96	4/9120	4/30	0.13	8.25
14	99.98	2/9120	2/30	0.07	1.00
15	99.98	2/9120	2/30	0.07	1.00
16	100.00	0/9120	0/30	0.00	–

The first fleet size that achieves 100% on-time completion in every stochastic run is sixteen water-service vehicles. At this fleet size, all 9,120 task instances are completed within their delay-adjusted service windows, with no late and no unserved tasks. Sixteen vehicles therefore represents the minimum robust fleet size for the water operation under the strict stochastic service-level criterion.

The near-threshold results are also operationally relevant. Eleven vehicles already result in only 8 late tasks out of 9,120, while fourteen vehicles reduce this to only 2 late tasks out of 9,120. The average lateness is also small for these near-threshold cases, except for thirteen vehicles, where the average is based on only four late tasks and should therefore be interpreted carefully. These fleet sizes do not provide a strict guarantee across all stochastic runs, but they may be defensible if a very small number of short internal service-window violations is acceptable.

The utilisation results show the cost of this additional robustness. Mean vehicle utilisation decreases from 49.72% with ten vehicles to 46.74% with eleven vehicles, 39.66% with fourteen vehicles, and 35.82% with sixteen vehicles. Additional vehicles therefore mainly provide buffer capacity rather than higher average workload per vehicle. From a strict service-level perspective, sixteen vehicles is the robust solution. From an operational-efficiency perspective, eleven to fourteen vehicles may be considered if the remaining late tasks are operationally acceptable.

Toilet Vehicles

Figure 6.12 shows the effect of toilet-service fleet size under combined task-timing and travel-time uncertainty. Each run contains 301 toilet-service tasks, meaning that each full fleet-size evaluation is based on 9,030 task instances.

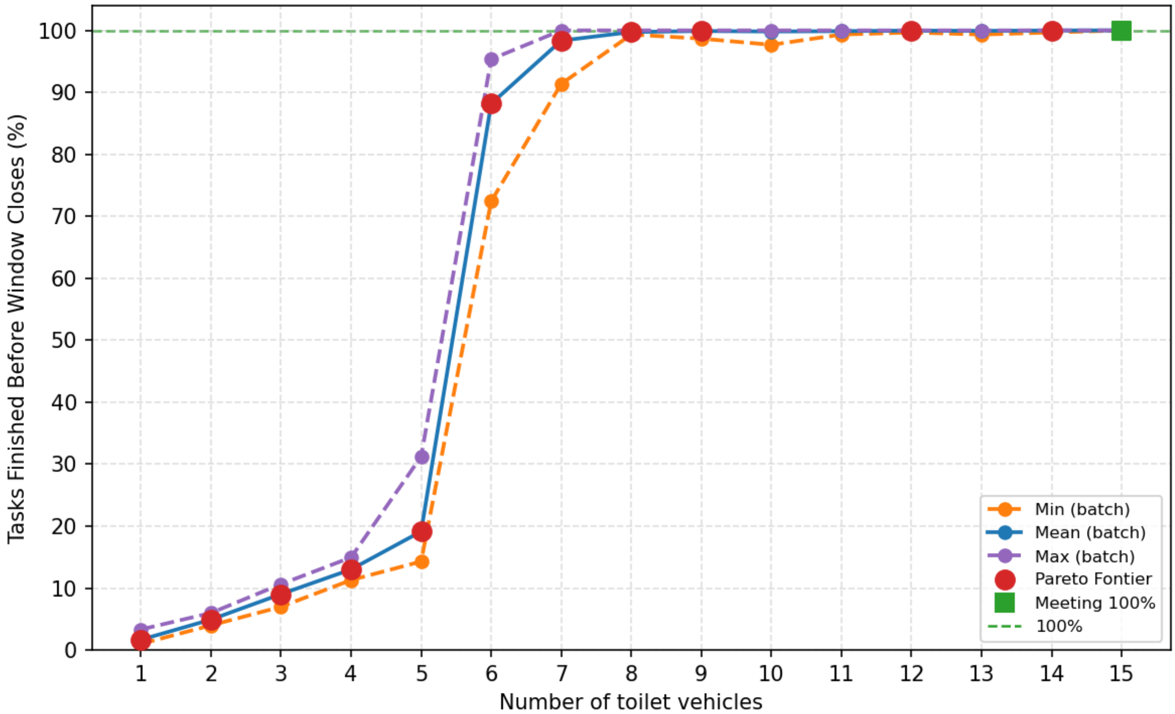


Figure 6.12: *Running different amounts of toilet vehicles under combined task-timing and travel-time uncertainty*

The results show a transition from structural undersizing to near-complete service reliability. With one to five toilet-service vehicles, part of the demand remains unserved. From six vehicles onward, all 9,030 task instances are completed, but not all tasks are completed within their delay-adjusted service windows. This means that the main bottleneck shifts from task completion to time-window compliance.

The strongest performance improvement occurs between five and seven vehicles. The mean on-time completion rate increases from 19.14% with five vehicles to 88.17% with six vehicles and 98.35% with seven vehicles. Around this transition region, one additional vehicle has a large operational effect. However, seven vehicles still result in 149 late tasks over 9,030 task instances, and 22 out of 30 runs contain at least one late task. This means that seven vehicles provide a high average service level, but not a robust solution under the tested stochastic conditions.

From eight vehicles onward, the operation is close to full reliability. With eight vehicles, only 24 tasks are late over all 9,030 task instances, resulting in an on-time completion rate of 99.73%. With nine vehicles, this improves further to 99.90%, with only 9 late tasks and 24 out of 30 runs achieving full on-time completion. The results for ten and eleven vehicles remain in the same high-performance range, but show small non-monotonic differences in the number of late tasks. These differences should not be interpreted as a structural deterioration when more vehicles are added. They occur in a region where the number of late tasks is already very small, and where stochastic task timing and travel-time variation can slightly change the final dispatching pattern.

With twelve to fourteen vehicles, the system is very close to full reliability. Twelve vehicles result in only 3 late tasks over 9,030 task instances, while fourteen vehicles reduce this to only 1 late task. With fourteen vehicles, 29 out of 30 runs achieve 100% on-time completion, and the only remaining late task has a delay of 4 minutes. Although this does not meet the strict 100% criterion, it represents a very high service level because the failure rate is limited to one late task in 9,030 task instances. The near-threshold stochastic fleet-size results for toilet vehicles are summarised in Table 6.9.

Table 6.9: *Toilet-service performance for fleet sizes close to full on-time completion*

Vehicles	On-time rate (%)	Late tasks (total)	Runs with late tasks	Avg. late tasks/run	Avg. late-task delay (min)
6	88.17	1068/9030	30/30	35.60	7.81
7	98.35	149/9030	22/30	4.97	5.87
8	99.73	24/9030	15/30	0.80	4.75
9	99.90	9/9030	6/30	0.30	3.22
10	99.82	16/9030	8/30	0.53	7.13
11	99.88	11/9030	10/30	0.37	2.73
12	99.97	3/9030	3/30	0.10	3.00
13	99.92	7/9030	6/30	0.23	5.43
14	99.99	1/9030	1/30	0.03	4.00
15	100.00	0/9030	0/30	0.00	–

The first fleet size that achieves 100% on-time completion in every stochastic run is fifteen toilet-service vehicles. At this fleet size, all 9,030 task instances are completed within their delay-adjusted service windows, with no late and no unserved tasks. Fifteen vehicles therefore represents the minimum robust fleet size for the toilet operation under the strict stochastic service-level criterion.

The near-threshold results are also operationally relevant. Twelve vehicles already result in only 3 late tasks out of 9,030, while fourteen vehicles reduce this to only 1 late task out of 9,030. These fleet sizes do not provide a strict guarantee across all stochastic runs, but they may be defensible if a very small number of short internal service-window violations is acceptable. Choosing fifteen vehicles instead of fourteen vehicles removes the final remaining late task, but requires one additional vehicle.

The utilisation results show the cost of this additional robustness. Mean vehicle utilisation decreases from 58.62% with six vehicles to 47.90% with eight vehicles, 35.83% with twelve vehicles, 32.03% with fourteen vehicles, and 30.85% with fifteen vehicles. Additional vehicles therefore mainly provide buffer capacity rather than higher average workload per vehicle. From a strict service-level perspective, fifteen vehicles is the robust solution. From an operational-efficiency perspective, twelve to fourteen vehicles may be considered if the remaining late tasks are operationally acceptable.

Loaders

Figure 6.13 shows the effect of loader fleet size under combined task-timing and travel-time uncertainty. Each run contains 147 loader tasks, meaning that each full fleet-size evaluation is based on 4,410 task instances.

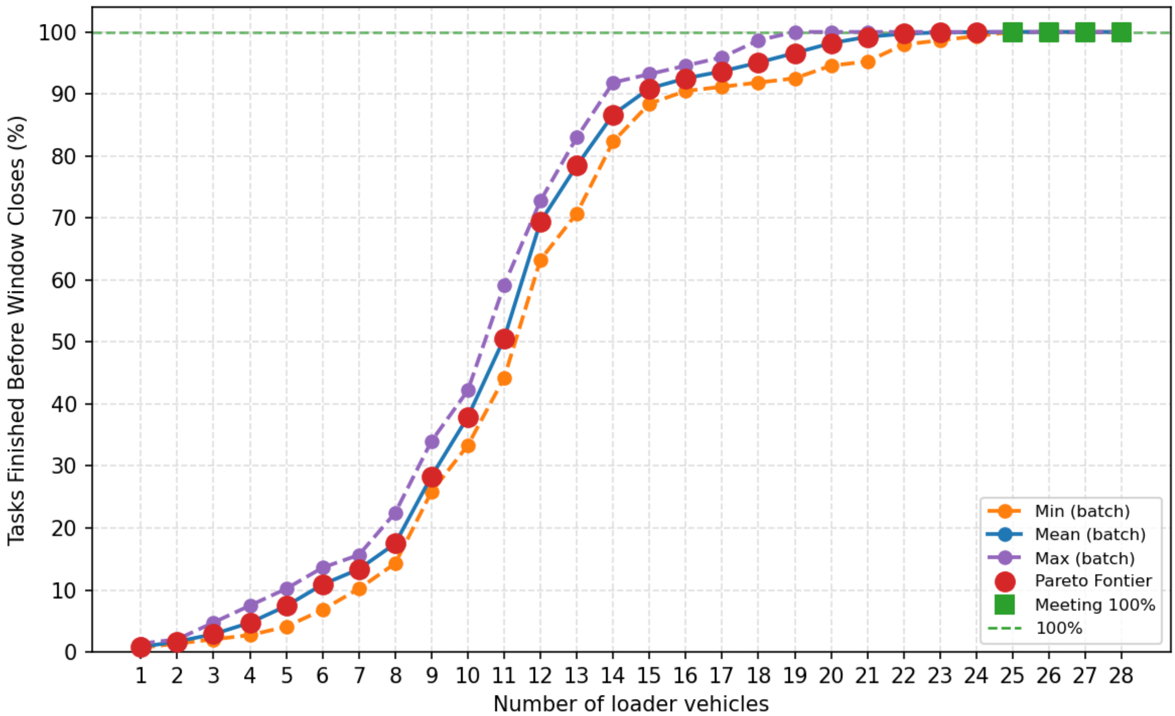


Figure 6.13: *Running different amounts of loader vehicles under combined task-timing and travel-time uncertainty*

The results show a transition from structural undersizing to near-complete service reliability. With one to five loaders, part of the demand remains unserved. From six loaders onward, all 4,410 task instances are completed, but many tasks are still completed after their delay-adjusted service-window closure. This indicates that the main bottleneck shifts from task completion to time-window compliance.

The strongest performance improvement occurs between 10 and 15 loaders. The mean on-time completion rate increases from 37.91% with 10 loaders to 90.88% with 15 loaders. After this point, additional vehicles continue to improve punctuality, but the marginal improvement becomes smaller. With 19 loaders, the on-time completion rate is 96.60%, but this still corresponds to 150 late tasks over 4,410 task instances. This means that 19 loaders provide a relatively high average service level, but not a robust solution under the tested stochastic conditions.

From 20 loaders onward, the system moves closer to full reliability. With 20 loaders, the number of late tasks decreases to 80, corresponding to 98.19% on-time completion. With 21 loaders, only 35 tasks are late, and 16 out of 30 runs achieve full on-time completion. With 22 loaders, the number of late tasks decreases further to 13, and 22 out of 30 runs achieve full on-time completion. These results show that 21 and 22 loaders already provide a high average service level, but they do not yet satisfy the strict 100% criterion across all stochastic runs.

The difference between 23 and 24 loaders is small, but relevant for the final fleet-size interpretation. With 23 loaders, only 3 tasks are late over all 4,410 task instances, giving a mean on-time completion rate of 99.93%. Only 2 out of 30 runs contain a late task. With 24 loaders, performance improves further to only 1 late task over all 4,410 task instances, corresponding to 99.98% on-time completion. Although 24 loaders does not meet the strict 100% criterion, it represents a near-perfect service level because the failure rate is limited to one late task in 4,410 task instances. The near-threshold stochastic fleet-size results for loaders are summarised in Table 6.10.

Table 6.10: *Loader performance for fleet sizes close to full on-time completion*

Loaders	On-time rate (%)	Late tasks (total)	Runs with late tasks	Avg. late tasks/run	Avg. late-task delay (min)
19	96.60	150/4410	29/30	5.00	5.66
20	98.19	80/4410	25/30	2.67	4.64
21	99.21	35/4410	14/30	1.17	2.94
22	99.71	13/4410	8/30	0.43	2.85
23	99.93	3/4410	2/30	0.10	3.67
24	99.98	1/4410	1/30	0.03	5.00
25	100.00	0/4410	0/30	0.00	–

The first fleet size that achieves 100% on-time completion in every stochastic run is 25 loaders. At this fleet size, all 4,410 task instances are completed within their delay-adjusted service windows, with no late and no unserved tasks. Fleet sizes above 25 loaders do not improve the service-level KPI, because 26, 27, and 28 loaders also achieve 100% on-time completion. Therefore, 25 loaders represents the minimum robust fleet size for the loader operation under the strict stochastic service-level criterion.

The near-threshold results are also operationally relevant. With 23 loaders, only 3 tasks are late out of 4,410, while 24 loaders reduce this to only 1 late task. These fleet sizes do not provide a strict guarantee across all stochastic runs, but they may be defensible if a very small number of short internal service-window violations is acceptable. However, because loader tasks are more closely linked to baggage and cargo handling around the aircraft, these near-threshold alternatives should be interpreted more cautiously than similar results for water or toilet vehicles.

The utilisation results show the cost of additional robustness. Mean vehicle utilisation decreases from 25.43% with 19 loaders to 20.22% with 24 loaders and 19.46% with 25 loaders. Additional vehicles therefore mainly provide buffer capacity rather than higher average workload per vehicle. Beyond 25 loaders, utilisation decreases further while the service-level KPI remains unchanged. From a strict service-level perspective, 25 loaders is the robust solution. From an operational-efficiency perspective, 23 or 24 loaders may be considered if the remaining late tasks are operationally acceptable.

6.3.6. Demand-Case Robustness Check for Water Vehicles

This subsection presents an additional robustness check for water vehicles only. It does not constitute a second full case study, because the simulation model, dispatching rules, uncertainty settings, and fleet assumptions remain unchanged. Only the operational input day is replaced. The reference case remains baseline day 14-06-2026 from week 24, while day 15-03-2027 from week 11 is used as a representative average-demand day. The purpose of this comparison is to assess whether lower demand reduces the required water-vehicle fleet size, and whether operational uncertainty remains relevant when demand pressure is lower.

Table 6.11 compares the input structure of the baseline day and the average day. The average day therefore represents a clearly lower-demand case in terms of total water-task volume and peak-hour workload, but it does not remove local concurrency pressure completely. Furthermore, the projected workload is reduced because the ground handling operations for Air France and Delta will no longer be carried out by KLM in the future operational setting, resulting in a lower expected number of water-service tasks.

Table 6.11: *Input comparison between the baseline day and the average day for water vehicles*

Metric	Baseline day	Average day	Change (%)
Total flight movements	811	727	−10.4
Arrivals	407	365	−10.3
Departures	404	362	−10.4
Water tasks	304	263	−13.5
Peak-hour water tasks	48	40	−16.7
Peak simultaneous water tasks	10	10	0.0
Average ground time (min)	81.2	109.1	+34.4
Median ground time (min)	55	80	+45.5
Commuter aircraft	352	334	−5.1
Narrow-body aircraft	312	280	−10.3
Wide-body aircraft	147	113	−23.1

The deterministic fleet-sizing results for the average day are shown in Table 6.12. All 263 water tasks are completed from fleet size 6 upward. The strict deterministic fleet size for the average day is 10 vehicles, compared with 12 vehicles on the baseline day.

Table 6.12: *Deterministic fleet-sizing results for water vehicles on the average day (15-03)*

Fleet size	Completed tasks	Unserved rate (%)	On-time rate (%)	Late tasks	Avg. lateness (min)	Distance (km)
6	263	0.00	41.06	155	34.5	310.8
7	263	0.00	70.72	77	16.4	320.2
8	263	0.00	86.69	35	9.1	338.8
9	263	0.00	95.06	13	2.7	355.1
10	263	0.00	100.00	0	0.0	381.7
11	263	0.00	100.00	0	0.0	403.0

Table 6.13 shows the average-day results under combined task-timing and travel-time uncertainty. The stochastic experiment consists of 30 replications per fleet size. At 11 vehicles, the rounded mean on-time completion rate reaches 99.90%. Therefore, 11 vehicles is interpreted as the pragmatic lower bound for the average day, while 13 vehicles is required for strict stochastic feasibility with 100% on-time completion in all replications.

Table 6.13: *Stochastic fleet-sizing results for water vehicles on the average day (15-03) under combined uncertainty*

Fleet size	Completed tasks/run	Unserved rate (%)	On-time rate (%)	Late tasks (total)	Avg. lateness (min)	Distance/run (km)
10	263	0.00	99.62	30	0.75	390.1
11	263	0.00	99.90	8	0.38	408.1
12	263	0.00	99.94	5	1.00	420.9
13	263	0.00	100.00	0	0.00	435.5

Note. Results are based on 30 replications per fleet size. On-time completion is reported as the mean across replications, late tasks are summed over all replications, and distance is the mean driven distance per replication.

The comparison between the baseline day and the average day is summarised in Table 6.14. The average day requires fewer vehicles under both deterministic and stochastic conditions. The strict deterministic fleet size decreases from 12 to 10 vehicles, a reduction of 16.7%. The strict stochastic fleet size decreases from 16 to 13 vehicles, a reduction of 18.8%. These reductions are broadly aligned with the reductions in water-task volume and peak-hour water demand. However, the relationship is not purely proportional to total task volume. Water-vehicle fleet requirements are driven not only by total daily demand, but also by temporal clustering, service windows, vehicle availability, refilling behaviour, charging behaviour, and travel-time constraints.

Table 6.14: *Fleet-size comparison between the baseline day and the average day for water vehicles*

Day type	Tasks	Det. strict	Stoch. strict	Pragmatic lower bound	Uncertainty buffer	Rel. uncertainty buffer (%)
Baseline day (14-06)	304	12	16	11	4	33.3
Average day (15-03)	263	10	13	11	3	30.0

The average-day experiment supports interpreting the baseline day water results as conservative upper-bound requirements for robust fleet sizing. The average day requires fewer vehicles, but uncertainty still has a material effect. The absolute uncertainty buffer decreases from four vehicles on the baseline day to three vehicles on the average day, while the relative uncertainty buffer decreases slightly from 33.3% to 30.0%. This means that lower demand reduces the uncertainty effect, but does not remove it. Even under average-day demand, strict stochastic feasibility still requires three vehicles above the deterministic strict fleet size.

A notable result is that the pragmatic lower bound is not lower on the average day. For the baseline day, 11 vehicles provide a pragmatic lower-bound solution, while the average day requires 11 vehicles for a more defensible pragmatic interpretation. This indicates that operating close to the deterministic minimum remains sensitive to uncertainty, even when total demand is reduced.

For KLM, the implication is that baseline day results are the more relevant basis for robust capacity planning, especially when 100% service reliability or peak robustness is required. The average-day results should not replace the baseline day fleet-size recommendation. Instead, they provide a useful benchmark for interpreting typical utilisation, efficiency, and the degree of conservatism in baseline day-based fleet sizing.

6.3.7. Operational-Control Variant: Idle Forward Staging

The fleet-sizing experiments show that late task completion can be caused not only by insufficient fleet size, but also by inefficient vehicle positioning between tasks. A diagnostic analysis of the baseline runs showed that several vehicles repeatedly returned to a central depot during short idle periods and were then dispatched back to the apron for the next demand peak. This created additional deadhead movements between the apron and the depot and reduced effective vehicle availability during busy periods.

To test whether this mechanism could be improved, an idle forward-staging policy was evaluated as an operational intervention. While there is still upcoming demand and its battery is healthy, an idle vehicle stages at its current apron position, the VOP, instead of deadheading back to the central depot. It still returns to the depot to charge when its state of charge falls below 40%, returns at end-of-shift, and still refills or dumps as usual when required. The vehicle is never interrupted while it is executing a task. The intervention is therefore a pure idle-positioning rule. It does not alter demand, task definitions, service windows, task durations, fleet size, dispatching logic, charging assumptions, or refill and dump rules.

For water vehicles, the intervention improves performance but does not reduce the deterministic fleet-size threshold. As shown in Table 6.15, the minimum fleet size for 100% on-time completion remains twelve vehicles. However, at equal fleet size, idle forward staging reduces driven distance by 25–38% and strongly reduces depot returns. At eleven vehicles, the number of late tasks decreases from seven to five, the average lateness decreases from 2.86 to 1.40 minutes, and the maximum lateness decreases from eight to two minutes. The detailed toilet results in Table 6.15 show why the intervention is operationally relevant. At a fleet size of nine vehicles, the baseline configuration still produces two late tasks and achieves 99.34% on-time completion. With idle forward staging, the same fleet reaches 100% on-time completion. At all tested toilet fleet sizes, the policy also strongly reduces driven distance and depot returns. This indicates that the improvement is not only a punctuality improvement at the threshold, but also a substantial reduction in unnecessary operational movement. For loaders, the fleet-size threshold also remains unchanged. It shows that the minimum deterministic fleet size remains 23 loaders. This is expected because loaders already have a low travel share and are initially distributed across multiple airside locations. The main benefit of forward staging is therefore not improved punctuality, but a large reduction in deadhead travel and depot returns. At the tested fleet sizes, driven distance decreases by 51–62%, while depot returns decrease from more than one hundred movements to thirteen.

Table 6.15: *Deterministic baseline and idle forward-staging results*

Group	Fleet	Variant	On-time (%)	Late tasks	Avg. late (min)	Max late (min)	Distance (km)	Depot returns
Water	10	Baseline	95.07	15	8.07	36	399.1	96
	10	Forward staging	96.05	12	5.00	13	298.5	15
	11	Baseline	97.70	7	2.86	8	425.4	119
	11	Forward staging	98.36	5	1.40	2	291.0	17
	12	Baseline	100.00	0	0.00	0	441.6	144
	12	Forward staging	100.00	0	0.00	0	272.4	20
Toilet	9	Baseline	99.34	2	2.50	3	451.9	129
	9	Forward staging	100.00	0	0.00	0	293.9	16
	10	Baseline	99.67	1	1.00	1	466.6	143
	10	Forward staging	100.00	0	0.00	0	286.2	18
	11	Baseline	100.00	0	0.00	0	489.3	171
	11	Forward staging	100.00	0	0.00	0	275.7	15
Loader	22	Baseline	98.64	2	3.00	4.0	54.9	112
	22	Forward staging	98.64	2	3.00	4.0	26.8	13
	23	Baseline	100.00	0	0.00	0	58.6	147
	23	Forward staging	100.00	0	0.00	0	22.5	13

The same-fleet results under combined uncertainty are summarised in Table 6.16. These results show that the benefit of idle forward staging is not limited to deterministic threshold changes. At equal fleet size under combined task-timing and travel-time uncertainty, the rule generally reduces the expected number of late tasks, increases the share of runs with 100% on-time completion for toilet and water vehicles, and consistently reduces driven distance and depot returns. For loaders, the service-level effect is limited and not strictly positive at every fleet size, but the reduction in operational movement remains substantial.

Table 6.16: *Idle forward staging under combined uncertainty (30 paired seeds, GSE_SIM_DDMM=1406)*

Group	Variant	Fleet	On-time rate	Late tasks	Runs with late tasks	Avg. late/run	Avg. delay	Distance (km)	Depot returns
Water	Baseline	10	98.60%	128/9120	25/30	4.27	3.70 min	411	108
	Forward staging	10	99.14%	78/9120	22/30	2.60	2.41 min	285	16
	Baseline	11	99.91%	8/9120	7/30	0.27	2.12 min	434	131
	Forward staging	11	99.96%	4/9120	4/30	0.13	1.50 min	273	18
	Baseline	12	99.96%	4/9120	4/30	0.13	1.75 min	455	153
	Forward staging	12	99.99%	1/9120	1/30	0.03	1.00 min	264	19
	Baseline	13	99.96%	4/9120	4/30	0.13	8.25 min	473	172
	Forward staging	13	99.98%	2/9120	2/30	0.07	1.00 min	248	21
	Baseline	14	99.98%	2/9120	2/30	0.07	1.00 min	491	190
	Forward staging	14	100.00%	0/9120	0/30	0.00	—	239	23
Toilet	Baseline	7	98.35%	149/9030	22/30	4.97	5.87 min	418	86
	Forward staging	7	99.01%	89/9030	16/30	2.97	4.60 min	324	20
	Baseline	8	99.73%	24/9030	15/30	0.80	4.75 min	443	115
	Forward staging	8	99.92%	7/9030	5/30	0.23	4.29 min	304	17
	Baseline	9	99.90%	9/9030	6/30	0.30	3.22 min	461	137
	Forward staging	9	99.99%	1/9030	1/30	0.03	8.00 min	283	15
	Baseline	10	99.82%	16/9030	8/30	0.53	7.12 min	486	159
	Forward staging	10	100.00%	0/9030	0/30	0.00	—	270	16
Loader	Baseline	22	99.71%	13/4410	8/30	0.43	2.85 min	53	121
	Forward staging	22	99.75%	11/4410	8/30	0.37	3.00 min	27	12
	Baseline	23	99.93%	3/4410	2/30	0.10	3.67 min	53	128
	Forward staging	23	99.89%	5/4410	4/30	0.17	2.60 min	24	13
	Baseline	24	99.98%	1/4410	1/30	0.03	5.00 min	54	133
	Forward staging	24	99.98%	1/4410	1/30	0.03	5.00 min	22	14
	Baseline	25	100.00%	0/4410	0/30	0.00	—	56	141
	Forward staging	25	100.00%	0/4410	0/30	0.00	—	20	14

Table 6.17 summarises the effect of idle forward staging on fleet-size thresholds under deterministic conditions and under combined operational uncertainty.

Table 6.17: *Effect of idle forward staging on fleet-size thresholds under deterministic and combined-uncertainty conditions.*

Vehicle group	Deterministic strict			Stochastic strict			Pragmatic lower bound		
	Base.	**VOP**	**Δ**	**Base.**	**VOP**	**Δ**	**Base.**	**VOP**	**Δ**
Water vehicles	12	12	0	16	14	−2	11	11	0
Toilet vehicles	11	9	−2	15	10	−5	9	8	−1
Loaders	23	23	0	25	25	0	23	23	0

Overall, idle forward staging is most effective for vehicle groups whose performance is reduced by unnecessary depot-centred deadheading during short idle periods. This effect is strongest for toilet vehicles. In the deterministic case, the required fleet size decreases from eleven to nine vehicles. Under combined uncertainty, the strict robust threshold decreases from 15 to 10 vehicles for toilet vehicles and from 16 to 14 vehicles for water vehicles. The pragmatic threshold also improves for toilet vehicles, while no fleet-size reduction is observed for loaders. For water vehicles, the deterministic fleet-size threshold remains unchanged. However, the combined-uncertainty results improve because vehicles spend less time returning unnecessarily to the depot and are therefore more likely to remain available near future demand locations. For loaders, the service-level effect is limited because their travel component is relatively small and their initial positioning is already more distributed across the apron. Nevertheless, the reduction in driven distance and depot returns remains operationally relevant.

These results show that idle forward staging should not be interpreted as a universal substitute for fleet capacity. Its effect depends on the underlying bottleneck. When lateness is caused by avoidable depot-return movements, forward staging can reduce fleet requirements. When lateness is caused by peak concurrency or stochastic travel-time variance, additional fleet capacity may still be required. The policy is therefore best interpreted as a conservative operational-control variant that reduces mileage and depot congestion, while improving robustness for selected vehicle groups.

6.3.8. Task-Selection Variant: Feasibility-Aware Triage

A second, smaller operational variant was tested to evaluate whether task ordering can improve performance under scarce-fleet conditions. In the baseline dispatcher, waiting tasks are ordered by urgency. This means that the task with the smallest remaining slack is considered first. When a task is already infeasible, however, assigning the nearest available vehicle to that task may prevent another task from being completed on time. The feasibility-aware triage variant therefore prioritises tasks that can still meet their service-window deadline and moves already-infeasible tasks later in the queue.

This variant changes only the ordering of waiting tasks. It does not change fleet size, demand, service windows, service durations, uncertainty settings, charging assumptions, refill and dump rules, or vehicle execution. It is therefore interpreted as an exploratory dispatching comparison rather than as a full redesign of the dispatching logic.

Table 6.18 shows the effect of the triage variant for all tested below-threshold fleet sizes. The results show that the rule can substantially reduce the number of late tasks when the fleet is scarce. The strongest effect is visible for water vehicles and toilet vehicles at very low fleet sizes, where the baseline dispatcher allows lateness to cascade across many tasks. The same mechanism remains visible closer to the deterministic threshold, but the improvement becomes smaller because fewer tasks are late in the baseline case.

Table 6.18: *Effects of feasibility-aware triage dispatching for all tested fleet sizes*

Group	Fleet	On-time (%) baseline → triage	Late tasks baseline → triage	Δ late tasks	Avg. lateness (min) baseline → triage	Δ avg. lateness (min)	Max late (min) baseline → triage
Water	4	6.9 → 57.9	178 → 30	−148	244.9 → 755.7	+510.8	677 → 1067
Water	5	11.8 → 70.7	210 → 42	−168	177.1 → 692.7	+515.6	451 → 908
Water	6	18.1 → 81.6	243 → 53	−190	81.1 → 409.9	+328.8	267 → 547
Water	7	42.1 → 91.4	176 → 26	−150	20.4 → 350.1	+329.7	57 → 462
Water	8	80.3 → 94.7	60 → 16	−44	14.6 → 265.1	+250.5	39 → 339
Water	9	94.1 → 97.4	18 → 8	−10	11.2 → 81.2	+70.1	20 → 130
Water	10	95.1 → 98.4	15 → 5	−10	8.1 → 32.6	+24.5	36 → 37
Water	11	97.7 → 99.0	7 → 3	−4	2.9 → 19.7	+16.8	8 → 21
Toilet	3	10.0 → 50.5	137 → 21	−116	275.2 → 870.2	+595.0	767 → 1071
Toilet	4	13.0 → 69.8	193 → 15	−178	184.4 → 591.0	+406.6	457 → 891
Toilet	5	17.3 → 80.4	213 → 40	−173	43.2 → 327.0	+283.8	250 → 455
Toilet	6	81.4 → 95.7	56 → 13	−43	11.5 → 87.5	+76.0	29 → 200
Toilet	7	96.7 → 98.7	10 → 4	−6	7.6 → 22.8	+15.2	12 → 25
Toilet	8	99.3 → 99.7	2 → 1	−1	2.0 → 17.0	+15.0	3 → 17
Toilet	9	99.3 → 99.3	2 → 2	0	2.5 → 13.5	+11.0	3 → 24
Toilet	10	99.7 → 99.7	1 → 1	0	1.0 → 2.0	+1.0	1 → 2
Loader	15	89.1 → 92.5	16 → 11	−5	15.7 → 27.1	+11.4	31 → 41
Loader	16	91.2 → 93.9	13 → 9	−4	14.6 → 23.1	+8.5	25 → 34
Loader	17	92.5 → 95.2	11 → 7	−4	13.0 → 21.3	+8.3	23 → 29
Loader	18	93.2 → 95.9	10 → 6	−4	10.0 → 18.5	+8.5	17 → 31
Loader	19	93.9 → 96.6	9 → 5	−4	8.2 → 15.0	+6.8	13 → 25
Loader	20	94.6 → 97.3	8 → 4	−4	3.5 → 11.8	+8.2	9 → 23
Loader	21	98.0 → 98.0	3 → 3	0	3.3 → 3.7	+0.3	6 → 6
Loader	22	98.6 → 99.3	2 → 1	−1	3.0 → 6.0	+3.0	4 → 6

The main trade-off is that the average lateness of the remaining late tasks increases. This is expected because the triage rule prioritises tasks that can still be completed on time and postpones tasks that are already infeasible. As a result, fewer tasks become late, but the tasks that remain late may wait longer. At very low fleet sizes, this effect becomes much stronger because the dispatcher effectively sacrifices a small number of already-lost tasks to prevent many other tasks from becoming late.

The distance effect is limited and mixed. Unlike idle forward staging, feasibility-aware triage does not directly reduce depot returns or deadhead travel. Its value lies mainly in saving recoverable tasks under scarcity, not in reducing operational mileage. Overall, the triage variant improves service reliability for several below-threshold fleet sizes, but it does not reduce the strict 100% fleet-size threshold. It should therefore be interpreted as a small exploratory dispatching variant rather than as a fleet-size reduction measure.

This variant is included as a small exploratory dispatching comparison. It is not developed as a full alternative dispatching framework and is not used to define the final fleet-size recommendations. Its purpose is to show whether task-ordering logic can reduce the number of late tasks under scarce-fleet conditions, and what trade-off this creates in terms of lateness severity.

6.3.9. Sensitivity Analysis of Operational and Energy Parameters

The sensitivity analysis is performed around fixed reference fleet configurations: 11 water vehicles, 11 toilet vehicles, and 25 loaders. These baseline configurations are used to evaluate how sensitive the model outcomes are to changes in key operational and energy-related assumptions. Task uncertainty and travel-time uncertainty are disabled in all sensitivity runs. The purpose of the sensitivity analysis is not to re-optimise the fleet size for each parameter setting, but to isolate the effect of individual assumptions on task punctuality, operational time allocation, and charging behaviour.

The results should therefore be interpreted as conditional on the selected baseline fleet sizes. The impact of a parameter change may differ at smaller or larger fleet sizes, because the amount of operational slack changes. For example, a travel-speed reduction may have a limited effect when sufficient vehicles are available, but may cause stronger punctuality losses when the fleet operates close to its minimum feasible size. The sensitivity analysis therefore indicates how the baseline configurations respond to parameter changes, but it does not replace the deterministic or stochastic fleet-sizing experiments.

The final fleet-size recommendation is based on the combined interpretation of the deterministic fleet-size results, the uncertainty experiments, and the marginal improvement of additional vehicles. The sensitivity analysis should therefore be interpreted as a robustness test around selected reference configurations, rather than as a standalone fleet-sizing method.

The tested parameters include service duration, travel speed, and an energy bundle consisting of the battery window, energy consumption, and charging speed. A reduction of the energy bundle means that the usable battery window and charging speed decrease, while energy consumption increases. An increase of the energy bundle represents the opposite combination. Parameters that directly affect task feasibility, such as service duration and travel speed, are expected to have a stronger effect than battery-related parameters when charging is not binding.

It should be noted that service times are fixed operational input values provided by KLM. Therefore, an increase in service time does not necessarily represent a realistic operational decision or expected change in the standard handling process. In some cases, increasing the service time may also make the required service duration longer than the available time within the corresponding service window. Such tasks can then no longer be completed on time by definition, regardless of dispatching performance or vehicle availability. This means that the negative effect of higher service times may partly overstate the operational impact. Nevertheless, the service-time sensitivity remains useful as a stress test of the planning structure. It indicates how much slack is available in the operational schedule and how vulnerable the operation is to longer-than-expected handling durations. The service-time sensitivity should therefore be interpreted primarily as an indicator of schedule robustness.

Water Vehicles

In the baseline run, all 304 water-service tasks are completed and the unserved task rate remains 0%. However, 7 tasks are completed after their service-window closure, resulting in an on-time task completion rate of 97.70%. The average delay of these late completions is 2.86 minutes. The baseline also shows that charging is not a binding operational constraint: there are 48 charging sessions, all charging sessions start immediately, and no charging queue occurs. The operational sensitivity results for water vehicles are reported in Table 6.19.

Table 6.19: *Operational sensitivity results for water vehicles*

Parameter	Change (%)	On-time rate (%)	Late tasks	Started within window (%)	Avg. late finish delay (min)
Travel speed	-20	96.38	11	98.36	5.64
Travel speed	-10	97.37	8	99.34	5.38
Travel speed	0	97.70	7	99.67	2.86
Travel speed	+10	99.01	3	100.00	2.33
Travel speed	+20	99.67	1	100.00	2.00
Service time	-20	100.00	0	100.00	0.00
Service time	-10	100.00	0	100.00	0.00
Service time	0	97.70	7	99.67	2.86
Service time	+10	93.42	20	98.36	5.35
Service time	+20	90.46	29	96.71	7.79

The operational sensitivity results show that the on-time task completion rate is mainly affected by the service time and, to a lesser extent, by travel speed. Reducing the service time by 10% or 20% eliminates all late completions and increases the on-time task completion rate to 100.00%. In contrast, increasing the service time by 10% reduces the on-time completion rate from 97.70% to 93.42%, while a 20% increase reduces it further to 90.46%. The number of late tasks increases from 7 in the baseline to 20 and 29 tasks respectively. The average delay of late completions also increases, from 2.86 minutes in the baseline to 5.35 minutes at +10% service time and 7.79 minutes at +20% service time.

Travel speed also affects punctuality, but its effect is smaller than the effect of service time. A 20% reduction in travel speed decreases the on-time task completion rate from 97.70% to 96.38%, corresponding to 11 late tasks. A 20% increase in travel speed improves the on-time task completion rate to 99.67%, with only one late task remaining. This indicates that travel time does influence the available operational slack, but that the water operation is more sensitive to changes in service duration than to changes in travel speed.

The operational time shares further support this interpretation. When travel speed is reduced by 20%, the travel time share increases to 12.19%, compared with 10.60% in the baseline. When travel speed is increased by 20%, the travel time share decreases to 8.84%. For service time, the service time share increases from 26.58% in the baseline to 29.26% at +10% service time and 31.49% at +20% service time. These changes reduce the time available for waiting, repositioning, and absorbing schedule deviations, which explains the decline in punctuality. The energy-related sensitivity results for water vehicles are reported in Table 6.20.

Table 6.20: *Energy sensitivity results for water vehicles*

Parameter	Change (%)	Charging sessions	Avg. charger occupancy (%)	Peak charger usage (%)
Charger power	0	48	13.75	63.64
Charger power	+100	37	6.96	45.45
Charger power	+200	33	4.68	36.36
Charger power	+300	32	3.74	27.27
Charger power	+400	31	3.01	36.36
Energy consumption	0	48	13.75	63.64
Energy consumption	+100	85	23.08	100.00
Battery bundle	-20	68	19.05	90.91
Battery bundle	-10	58	15.92	72.73
Battery bundle	0	48	13.75	63.64
Battery bundle	+10	38	11.78	72.73
Battery bundle	+20	34	10.39	63.64

The energy-related parameters mainly affect charging demand rather than task punctuality. The on-time completion rate remains 97.70% in all energy sensitivity runs, and all 304 tasks remain completed. Higher charger power reduces the number and duration of charging sessions, while doubled energy consumption increases charging sessions, average charger occupancy, and peak charger use. This indicates that, in the tested water configuration, energy-related assumptions are mainly relevant for charging-infrastructure robustness rather than for explaining late task completion.

Toilet Vehicles

In the baseline run, all 301 toilet-service tasks are completed and the unserved task rate remains 0%. All tasks are started and finished within their service windows, resulting in an on-time task completion rate of 100.00%. The baseline contains 48 charging sessions, no charging queue, and all charging sessions start immediately. This indicates that neither vehicle availability nor charging availability is a binding constraint in the deterministic baseline case. The operational sensitivity results for toilet vehicles are reported in Table 6.21.

Table 6.21: *Operational sensitivity results for toilet vehicles*

Parameter	Change (%)	On-time rate (%)	Late tasks	Service share (%)	Travel share (%)
Travel speed	-20	100.00	0	18.37	14.61
Travel speed	-10	100.00	0	18.18	13.19
Travel speed	0	100.00	0	18.28	12.05
Travel speed	+10	100.00	0	18.28	10.98
Travel speed	+20	100.00	0	18.28	10.40
Service time	-20	100.00	0	14.96	12.35
Service time	-10	100.00	0	16.47	12.32
Service time	0	100.00	0	18.28	12.05
Service time	+10	100.00	0	20.27	12.34
Service time	+20	100.00	0	21.50	11.89

The operational sensitivity results show that the toilet-service operation is robust to the tested changes in travel speed and service time. In all operational sensitivity runs, the on-time task completion rate remains 100.00%, no tasks are completed late, and the unserved task rate remains 0%.

This means that, within the tested parameter range, the toilet fleet has sufficient operational slack to absorb slower travel speeds and longer service durations.

Travel speed mainly affects the distribution of operational time rather than task punctuality. When travel speed is reduced by 20%, the travel time share increases from 12.05% in the baseline to 14.61%. When travel speed is increased by 20%, the travel time share decreases to 10.40%. However, these changes do not affect whether tasks are completed on time. This indicates that the toilet-service planning contains enough buffer to absorb the tested variation in travel time.

The service-time sensitivity shows a similar pattern. Increasing the service time by 10% and 20% increases the service time share from 18.28% in the baseline to 20.27% and 21.50% respectively. However, all 301 tasks are still completed within their service windows. This suggests that, in contrast to the water case, the toilet-service operation has more available slack in the deterministic schedule. The energy-related sensitivity results for toilet vehicles are reported in Table 6.22.

Table 6.22: *Energy sensitivity results for toilet vehicles*

Parameter	Change (%)	Charging sessions	Avg. charger occupancy (%)	Peak charger usage (%)	Queue time (min)
Charger power	0	48	18.98	77.78	0
Charger power	+100	38	10.24	77.78	0
Charger power	+200	36	7.01	66.67	0
Charger power	+300	34	5.33	66.67	0
Charger power	+400	34	4.30	66.67	0
Energy consumption	0	48	18.98	77.78	0
Energy consumption	+100	91	33.70	100.00	39
Battery bundle	-20	67	26.57	100.00	0
Battery bundle	-10	66	22.27	88.89	0
Battery bundle	0	48	18.98	77.78	0
Battery bundle	+10	46	16.82	77.78	0
Battery bundle	+20	40	13.77	77.78	0

The energy-related parameters mainly affect charging infrastructure load rather than task punctuality. The on-time completion rate remains 100.00% in all toilet energy sensitivity runs. Higher charger power reduces charging sessions and average charger occupancy, while doubled energy consumption increases charging sessions, average charger occupancy, peak charger use, and queueing. This shows that the tested toilet configuration contains enough operational slack to absorb the increased charging burden, although the charging infrastructure becomes more heavily loaded.

Loaders

In the baseline run, all 147 loader tasks are completed and the unserved task rate remains 0%. All tasks are started and finished within their service windows, resulting in an on-time task completion rate of 100.00%. The baseline contains 32 charging sessions, no charging queue, and all charging sessions start immediately. This indicates that neither vehicle availability nor charging availability is a binding constraint in the deterministic baseline case. The operational sensitivity results for loaders are reported in Table 6.23.

Table 6.23: *Operational sensitivity results for loaders*

Parameter	Change (%)	On-time rate (%)	Late tasks	Started within window (%)	Avg. late finish delay (min)
Travel speed	-20	100.00	0	100.00	0.00
Travel speed	-10	100.00	0	100.00	0.00
Travel speed	0	100.00	0	100.00	0.00
Travel speed	+10	100.00	0	100.00	0.00
Travel speed	+20	100.00	0	100.00	0.00
Service time	-20	100.00	0	100.00	0.00
Service time	-10	100.00	0	100.00	0.00
Service time	0	100.00	0	100.00	0.00
Service time	+10	0.00	147	100.00	3.29
Service time	+20	0.00	147	100.00	7.87

The operational sensitivity results show that the loader operation is robust to changes in travel speed, but highly sensitive to increases in service time. Changing travel speed between −20% and +20% does not affect task punctuality. In all travel-speed sensitivity runs, all 147 tasks are completed, all tasks are started within their service windows, and the on-time task completion rate remains 100.00%. This indicates that, in the deterministic loader case, travel time is not a binding constraint for task completion.

The effect of the service-time sensitivity is fundamentally different. Reducing the service time by 10% or 20% keeps the on-time task completion rate at 100.00%. However, increasing the service time by 10% or 20% reduces the on-time task completion rate to 0.00%. In both cases, all 147 tasks are still completed and all tasks are still started within their service windows, but all tasks finish after the service-window closure. The average finish delay outside the service window is 3.29 minutes for the +10% service-time case and 7.87 minutes for the +20% service-time case.

The operational time shares support this interpretation. In the baseline, the service time share is 16.18% and the travel time share is 1.20%. When the service time is increased by 10%, the service time share increases to 17.53%, and at +20% it increases to 19.40%. The travel time share remains low in all operational sensitivity runs. This confirms that the reduction in punctuality in the service-time sensitivity is caused by longer handling durations rather than by travel-time limitations. The energy-related sensitivity results for loaders are reported in Table 6.24.

Table 6.24: *Energy sensitivity results for loaders*

Parameter	Change (%)	Charging sessions	Avg. charger occupancy (%)	Peak charger usage (%)
Charger power	0	32	10.02	75.00
Charger power	+100	34	5.07	71.43
Charger power	+200	35	3.42	75.00
Charger power	+300	34	2.57	67.86
Charger power	+400	34	2.07	67.86
Energy consumption	0	32	10.02	75.00
Energy consumption	+100	68	16.75	57.14
Battery bundle	-20	50	13.93	64.29
Battery bundle	-10	41	12.27	67.86
Battery bundle	0	32	10.02	75.00
Battery bundle	+10	26	8.62	82.14
Battery bundle	+20	25	7.38	89.29

The energy-related parameters mainly affect charging demand rather than task punctuality. The on-time completion rate remains 100.00% in all loader energy sensitivity runs. Higher charger power reduces average charger occupancy, while doubled energy consumption and negative battery-bundle settings increase charging sessions and charger occupancy. Because no charging queues occur in these runs, the energy-related changes do not affect task completion in the deterministic loader configuration.

6.3.10. Result Synthesis

The result synthesis links the simulation outcomes back to the modelling requirements established earlier in the thesis. The operational foundations in Chapter 2 explain why fleet capacity depends on time-window feasibility, stand locations, vehicle movements, and service-specific constraints. The electrification section in Section 2.2 explains why vehicle availability must also be interpreted through SOC, charging access, and charging duration. The uncertainty section in Section 2.3 explains why deterministic sufficiency is not necessarily equivalent to stochastic robustness. The synthesis therefore interprets fleet-size thresholds as the combined outcome of operational demand, charging constraints, uncertainty, and control logic. This section synthesises the main findings from the deterministic fleet-sizing experiments, the stochastic fleet-sizing experiments, and the sensitivity analysis. The aim is to compare the strict fleet-size requirements with more pragmatic lower-bound alternatives and to clarify how uncertainty and operational assumptions affect the interpretation of the results.

Fleetsize Outcomes

The fleet-sizing results show that the required number of eGSE vehicles depends strongly on the interpretation of the service-level requirement. A strict interpretation requires 100% on-time task completion in all tested runs. A more pragmatic interpretation accepts a very small number of short internal service-window violations, provided that the overall service level remains high and the delays are operationally minor.

Under deterministic operating conditions, without task-timing uncertainty or travel-time uncertainty, the minimum strict fleet sizes are 12 water vehicles, 11 toilet vehicles, and 23 loaders. These are the smallest tested configurations that complete all tasks within their service windows under nominal assumptions. The deterministic results also show that some smaller configurations already perform close to the strict requirement. However, these configurations still produce late task completions and therefore do not satisfy the 100% on-time criterion.

When combined task-timing and travel-time uncertainty is included, the strict fleet-size requirements increase to 16 water vehicles, 15 toilet vehicles, and 25 loaders. This shows that uncertainty affects the required fleet capacity, even when the total number of daily tasks remains unchanged. The main effect of uncertainty is that it changes the timing of task demand and vehicle availability. As a result, vehicles may become unavailable during short demand peaks, which increases the need for operational buffer capacity.

The stochastic experiments also show that smaller fleets may still be operationally relevant under a less strict service-level interpretation. If a minimum on-time rate of 99.90% is accepted, the pragmatic stochastic lower bounds are 11 water vehicles, 9 toilet vehicles, and 23 loaders. These configurations achieve high service levels, but they are not equivalent to the strict robust fleet sizes. They should be interpreted as lower-bound alternatives that require operational judgement about the acceptability of the remaining late tasks. The fleet-size thresholds with and without idle forward staging are summarised in Table 6.25.

Table 6.25: *Fleet-size thresholds with and without idle forward staging*

Vehicle group	Policy	Deterministic strict	Stochastic strict	Pragmatic stochastic (≥ 99.90%)
Water vehicles	Baseline	12	16	≥ 11
	Idle forward staging	12	14	≥ 11
	Change	0	−2	0
Toilet vehicles	Baseline	11	15	≥ 9
	Idle forward staging	9	10	≥ 8
	Change	−2	−5	−1
Loaders	Baseline	23	25	≥ 23
	Idle forward staging	23	25	≥ 23
	Change	0	0	0

It is important to interpret these values as operational fleet-size requirements produced by the simulation model, not as final procurement recommendations. The table identifies the number of vehicles required to satisfy the tested service-level criteria during simulated daily operations. In practice, an additional technical reserve would still be required to account for maintenance, failures, battery degradation, charger unavailability, and other causes of temporary vehicle unavailability. The simulated fleet sizes therefore form the operational basis for fleet planning, after which a reserve margin should be added depending on vehicle criticality and maintenance policy.

The results therefore support two fleet-size interpretations. From a strict service-level perspective, the recommended fleet sizes are 16 water vehicles, 15 toilet vehicles, and 25 loaders. These are the smallest tested configurations that achieve 100% on-time task completion under combined operational uncertainty. From a more pragmatic operational perspective, smaller fleets may also be defensible if the remaining delays are rare, short, and do not affect aircraft departure punctuality. In the tested stochastic runs, the pragmatic lower-bound configurations achieve 99.91% on-time completion for water vehicles, 99.90% for toilet vehicles, and 99.93% for loaders, with average lateness values of 2.13, 3.22, and 3.67 minutes, respectively.

To assess the precision of the stochastic fleet-sizing results, replication-level 95% confidence intervals were calculated for the on-time completion rate at the main fleet-size thresholds. This check focuses on the combined uncertainty setting, because this experiment is used to determine the stochastic strict and pragmatic lower-bound fleet sizes. All stochastic experiment groups contained 30 replications with 30 unique random seeds.

Table 6.26 shows the confidence intervals for the key fleet sizes. These include the pragmatic lower-bound fleet sizes and the strict robust fleet sizes for each vehicle group. The confidence intervals are used only to assess the precision of the estimated mean on-time completion rate. They are not used to determine the strict robust threshold, which remains based on the requirement of 100% on-time completion in all sampled stochastic replications.

Table 6.26: *Replication-level confidence intervals for on-time completion rate under combined uncertainty*

Vehicle group	Stochastic strict	Pragmatic stochastic (≥ 99.90%)	Pragmatic mean on-time rate (%)	Pragmatic 95% CI (%)
Water vehicles	16	≥ 11	99.91	[99.85, 99.98]
Toilet vehicles	15	≥ 9	99.90	[99.80, 100.00]
Loaders	25	≥ 23	99.93	[99.83, 100.00]

The confidence intervals indicate that 30 replications provide a sufficiently narrow estimate of the mean on-time completion rate for the purpose of comparing the tested fleet-size thresholds, but they do not establish formal reliability guarantees. The half-widths of the confidence intervals are small, with values of approximately 0.06 percentage points for water vehicles and approximately 0.10 percentage points for toilet vehicles and loaders at the pragmatic lower-bound fleet sizes. The strict robust fleet sizes show no variation in on-time completion rate across the sampled replications, because all 30 replications achieve 100% on-time completion.

However, the confidence intervals also show that the pragmatic lower-bound fleet sizes are boundary cases. Water vehicles with 11 vehicles, toilet vehicles with 9 vehicles, and loaders with 23 vehicles all achieve a mean on-time completion rate around or above the 99.90% pragmatic threshold, but their lower confidence bounds fall below this threshold. These fleet sizes should therefore be interpreted as pragmatic lower bounds based on high average sampled performance, not as statistically strict guarantees that the 99.90% threshold will be exceeded under every possible stochastic realisation. This distinction is important for interpreting the fleet-size results. The pragmatic lower-bound fleet sizes are useful because they show that near-perfect service performance can be achieved with substantially fewer vehicles than the strict robust fleet size. At the same time, the strict robust threshold remains the more conservative criterion: it requires 100% on-time completion in every sampled stochastic replication.

The two exploratory operational-control variants provide additional insight into possible ways to improve daily execution, but they do not replace the main fleet-sizing results. Idle forward staging shows that vehicle positioning between tasks can influence punctuality and operational mileage. In particular, it reduces unnecessary depot-centred deadheading and can improve performance for vehicle groups where lateness is partly caused by repeated depot returns. Feasibility-aware triage dispatching shows that task-ordering logic can reduce the number of late tasks under scarce-fleet conditions by prioritising tasks that can still meet their service-window deadline. However, this also increases the lateness of the remaining late tasks. These variants therefore indicate promising operational-control directions, but they are not used as the basis for the final fleet-size recommendations.

Sensitivity Interpretation

The sensitivity analysis provides context for interpreting the fleet-size outcomes, but it should be read as a local robustness test around selected reference configurations. The tested configurations are 11 water vehicles, 11 toilet vehicles, and 25 loaders. Therefore, the sensitivity results do not provide general conclusions for all possible fleet sizes. Sensitivity to operational parameters depends on the amount of available slack. Smaller fleets have less idle time and are generally more vulnerable to changes in service duration, travel speed, and depot-related assumptions. Larger fleets may absorb part of these changes through additional vehicle availability.

Across the tested reference configurations, charging is not the main cause of late task completion. Energy-related parameters mainly affect the number of charging sessions, charger occupancy, peak charger use, and charging infrastructure utilisation, while the on-time task completion rate remains unchanged in the tested runs. This indicates that, for the analysed experiments, the main operational constraint is vehicle availability during demand peaks rather than energy feasibility.

The charger-use results also suggest that the assumed charging infrastructure may be conservative for the tested conditions. Even in the most constrained charging case, which occurs for toilet vehicles because this is the only tested group with fewer chargers than vehicles, charging does not become a relevant operational bottleneck. The total charging waiting time is only 39 minutes, and this waiting occurs while vehicles still have relatively high SOC levels of 59.8% and 69.0%. The vehicles therefore request charging during idle time rather than because they are critically low on energy.

For the water operation, the sensitivity analysis is directly relevant to the pragmatic lower bound because it is based on 11 water vehicles. The results show that this configuration is sensitive to service duration and travel speed. This means that the operational acceptability of an 11-vehicle water fleet depends strongly on whether the assumed service times, travel times, and depot-return behaviour are realistic. It does not imply that larger water fleets, such as 12 or 16 vehicles, would be equally sensitive.

For the toilet operation, the sensitivity analysis is based on 11 vehicles. This configuration corresponds to the deterministic strict fleet size, while the pragmatic stochastic lower bound is 9 vehicles and the strict stochastic fleet size is 15 vehicles. The 11-vehicle configuration remains relatively stable under the tested changes in travel speed and service duration. This supports the interpretation that the deterministic strict toilet fleet contains operational slack under nominal conditions. However, it does not prove that the 9-vehicle lower bound would remain equally robust under changed assumptions.

For loaders, the sensitivity analysis is based on 25 loaders, which corresponds to the strict stochastic fleet size. The results are therefore most useful for interpreting the robustness of the strict loader configuration. The analysis shows that 25 loaders are robust to changes in travel speed, but highly sensitive to increases in service duration. This indicates that loader performance is mainly constrained by the relationship between service duration and service-window length. The pragmatic lower bound of 23 loaders should therefore be interpreted cautiously, because it has less operational buffer than the tested 25-loader configuration.

Overall, the deterministic experiments identify the nominal fleet-size requirement, the stochastic experiments show the additional capacity needed for robustness under uncertainty, and the sensitivity analysis identifies which assumptions are most important for interpreting these outcomes. The final fleet-size decision should therefore balance service reliability, vehicle utilisation, operational slack, charging infrastructure use, and the operational relevance of the remaining late tasks.

7 Discussion

This chapter discusses the transferability, limitations, modelling choices, and interpretation of the simulation results for eGSE fleet sizing. First, the generalisability of the simulation framework is examined by distinguishing between the transferable model structure and the case-specific numerical outcomes for the KLM Schiphol case. Second, the main general and case-specific limitations are discussed. Third, the principal modelling choices are reflected upon, with particular attention to the balance between operational realism, transparency, data availability, and computational tractability. Finally, the fleet-sizing outcomes are interpreted by distinguishing between deterministic minimum fleet sizes, stochastic strict thresholds, pragmatic lower-bound fleet sizes, and technical fleet requirements. The implications of operational uncertainty, the average-day check, idle forward staging, feasibility-aware triage, charging behaviour, and the sensitivity analyses are considered throughout these sections.

7.1. Generalisability of the Model

The developed simulation framework is transferable as a modelling approach, but the numerical fleet-sizing outcomes are case-specific. The transferable contribution is the model structure, in which eGSE operations are represented through time-window-based service tasks, individual vehicle states, vehicle movements, battery state, charging behaviour, service-specific resource constraints, and capacity-constrained infrastructure. These mechanisms are not unique to the KLM Schiphol case. They occur in many airport ground-handling systems in which mobile service resources must complete time-critical tasks under operational and energy-related constraints.

The resulting fleet requirements are nevertheless determined by local operating conditions. Relevant factors include the flight schedule, aircraft mix, stand distribution, apron layout, travel-time assumptions, service rules, vehicle characteristics, charger availability, refill and dump infrastructure, operational-control policies, and uncertainty profiles. The fleet sizes identified in this study should therefore not be transferred directly to another airport, airline, season, or operating day. They represent the simulated operational requirements for the tested KLM Schiphol configurations rather than universal eGSE fleet-size values.

This distinction also applies to the broader set of GSE types identified in the case-study chapter. The model framework is designed so that additional vehicle groups can be represented within the same general structure, provided that suitable vehicle-specific input data and operational rules are available. However, the validated experimental scope of this thesis is limited to water vehicles, toilet vehicles, and loaders. The results should therefore not be interpreted as validated fleet-sizing conclusions for all thirteen potential GSE types.

The selected vehicle groups nevertheless provide a useful basis for assessing the broader applicability of the framework. Water and toilet vehicles represent depot-based and resource-constrained operations with onboard capacity limitations and intermediate refill or dump visits. Loaders represent stand-side service equipment without onboard water or waste-resource constraints. Together, these groups test two distinct operational modes within one simulation structure: depot-based replenishment operations and stand-side task execution.

The effort required to add another vehicle type depends on whether its behaviour can be represented through the existing model logic. Some additional GSE types may primarily require new input parameters, such as service duration, aircraft compatibility, starting location, travel speed, battery capacity, charging thresholds, and energy consumption. Other vehicle types may require additional model mechanisms. Examples include stronger coupling to departure events, multi-vehicle task dependencies, weather-dependent task generation, towing-specific safety constraints, stand-access restrictions, or explicit precedence relations with other turnaround activities. Extending the model is therefore not necessarily a parameterisation exercise only. Additional modelling choices, verification, and validation may be required.

The same principle applies when the framework is transferred to another airport or operational context. The local operation must first be translated into suitable task definitions, spatial mappings, vehicle rules, infrastructure constraints, and uncertainty assumptions. Each application also requires validation using local operational data and expert judgement. Without this step, the model structure may remain relevant, but its outputs would not be sufficiently reliable for fleet-capacity decisions.

The framework should therefore be interpreted as a reusable decision-support structure rather than as a source of universal fleet-size values. Its main generalisable contribution is the integrated representation of task execution, vehicle availability, charging, service-specific resources, supporting infrastructure, and operational uncertainty within a dynamic simulation environment. The numerical results remain valid only within the tested case-study scope, while the framework provides a structured basis for evaluating other airports, demand cases, seasons, and vehicle groups after appropriate parameterisation, verification, and validation.

7.2. Limitations

Several limitations must be considered when interpreting the results. These limitations do not invalidate the simulation outcomes, but they define the scope within which the results can be understood and used for decision support.

7.2.1. General Modelling Limitations

A first limitation is the use of rule-based operational-control logic. Dispatching, task selection, vehicle selection, charging decisions, refill and dump decisions, and depot-return behaviour are implemented through transparent rules. This improves interpretability, enables consistent comparison between experiments, and keeps the model computationally manageable for repeated fleet-sizing and uncertainty analyses.

However, rule-based control does not fully represent human decision-making, informal coordination, operator experience, radio communication, local situational awareness, or ad hoc real-time interventions. In practice, operators and dispatchers may deviate from standard procedures when they anticipate congestion, identify an operationally critical task, know that another vehicle will soon become available, or coordinate with other turnaround activities. The model should therefore be interpreted as a structured approximation of operational control rather than as a complete representation of daily human decision-making.

A second limitation concerns the uncertainty scope. The active stochastic inputs are task-timing uncertainty and travel-time uncertainty. These sources were selected because they directly affect the two principal drivers of dynamic vehicle availability. Task-timing uncertainty changes when aircraft-dependent service demand occurs, while travel-time uncertainty changes when vehicles reach tasks and become available after previous assignments. Historical operational flight data were available to support the task-timing profiles, while travel-time uncertainty could be implemented as a transparent and bounded operational assumption.

Other relevant uncertainty sources are not sampled stochastically. These include service-duration variation, task-occurrence uncertainty, variation in water and waste quantities, variation in energy consumption, operator behaviour, operator availability, weather effects, charger failures, vehicle breakdowns, battery degradation, aircraft stand changes, and disruption-driven rescheduling. These factors are addressed through deterministic assumptions, sensitivity analyses, validation checks, expert review, or explicit scope limitations, but they are not jointly sampled in the stochastic experiments.

The excluded uncertainty sources could affect the fleet-size conclusions in different ways. Vehicle breakdowns, charger failures, and battery degradation would mainly increase temporary vehicle unavailability and could consequently increase the required technical reserve. Service-duration variation could create additional task overlap and is particularly relevant for loaders because their service windows contain limited operational slack. Aircraft stand changes could increase travel demand and alter the spatial concentration of tasks. Weather conditions could affect travel time, service duration, battery performance, and energy consumption. Task-occurrence uncertainty could change the total operational workload, whereas the implemented task-timing uncertainty mainly changes the timing of demand. The stochastic fleet-sizing outcomes should therefore be interpreted as robustness under the two uncertainty mechanisms tested in this thesis, not as robustness against the complete range of real-world disturbances.

The primary performance metric also creates an interpretive limitation. On-time task completion measures whether a task is completed within its internal service window. This provides a clear and consistent basis for comparing fleet sizes, uncertainty settings, sensitivity cases, and operational-control variants. However, an internal service-window violation is not equivalent to an aircraft departure delay. A task may be completed shortly after its internal deadline without affecting OB punctuality if sufficient turnaround slack remains. Conversely, a short delay in an operationally critical task may have a greater impact than a longer delay in a less critical activity.

Because the model does not represent the complete turnaround critical path or explicit aircraft-delay propagation, its outputs should be interpreted as ground-handling service-performance outcomes rather than as direct predictions of aircraft punctuality. The on-time completion percentage also treats all service-window violations equally. A task that is one minute late receives the same binary status as a task that is substantially late. Supporting indicators such as the number of late tasks, average lateness, maximum lateness, and unserved tasks partly address this limitation, but the model does not explicitly weight tasks according to operational criticality or potential departure impact. Linking internal service-window violations more directly to aircraft departure punctuality would improve the operational interpretation of the results.

Charging behaviour is represented explicitly, but the battery and charging model remains simplified. The model includes usable battery capacity, energy consumption, SOC, charging thresholds, charger power, charger capacity, charging sessions, charging queues, and charging-related vehicle unavailability. This level of detail is appropriate for the operational fleet-sizing objective of the study. However, the model does not include detailed electrochemical battery behaviour, temperature-dependent performance, long-term battery degradation, charger reliability, grid constraints, or detailed charger-sharing strategies across all possible GSE users. Under conditions such as fewer chargers, charger failures, cold-weather performance losses, degraded batteries, higher energy consumption, or more intensive demand patterns, charging could become a more restrictive operational constraint.

The overall model scope also limits the interpretation of the results. The simulation represents selected GSE operations and their interaction with vehicle availability, charging, resource replenishment, infrastructure constraints, and operational uncertainty. It does not represent the complete aircraft turnaround process, detailed airside traffic interactions, crew availability, passenger boarding, baggage dependencies, fuelling interactions, aircraft stand changes, or multi-actor operational coordination. This scope is appropriate for analysing fleet capacity and eGSE availability, but the model should not be used as a real-time dispatching system or a complete aircraft-delay prediction model without further operational integration, validation, and model development.

The tested operational-control variants should also be interpreted as exploratory. Idle forward staging demonstrates that depot-centred repositioning can create avoidable inefficiency and that vehicles may benefit from remaining closer to future demand. Practical implementation may nevertheless be constrained by apron safety regulations, available parking space near VOP positions, driver handover procedures, dispatcher visibility, charging access, and local operating agreements.

Feasibility-aware triage is similarly an exploratory task-selection rule. It can increase the number of on-time completions by prioritising tasks that can still be completed within their service windows. However, this may increase the lateness of tasks that are already infeasible. The rule therefore improves one performance objective, but does not necessarily minimise total lateness, average lateness, maximum lateness, or aircraft-level delay risk. Practical implementation would require additional safeguards related to task criticality, maximum acceptable lateness, operational priority, departure impact, and fairness between tasks.

7.2.2. Case-Study Modelling Limitations

A principal limitation of the case-study application is the number of operating days included. The main experiments are based on one selected peak-demand day. This is appropriate for stress testing because operationally critical eGSE fleets should not be sized only for average demand. However, one day cannot represent the complete variation in KLM operations across seasons, weekdays, weekends, holiday periods, winter conditions, irregular disruption days, and different aircraft mixes.

The average-day robustness check for water vehicles partly addresses this limitation, but it does not replace a broader multi-day analysis across a representative portfolio of operating days. The fleet-size outcomes should therefore be interpreted as configuration-specific results for the tested demand conditions rather than as universal fleet requirements for all KLM operating days.

Validation is also limited by the available case-study data. The model was assessed using operational data, expert judgement, service-demand checks, energy and charging behaviour, water-demand checks, and uncertainty validation. This provides an adequate basis for system-level scenario analysis. However, not every process could be validated using detailed vehicle-level observations. Travel-time variability, exact service execution times, operator behaviour, vehicle-level energy consumption, and several resource-use assumptions remain partly dependent on modelling assumptions and expert review.

The strength of the empirical basis also differs between vehicle groups and processes. It is strongest for water vehicles, more limited for toilet vehicles, and partly qualitative for loaders. Service-resource validation is stronger for water and toilet vehicles than for loaders because the first two operations include measurable onboard resource processes. Detailed validation of charging behaviour and SOC development is also constrained by the availability of vehicle-specific operational data.

The model should consequently be interpreted as a credible decision-support approximation rather than as a fully calibrated digital twin of KLM ground operations. Additional validation of service times, travel times, depot dwell assumptions, vehicle-level charging behaviour, and SOC development would strengthen the empirical basis of the model.

The conclusion that charging is not a binding punctuality constraint is also case-specific. In the tested KLM Schiphol configurations for water vehicles, toilet vehicles, and loaders, late task completion is mainly associated with vehicle availability during demand peaks, travel time, service duration, positioning logic, and depot-return behaviour. Charging primarily affects charging sessions, charger occupancy, charging queues, peak charger use, and the distribution of vehicle time. This finding is conditional on the tested assumptions for battery capacity, charger availability, charging power, energy consumption, fleet size, and operational demand. It should therefore be reassessed under reduced charger availability, charger failures, vehicle unavailability, battery degradation, higher energy consumption, shared infrastructure use, and alternative demand days.

Finally, the simulated operational fleet requirements do not include a technical reserve. The simulation estimates how many vehicles must be operationally available to execute the modelled tasks. It does not explicitly simulate planned maintenance, preventive maintenance, unexpected breakdowns, inspections, cleaning, long-term battery degradation, or extended charger and connector failures. The operational fleet-size results should therefore not be used directly as procurement quantities. A technical reserve must be added separately based on fleet criticality, maintenance performance, expected technical availability, and access to backup vehicles. The required reserve may differ between vehicle groups and may be particularly important for small or operationally critical fleets.

Overall, the case-study results should be used as structured decision-support evidence rather than as final operational prescriptions. Final implementation decisions require testing across additional representative operating days, a separate technical-reserve assessment, operational feasibility checks, and expert review of the proposed control policies. Further testing should include disruption scenarios, charging-infrastructure scenarios, and vehicle-availability scenarios.

7.3. Reflection on Model Scope and Modelling Choices

This section reflects on the rationale behind the main modelling choices and their influence on the usefulness, generalisability, and interpretation of the results. The reflection first addresses the overall modelling approach and subsequently discusses specific choices made for the KLM Schiphol implementation.

7.3.1. General Reflection

The principal value of the developed simulation framework is that it combines fleet sizing, task execution, vehicle availability, charging behaviour, supporting infrastructure, service-specific resource constraints, and operational uncertainty within one dynamic model. This integrated perspective is required because these elements are strongly interdependent in electrified ground-handling operations.

A vehicle is operationally available only when several conditions are satisfied. It must be located appropriately, have sufficient battery state, satisfy service-specific resource constraints, and be able to complete the assigned task within its time window. Charging is similarly not only an energy-management process. It determines when vehicles are unavailable for service and how charging infrastructure is used during the operating day.

The selected model scope reflects a deliberate balance between operational realism and tractability. The model includes the mechanisms considered most relevant to the research objective: time-window-based tasks, individual vehicle states, travel between aggregated airport locations, battery state, charging, refill and dump processes, depot-return behaviour, and stochastic variation in task timing and travel time. It excludes the complete turnaround process, detailed operator coordination, and detailed battery degradation behaviour.

This abstraction allows the model to remain transparent and computationally manageable while representing the main causes of temporary vehicle unavailability. Adding more detail would not automatically improve the quality of the fleet-size conclusions. Additional mechanisms are useful only when they can be supported by sufficiently reliable data and when they materially affect the research objective.

The use of a rule-based DES is also a deliberate choice. An optimisation model could search for globally optimal task assignments, routes, or charging schedules. However, such a model would require stronger assumptions regarding objectives, constraints, future information, and operational feasibility. Integrated optimisation of task assignment, routing, charging, individual vehicle states, stochastic disturbances, and service windows may also become computationally demanding.

The rule-based simulation instead represents operational execution through transparent dispatching and control logic. This is consistent with day-of-operations ground handling, where decisions are made under time pressure and with limited information about future disturbances rather than under perfect global coordination. The model therefore does not assume optimal airport operations. It evaluates how specified operational rules perform under dynamic conditions.

This distinction is important when interpreting fleet-size outcomes. A simulated fleet requirement is determined partly by physical demand and partly by the operational rules used to control the fleet. Inefficient positioning, depot-return, charging, or dispatching rules can result in a higher apparent fleet requirement than more effective control policies. Fleet sizing and operational policy should therefore not be treated as completely separate decisions.

A second central modelling choice concerns the definition of acceptable performance. The same simulation outcomes can lead to different fleet-size conclusions depending on whether the decision-maker applies a deterministic minimum, a stochastic strict threshold, or a pragmatic lower-bound interpretation. This is not solely a technical modelling issue. It reflects the decision-maker's tolerance for operational risk, vehicle investment, underutilisation, and internal service-window violations.

A strict 100% on-time requirement provides a conservative service-level criterion. However, it may result in an additional vehicle being added to eliminate only a small number of short service-window violations. A pragmatic lower-bound interpretation can use the fleet more efficiently, but accepts a limited residual risk. The required service level should therefore be defined before simulation outcomes are translated into fleet or investment decisions.

The results also demonstrate why one performance indicator is insufficient. The on-time completion rate provides a clear primary KPI, but does not explain why tasks are late. Supporting indicators such as late-task counts, average lateness, maximum lateness, unserved tasks, vehicle utilisation, driven distance, depot returns, charging sessions, charging-threshold behaviour, charger occupancy, peak charger use, charging queues, and infrastructure interaction are needed to diagnose the underlying bottleneck.

A lower service level may be caused by insufficient fleet capacity, but may also result from peak clustering, unnecessary depot returns, long service durations, travel assumptions, onboard resource constraints, infrastructure interaction, or imposed depot dwell behaviour. The model is therefore most useful as a diagnostic decision-support tool rather than only as a method for selecting one fleet-size number.

The modelling approach should not replace operational judgement. Its primary strength is that assumptions are made explicit and can be tested consistently across experiments. It can quantify how fleet size, uncertainty assumptions, charging parameters, service assumptions, depot-return rules, charging thresholds, infrastructure constraints, and task-selection logic affect operational performance. The results are most valuable when combined with expert judgement, operational feasibility checks, and testing across multiple representative operating days.

The modelling choices made in this thesis are therefore appropriate for evaluating eGSE fleet capacity under operational uncertainty and charging constraints. The model does not reproduce every detail of airport ground handling, but it captures the principal mechanisms through which electrification, timing variability, vehicle availability, depot-return behaviour, resource constraints, charging, and operational control interact. The most useful next step is not necessarily to add model detail indiscriminately, but to strengthen the empirical basis of the existing structure through broader operating-day coverage and more extensive validation.

7.3.2. Case-Study Reflection

A specific modelling choice in the case-study application concerns the depot dwell time for water and toilet vehicles. When these vehicles return to the depot, they remain there for a minimum of 15 minutes before becoming available again. This prevents an unrealistic process in which a vehicle returns to the depot, becomes available for only a very short period, and immediately departs for another task.

The dwell period approximates operational processes that are not represented explicitly, such as parking, operator handover, administrative handling, vehicle changeover, or the practical time required before redeployment. In reality, this process may occasionally be shorter, for example when another operator can use the vehicle immediately. The 15-minute dwell time should therefore be interpreted as a simplifying and partly conservative operational assumption.

This assumption also affects the interpretation of idle forward staging. The benefit of the variant does not arise only from improved vehicle positioning. It combines three effects: avoiding unnecessary travel to the depot, keeping vehicles closer to future demand, and avoiding the modelled depot dwell period. The results therefore demonstrate that depot-return policy can influence the required operational capacity, but they do not prove that the complete simulated fleet reduction can be achieved in practice without further operational testing.

Another case-study insight concerns the role of charging. Electrification may create the expectation that battery capacity, charging duration, and charger availability will automatically dominate fleet-sizing outcomes. These elements are important and must be represented in an eGSE model. However, the tested KLM Schiphol experiments show that charging is not necessarily the binding constraint for task punctuality.

In the reference configurations, service punctuality is mainly influenced by temporal vehicle availability, task overlap, travel time, service duration, vehicle positioning, and depot-return behaviour. Charging remains operationally relevant because it changes the distribution of vehicle time and determines charging sessions, charger occupancy, peak charger use, charging queues, and infrastructure utilisation. The relevant modelling lesson is therefore not that charging can be ignored, but that the dominant bottleneck should be identified from the interaction between demand, fleet capacity, vehicle states, charging rules, and infrastructure rather than assumed in advance.

The two operational-control variants illustrate different intervention mechanisms. Idle forward staging changes where vehicles wait and how quickly they can reach future demand. Feasibility-aware triage changes which tasks receive scarce capacity when not every task can still be completed on time. The first variant mainly reduces avoidable vehicle movement and temporary unavailability. The second changes how lateness is distributed across tasks. Both demonstrate that operational policy can influence the apparent fleet-capacity requirement, but neither should be treated as a final implementation policy without further assessment of safety, parking availability, operator deployment, dispatcher procedures, charging access, task criticality, and operational acceptance.

7.4. Interpretation of Fleet-Sizing Outcomes

This section interprets the fleet-sizing outcomes beyond the numerical vehicle requirements reported in the results chapter. It explains how the different simulated fleet sizes should be understood and how service-level definitions, uncertainty, operational-control rules, and technical availability affect the resulting capacity conclusions.

7.4.1. General Interpretation

Four fleet-size concepts should be distinguished. The deterministic minimum fleet size is the smallest fleet that achieves 100% on-time task completion under fixed operating assumptions. It represents nominal operational sufficiency without stochastic task-timing or travel-time variation.

The stochastic strict threshold is the smallest tested fleet that achieves 100% on-time task completion in all 30 sampled replications of the relevant uncertainty setting. It provides a conservative, experiment-based robustness criterion. However, it is not a formal probabilistic guarantee across all possible operating days, disturbances, unmodelled uncertainty sources, or stochastic realisations.

The pragmatic lower-bound fleet size is a smaller fleet that achieves a very high average service level while accepting a limited number of internal service-window violations. It is useful for evaluating the trade-off between operational reliability and fleet efficiency. However, it should not be considered equivalent to the stochastic strict threshold, particularly when the lower confidence bound falls below the selected service-level target.

The technical fleet requirement consists of the operational fleet requirement plus a separately determined technical reserve. This reserve accounts for vehicles that are temporarily unavailable because of maintenance, failures, inspections, cleaning, battery degradation, charger unavailability, or other technical causes. Because these processes are not represented explicitly in the operational simulation, the technical reserve cannot be derived directly from the simulated operational fleet requirement.

The difference between deterministic and stochastic fleet requirements is primarily caused by changes in temporal vehicle availability. Operational uncertainty does not only make individual trips longer or individual tasks later. It changes when demand occurs and when vehicles become available following previous assignments. Even when the total daily workload remains unchanged, relatively small timing changes can create local demand peaks.

During these peaks, vehicles may already be servicing, travelling, charging, refilling, dumping, returning to the depot, or waiting for infrastructure. Operational slack is consequently reduced. This confirms that eGSE fleet sizing is a dynamic availability problem rather than a static comparison between total task volume and fleet size.

The task-timing uncertainty results require careful interpretation. In the model, task occurrence and the associated service window shift together when an aircraft arrives early or late. The stochastic experiment therefore evaluates whether the fleet can handle a changed temporal demand pattern under delay-adjusted service windows. It does not evaluate whether delayed tasks can still be completed within their original deterministic service windows.

This choice is reasonable because the service opportunity for many aircraft-dependent activities moves with the aircraft arrival time. However, it also means that task-timing uncertainty can occasionally improve the measured on-time completion rate by reducing overlap in the original demand pattern. This does not imply that uncertainty is operationally beneficial. It demonstrates that fleet performance depends on the temporal distribution of demand and the amount of schedule slack.

The strict 100% criterion creates a conservative fleet-size interpretation. Near the performance threshold, an additional vehicle may eliminate only a small number of short internal service-window violations. This creates a trade-off. A larger fleet improves reliability but increases investment, apron parking demand, potential charger demand, and underutilisation. A smaller fleet improves fleet efficiency but accepts residual service risk.

As discussed in the limitations, an internal service-window violation does not automatically cause an aircraft departure delay. The operational significance of residual lateness depends on its frequency, duration, and criticality. A small number of short violations may be acceptable for some vehicle groups, but this should be assessed using operational expertise and aircraft-level impact rather than the on-time completion percentage alone.

The confidence-interval results reinforce the interpretation of pragmatic lower-bound fleet sizes as boundary-case configurations. Their mean on-time completion rates may be close to the 99.90% service-level target, while their lower confidence bounds fall below this value. These configurations therefore indicate that near-perfect average performance may be possible with fewer vehicles, but they should not be interpreted as statistically strict service-level guarantees. Confidence intervals quantify the precision of the estimated mean performance; they do not provide a guarantee for all future operating conditions.

The sensitivity results further demonstrate that fleet-size outcomes depend on operational assumptions. Service duration, travel speed, depot-return rules, depot dwell time, task-selection logic, charging thresholds, and infrastructure interaction all affect when vehicles become available again. Part of the apparent fleet-capacity requirement may therefore be created by controllable operating rules rather than by unavoidable physical demand.

Idle forward staging illustrates this mechanism. It can reduce depot-centred deadheading, avoid unnecessary depot dwell, and keep vehicles closer to future demand. Its effect is greatest when performance losses are caused by avoidable depot returns during short idle periods. It is less effective when the binding constraint is simultaneous peak demand, long service duration, onboard resource capacity, or infrastructure capacity. For vehicle groups whose tasks are already spatially distributed or whose travel component is relatively small, the strict fleet-size threshold may remain unchanged even when driven distance and depot-return movements decrease substantially.

Idle forward staging should therefore not be interpreted as a universal substitute for fleet capacity. Its value depends on the underlying bottleneck and on whether forward parking is operationally feasible. The relevant conclusion is that fleet size and positioning policy should be evaluated jointly.

Feasibility-aware triage has a different objective. It prioritises tasks that can still be completed within their service windows and postpones tasks for which on-time completion is already infeasible. Under scarce capacity, this can increase the total number of on-time completions because vehicles are not assigned immediately to tasks whose deadlines can no longer be met.

However, the postponed tasks may become more severely late. The rule therefore does not necessarily minimise total lateness, average lateness, maximum lateness, or aircraft-delay risk. It primarily demonstrates that task-selection logic affects how scarce capacity and lateness are distributed. A practical dispatching policy would need to include task criticality, departure impact, maximum acceptable lateness, and fairness between tasks.

Charging was not a binding constraint for service punctuality under the assumed fleet sizes, battery capacities, energy consumption, charger power, charging thresholds, and charger availability. Energy-related parameters mainly affected charging sessions, charger occupancy, peak charger use, charging queues, and infrastructure utilisation. This does not imply that charging is generally unimportant. Fewer chargers, charger failures, higher energy consumption, cold-weather conditions, degraded batteries, increased demand, or shared infrastructure use could make charging restrictive. Fleet and charging-infrastructure decisions should therefore still be evaluated together.

Two general implications follow from these results. First, a fleet-size decision should not be based on one deterministic simulation or one isolated KPI. A defensible decision requires a combined interpretation of deterministic sufficiency, stochastic performance, marginal fleet-size improvement, late-task severity, unserved tasks, vehicle utilisation, driven distance, charging activity, charger occupancy, queueing behaviour, and infrastructure interaction.

Second, fleet size should be evaluated together with operational-control rules. Positioning can reduce avoidable inefficiency, while task-selection logic changes how lateness is distributed under scarce capacity. The preferred fleet size therefore depends on both the accepted service-level criterion and the way the fleet is controlled during daily operations.

7.4.2. Case-Study Interpretation

The interpretation of the KLM Schiphol results differs between the three vehicle groups. Water vehicles are sensitive to peak demand, travel time, service duration, refill behaviour, and depot-return logic. A pragmatic water-vehicle fleet is therefore defensible only when assumptions concerning service duration, travel time, water demand, and refill operations are sufficiently reliable.

Toilet vehicles are more stable under the tested deterministic sensitivity settings, but stochastic uncertainty still increases their capacity requirement. Their results demonstrate that deterministic sufficiency does not automatically imply adequate performance under variable task timing and travel conditions.

Loader performance shows only a limited difference between the deterministic and stochastic strict thresholds, but it is highly sensitive to service duration. This reflects the limited slack between loader service durations and the available service windows. A relatively small increase in handling time can therefore remove operational slack and create additional late tasks.

The water-vehicle baseline also illustrates why task-timing uncertainty must be interpreted carefully. Because task times and service windows shift together, the changed demand distribution can reduce local overlap in the original peak pattern. This can improve the measured on-time completion rate in some replications. The result does not indicate that uncertainty improves operations. It confirms that local concurrency and the temporal distribution of demand are important drivers of fleet performance.

Idle forward staging produces clear effects in the case study. For toilet vehicles, the deterministic fleet-size requirement decreases from eleven to nine vehicles. Under combined uncertainty, the variant also lowers the stochastic strict thresholds for water and toilet vehicles. This indicates that positioning can create operational buffer when lateness is caused partly by avoidable repositioning rather than by unavoidable peak concurrency.

However, the magnitude of this improvement is linked to the baseline depot-return logic and the assumed 15-minute depot dwell period. The observed benefit combines reduced travel, improved positioning, and avoided dwell time. The outcome should therefore be interpreted as evidence that vehicle-positioning and depot-return policies matter, not as proof that the complete simulated fleet reductions can be implemented directly without further testing.

The average-day check for water vehicles confirms that demand-day selection affects the fleet-size outcome. The lower-demand day requires fewer vehicles, showing that the peak-day analysis provides a conservative capacity basis. However, the reduction in fleet requirement is not directly proportional to the reduction in total daily tasks. Local concurrency, peak-hour demand, stand distribution, travel requirements, and available schedule slack remain important on the average day. Fleet size is therefore determined by the temporal and spatial distribution of demand rather than by total daily workload alone.

For the tested KLM Schiphol configurations, charging is not a binding cause of late task completion. This conclusion applies to the tested water, toilet, and loader configurations under the assumed battery capacities, charger availability, charging power, charging thresholds, energy consumption, and demand conditions. It should not be generalised to all eGSE operations or all future Schiphol configurations. Changes in charging infrastructure, battery condition, shared charger use, fleet composition, weather conditions, or operational demand could alter the dominant bottleneck.

Overall, the KLM Schiphol outcomes should be used as structured evidence for operational fleet planning rather than as final procurement advice. The tested thresholds show how service-level interpretation, operational uncertainty, vehicle positioning, depot-return behaviour, charging assumptions, and vehicle-specific task characteristics interact.

Before implementation, the operational fleet sizes should be tested across additional representative operating days, combined with a technical-reserve assessment, and reviewed for practical feasibility by operational experts. Additional testing should include disruption scenarios, charging-infrastructure scenarios, vehicle-availability scenarios, and alternative operational-control policies. Further validation of service times, travel times, depot dwell behaviour, charging patterns, and SOC development would strengthen the practical basis of the recommendations. Linking late internal service tasks more directly to aircraft departure punctuality would further improve the translation from simulated service performance to operational fleet-planning decisions.

8 Conclusion

The chapter is structured as follows. It first reiterates the six SQs that operationalise the central RQ, ensuring that the conclusion is explicitly anchored to the study's analytical framework. The subsequent subsections then consolidate the main findings into concise answers to SQ 1–SQ 6, moving from the operational characterisation of GSE and the required model inputs to the simulation design, uncertainty representation, performance indicators, and fleet-sizing interpretation. Building on these synthesised answers, the chapter finally addresses the RQ directly by translating the simulation outcomes into broader conclusions on how operational requirements, uncertainty, and charging constraints affect required eGSE fleet capacity and demand in airport turnaround operations.

8.1. Sub-questions (SQs)

In order to answer the RQ, the following set of SQs is formulated:

SQ 1. **Which GSE types exist and what are the functions and duty cycles of these types?**

SQ 2. **What input requirements and data sources are necessary to model an eGSE fleet while capturing operational requirements, uncertainty, and charging processes?**

SQ 3. **How can airport eGSE operations be represented in a simulation model using appropriate charging control rules, decision rules and parameters?**

SQ 4. **How can task-timing and travel-time uncertainty in eGSE operations be incorporated into a simulation model?**

SQ 5. **Which KPIs best capture the impact of eGSE implementation on operational performance and resource requirements?**

SQ 6. **How can the required eGSE fleet capacity be determined from simulation outcomes, and how do demand conditions, operational positioning, and task-selection rules affect service performance?**

Together, the SQs connect the literature review, simulation-framework design, case-study implementation, validation, simulation experiments, and final fleet-sizing interpretation. They therefore support the RQ by moving from conceptual model requirements to operational and stochastic fleet-size outcomes.

8.1.1. Sub-question (SQ) 1

GSE consists of the vehicles and equipment required to support aircraft during ground handling operations. These vehicles are not one homogeneous group, but can be classified according to their function in the turnaround process and according to their operational duty cycle. The required GSE configuration depends on the aircraft type, stand configuration, flight type, airline procedures, and the specific services required during the turnaround.

From a functional perspective, GSE can be divided into several main groups. First, aircraft support equipment provides technical support to the aircraft while it is parked at the stand. Examples are Ground Power Unit (GPU), Air Starter Unit (ASU), and Pre-Conditioned Air Unit (PCA). These units supply electricity, air, or conditioned air to the aircraft and are usually connected to the aircraft for part of the turnaround process. Their duty cycle is therefore mainly stand-based and time-dependent: they remain at or near the aircraft while the service is required.

Second, aircraft movement equipment is used to move or reposition aircraft. This includes push-back tractors, and towing vehicles. These vehicles perform short but operationally critical tasks, often directly linked to aircraft departure or repositioning. Their duty cycle is characterised by relatively short service events, high power demand, strict timing, and strong dependence on the flight schedule.

Third, aircraft servicing equipment performs physical service operations on the aircraft. This category includes potable-water vehicles, toilet or lavatory-service vehicles, fuel trucks, catering vehicles, de-icing vehicles, and cleaning-related support equipment. These vehicles are usually stand-based during service execution, but their duty cycles differ by resource requirement. Water vehicles need to refill the aircraft and may need to visit a water refill point. Toilet vehicles empty the lavatory system and may need to visit a dump point. Catering and fuel vehicles also depend on resource quantities and may require replenishment between tasks. These vehicles therefore often follow a resource-limited duty cycle, where the number of aircraft that can be served depends not only on time availability, but also on onboard capacity.

Fourth, passenger handling equipment supports passenger boarding and disembarking. Examples are passenger stairs, passenger buses, and passenger lift vehicles. Their function depends strongly on the stand type. At contact stands, passengers may board through a passenger boarding bridge, while remote stands often require stairs and buses. The duty cycle of passenger handling equipment is therefore strongly connected to passenger flows, boarding procedures, and the distinction between terminal and remote stands.

Fifth, baggage and cargo handling equipment is used to load, unload, and transport baggage, cargo, and unit load devices. This category includes belt loaders, lower-deck loaders, baggage tractors, baggage carts, cargo transporters, and loader-transporters. Some of these vehicles, such as loaders and belt loaders, mainly operate at the aircraft stand during loading or unloading. Others, such as baggage tractors and transporters, follow repeated movements between the aircraft stand and baggage or cargo handling areas. Their duty cycle is therefore often round-trip based and depends on baggage volume, aircraft type, and the timing of loading and unloading activities.

From an operational perspective, the duty cycles of GSE can be summarised in three broad working modes. The first is a continuous working mode, where vehicles can move from one aircraft stand to the next without needing an intermediate resource stop. Examples include towing tractors, stairs, transporters, and loaders. The second is a resource-limited working mode, where vehicles can only serve a limited number of aircraft before they must refill, dump, or replenish. Water, toilet, catering, and fuel vehicles are typical examples. The third is a round-trip working mode, where vehicles repeatedly travel between the aircraft stand and another operational location, such as a terminal, baggage hall, cargo area, depot, or service point. Passenger buses, and baggage tractors are examples of this mode.

The answer to SQ1 is therefore that GSE types differ both in function and in operational behaviour. Functionally, they can be grouped into aircraft support, aircraft movement, aircraft servicing, passenger handling, and baggage or cargo handling equipment. Operationally, their duty cycles can be described as stand-based, continuous, resource-limited, or round-trip based. This distinction is important for eGSE modelling because each duty cycle creates different requirements for fleet size, travel time, service timing, battery capacity, charging opportunities, and supporting infrastructure. In this thesis, the case-study model focuses on water vehicles, toilet vehicles, and loaders, but the same modelling logic can be extended to other GSE types when their task rules, resource requirements, movement patterns, and charging characteristics are available. In the case-study application, water vehicles, toilet vehicles, and loaders are selected as representative examples of resource-limited and stand-side GSE operations.

8.1.2. Sub-question (SQ) 2

To model an eGSE fleet in a realistic way, several input layers are required. These inputs must describe the demand for ground handling tasks, the operational rules that translate this demand into service tasks, the spatial structure of the airport, the characteristics of the vehicles, the available infrastructure, and the uncertainty affecting daily operations. Together, these data sources determine when tasks occur, where they take place, which vehicles can perform them, and whether the available fleet and charging infrastructure are sufficient.

The first required input is a flight schedule. This forms the basis for operational demand. The flight schedule should include scheduled arrival and departure times, aircraft stand, aircraft subtype, operational classification, and information on connected flights. These data are needed to determine when and where service demand occurs. For turnaround operations, connected arrivals and departures must be treated carefully to avoid double-counting the same aircraft service demand. After preprocessing, each relevant flight movement or turnaround can be translated into one or more service tasks.

The second input layer consists of service data. These data define how flight-level demand is converted into operational tasks. For each combination of aircraft subtype, operation type, and service process, the service data should specify the nominal service duration and the time-window offsets relative to the scheduled flight time. This is necessary because not all aircraft require the same service duration or the same service window. The service data also distinguish between different service types, such as water service, toilet service, and loader operations. These rules are essential for modelling operational requirements at the task level.

The third required input concerns aircraft and service characteristics. Aircraft subtype and body category are needed to approximate service requirements, such as water volume, waste volume, or loading-related work. Larger aircraft generally require more service time and larger resource quantities than smaller aircraft. Therefore, a mapping between aircraft subtype, aircraft class, service duration, and resource requirement is needed. This allows the model to represent differences between commuter, narrow-body, and wide-body aircraft without modelling every aircraft in full technical detail.

The fourth input layer is the spatial representation of the airport. The model needs information on aircraft stands, depot locations, charging locations, refill points, dump points, and the travel distances between these locations. Because a full microscopic traffic model is often too detailed for fleet-sizing analysis, stands can be grouped into location areas and connected through a distance matrix. Travel times can then be calculated from these distances using vehicle-specific speeds. This spatial input is necessary because travel time directly affects vehicle availability, task feasibility, and service punctuality.

The fifth required input describes the eGSE fleet itself. For each vehicle type, the model needs the number of vehicles, starting location, vehicle compatibility, operational speed, battery capacity, energy consumption, charging behaviour, and initial state of charge. For resource-limited vehicles, additional onboard capacity parameters are required. Water vehicles need water tank capacity and refill thresholds, while toilet vehicles need waste tank capacity and dumping thresholds. These parameters determine whether a vehicle can execute a task, whether it must first recharge, refill, or dump, and when it becomes available again.

The sixth input layer concerns supporting infrastructure. Charging stations, refill points, and dump points must be represented with their locations and capacities. For charging infrastructure, the model requires charger power, number of chargers, charging efficiency, and charging thresholds. For refill and dump infrastructure, the model requires service durations and resource capacities. These inputs are important because infrastructure is shared between vehicles and can therefore create queues. Limited infrastructure capacity can reduce effective fleet availability, even if enough vehicles are present.

The seventh input layer consists of operational policies and control parameters. These include task release time, dispatching rules, task prioritisation logic, vehicle selection rules, charging thresholds, depot-return logic, refill and dump rules, and stop or end-of-shift conditions. These parameters determine how the model behaves during simulation. They are necessary because the same demand pattern and fleet size can lead to different outcomes depending on how vehicles are dispatched, when they are allowed to charge, and how supporting resource constraints are handled.

The eighth input layer concerns uncertainty. To capture operational variability, the model requires data or assumptions for stochastic task timing and travel-time variability. In this thesis, arrival-time uncertainty is derived from historical flight data and segmented by time of day. Delay magnitudes are sampled from capped empirical distributions. Travel-time uncertainty is represented through a bounded stochastic disturbance applied to individual vehicle movements. Random seeds are used to make simulation runs reproducible. These uncertainty inputs allow the model to evaluate whether the fleet remains robust when demand timing and vehicle availability differ from the deterministic baseline.

Finally, validation data are required to assess whether the model outputs are plausible. These data can include operational task counts, fleet availability, measured or estimated driven distances, energy consumption data, charging behaviour, water demand, and expert judgement from operational stakeholders. Validation data do not only check whether the model produces reasonable outputs, but also help identify which assumptions are reliable and which assumptions remain uncertain.

The answer to SQ2 is therefore that an eGSE fleet model requires more than a flight schedule alone. A realistic model needs integrated input data on demand, service rules, aircraft characteristics, airport layout, vehicle properties, charging and service infrastructure, operational control policies, uncertainty profiles, and validation references. These inputs allow the simulation to represent the interaction between task timing, vehicle availability, energy constraints, charging behaviour, and operational uncertainty. Without this combination of input layers, the model would not be able to evaluate fleet capacity and service reliability under realistic electrified ground handling conditions.

8.1.3. Sub-question (SQ) 3

Airport eGSE operations can be represented as a rule-based discrete-event simulation in which service tasks, individual vehicles, and supporting infrastructure interact over time. This modelling approach is suitable because ground handling operations consist of time-dependent events, such as task releases, vehicle dispatches, vehicle arrivals, service starts, service completions, charging sessions, refilling actions, and dumping actions. By representing these events explicitly, the model can evaluate how vehicle availability, travel time, service windows, battery state, and infrastructure constraints jointly affect operational performance.

The model structure consists of three main entity types: tasks, vehicles, and infrastructure. Tasks represent realised service demand at aircraft stands. Each task is defined by a service type, aircraft stand, service duration, time-window opening, time-window closing, and, where relevant, a resource requirement. Vehicles represent individual eGSE units and are described by their current location, operational status, battery state of charge, onboard resource level, and compatibility with specific task types. Supporting infrastructure consists of chargers, refill points, dump points, depots, and other service locations. These infrastructure elements are modelled as capacity-constrained resources, meaning that vehicles may have to wait when several vehicles require the same facility at the same time.

The decision structure of the simulation is organised around four main components: the dispatcher, task selection, vehicle selection, and task execution. The dispatcher acts as the central coordination mechanism. It monitors released tasks and available vehicles and assigns tasks when this is operationally feasible. Tasks are not necessarily assigned immediately after they enter the system. If no suitable vehicle is available, the task remains in the task queue and is reconsidered in a later dispatching round. This makes the model adaptive, because assignments depend on the current state of the system rather than on a fixed pre-planned schedule.

Task selection determines which waiting task should be prioritised. In this thesis, tasks are ranked according to their remaining slack before the latest feasible service start. Slack is defined as the time remaining until the task must start in order to finish before the service-window closing time. Tasks with less slack are more urgent and are therefore considered first. If two tasks have equal slack, the task with the longer minimum travel time is prioritised, because it requires more lead time before service can begin. This rule supports time-window feasibility by giving priority to tasks that are both urgent and less accessible.

Vehicle selection determines which available vehicle should execute the selected task. The model first constructs a candidate set of vehicles that are operationally available and compatible with the task. Vehicles can be excluded if they have insufficient battery charge, insufficient onboard resource capacity, are returning to the depot, are in a stop-cutoff state, or violate service-specific constraints. For each eligible candidate, the model checks whether the vehicle can travel to the task location, wait if necessary, perform the service, and complete the task before the service-window closing time. If one or more vehicles can complete the task on time, the vehicle with the shortest travel time is selected. If no vehicle can complete the task within the service window, the model selects the vehicle that minimises expected lateness. This fallback rule allows the simulation to represent constrained operational situations in which late service is still performed when no on-time option exists.

Task execution is represented as a sequence of operational events. After a vehicle is assigned to a task, it travels to the aircraft stand. If it arrives before the service window opens, it waits until service is allowed to start. The service then begins and continues for the task-specific service duration. After completion, the vehicle state is updated. This includes its location, battery state of charge, onboard resource level, and operational status. The vehicle can then be assigned to a new task or can be sent to charge, refill, dump, or return to the depot, depending on its state and the operational rules.

Charging control is modelled explicitly because electrification makes vehicle availability dependent on battery state and charging infrastructure. Vehicles are allowed or required to charge based on state-of-charge thresholds, location, and expected operational need. A lower battery threshold prevents vehicles from being assigned when their remaining charge is too low for safe operation. Depot-based charging can be triggered when an idle vehicle is below a higher charging threshold. The model can also include proactive idle charging, where vehicles charge during idle periods if no urgent tasks are expected. Charging is constrained by charger location and charger capacity, which means that vehicles may queue before charging can start. Charging can also be interrupted when a task becomes operationally more important than continuing the charging session.

For resource-limited vehicles, additional control rules are required. Water vehicles must monitor their remaining onboard water level and visit a refill point when the level falls below a threshold or when end-of-shift rules require the vehicle to return in a filled state. Toilet vehicles must monitor their waste tank level and visit a dump point when the tank exceeds a defined threshold or before shift completion. These rules ensure that resource constraints are represented as operational limitations rather than only as accounting variables. They also allow the model to capture queueing at refill and dump locations.

The model is controlled by a set of operational parameters. These include fleet size, vehicle speed, service duration, task release offset, battery capacity, energy consumption, charging power, charging thresholds, charger capacity, refill and dump duration, onboard tank capacity, refill and dump thresholds, depot-return rules, and uncertainty settings. By changing these parameters between experiments, the model can evaluate how operational decisions and infrastructure assumptions affect performance. This makes the simulation suitable for fleet-sizing analysis, charging strategy evaluation, and sensitivity analysis.

The answer to SQ3 is therefore that airport eGSE operations can be represented by modelling the system as a dynamic interaction between time-window-based service tasks, individual battery-electric vehicles, and capacity-constrained infrastructure. Appropriate decision rules are needed for task prioritisation, vehicle assignment, task execution, charging, refilling, dumping, and depot return. Charging control rules must be linked to battery state, charger availability, operational urgency, and idle time. This representation allows the simulation to capture the main operational mechanisms through which electrification affects ground handling performance: limited vehicle availability, charging downtime, infrastructure competition, and reduced flexibility during demand peaks. In this thesis, this representation is implemented for water vehicles, toilet vehicles, and loaders; extension to other GSE types requires vehicle-specific rules and renewed validation.

8.1.4. Sub-question (SQ) 4

Uncertainties in eGSE operations can be incorporated by separating stochastic operational variability from uncertain model parameters. Not every uncertainty source has to be modelled as a random process. Some uncertainties directly affect the timing and availability of vehicles during a simulation run and are therefore suitable for stochastic modelling. Other uncertainties are better represented through deterministic assumptions, scenario analysis, sensitivity analysis, or validation checks. This distinction keeps the model interpretable while still capturing the main operational effects of uncertainty.

In airport ground handling, uncertainty can arise from several sources. Flight timing uncertainty affects when aircraft-dependent service tasks can start. Task demand uncertainty affects whether a service is required and how much work is needed. Service duration uncertainty affects how long a vehicle remains occupied at an aircraft stand. Travel time uncertainty affects when vehicles arrive at task locations and when they become available for later tasks. Energy and charging uncertainty affects whether vehicles have enough battery capacity and whether charging infrastructure is available when needed. These uncertainty sources interact with each other. For example, a delayed task can reduce idle time, postpone charging, increase charger competition, and reduce vehicle availability later in the day.

In this thesis, the active stochastic uncertainty scope is limited to task timing uncertainty and travel time uncertainty. These two uncertainty sources are selected because they directly influence the timing of service demand and the availability of vehicles, which are central to fleet sizing. Task cancellations are not included as an active stochastic process, because the aim is to evaluate fleet performance under maximum demand conditions. Including cancellations would reduce the effective workload and could therefore underestimate the required fleet capacity.

Task timing uncertainty is incorporated by applying stochastic delays to planned service tasks. For each task, a delay occurrence is first determined using a Bernoulli trial. The delay probability depends on the task profile and the time-of-day bucket. If a delay occurs, the delay magnitude is sampled from an empirical delay distribution derived from historical operational data. These empirical delay magnitudes are capped to reduce the influence of extreme outliers or timestamp errors. The sampled delay is then rounded to whole minutes and applied to the task timing. In this model, the task timing and the corresponding service window are shifted together. This means that task uncertainty mainly changes the temporal distribution of demand, while the relative duration of the service window remains unchanged.

Travel time uncertainty is incorporated during simulation execution. The deterministic travel time is first calculated from the distance matrix and vehicle speed. Then, for each vehicle movement, a stochastic disturbance may occur. If a travel disturbance is triggered, the deterministic travel time is increased by a predefined multiplicative factor. This represents incidental apron delays such as congestion, temporary blockages, routing inefficiencies, waiting time, or interactions with other airside traffic. Because travel uncertainty is applied dynamically, it directly affects vehicle availability: a longer movement keeps a vehicle occupied for longer and can influence later dispatching decisions.

The stochastic uncertainty mechanisms are controlled through random seeds. This makes the simulation reproducible. Running the same experiment with the same seed gives the same uncertainty realisation, while running the same experiment with different seeds gives different realisations of the stochastic processes. This allows the model to be used for Monte Carlo-style experiments, where the same fleet size and parameter setting are tested over multiple simulation runs. The output can then be analysed not only as a single value, but as a distribution of possible operational outcomes.

Other uncertainty sources are not ignored, but are treated differently. Service durations, task occurrence rules, water demand, waste quantities, energy consumption, battery capacity, charging power, charger availability, refill capacity, and dump capacity are represented as model parameters. Their influence can be tested through scenario analysis or sensitivity analysis. For example, changing travel speed, service duration, charging power, or battery-related parameters allows the model to evaluate whether the fleet-sizing results are robust to these assumptions. This approach separates short-term stochastic variability from longer-term parameter uncertainty.

Charging-related uncertainty is incorporated mainly through the interaction between stochastic task timing, stochastic travel time, and deterministic charging constraints. The model explicitly represents battery state of charge, charging thresholds, charger locations, charger capacity, and charging duration. These elements are not sampled randomly in each run, but they constrain vehicle availability. When tasks are delayed or travel times increase, idle periods can become shorter, charging can be postponed, and charger queues can occur. In this way, uncertainty can still affect charging behaviour indirectly through the dynamic operation of the energy-constrained fleet.

The answer to SQ4 is therefore that uncertainties in eGSE operations can be incorporated by combining stochastic simulation with scenario and sensitivity analysis. In this thesis, task timing uncertainty and travel time uncertainty are modelled stochastically because they directly affect demand timing and vehicle availability. Other uncertainties are represented as deterministic model parameters and tested through controlled experiment variations. This approach captures the main operational mechanisms through which uncertainty affects eGSE fleet performance, while keeping the simulation model transparent, reproducible, and suitable for fleet-sizing analysis.

8.1.5. Sub-question (SQ) 5

The impact of eGSE implementation cannot be captured by one single KPI. Electrification affects both operational service performance and the use of vehicles, energy, and infrastructure. Therefore, a useful KPI set must include indicators for service reliability, fleet efficiency, vehicle movement, and charging performance. Together, these indicators make it possible to evaluate whether the fleet can complete the required tasks on time, whether the available vehicles are used efficiently, and whether charging infrastructure creates additional operational constraints.

The most important KPI is the on-time task completion rate. This metric measures the percentage of tasks that are completed before the end of their required service time window. It is the primary service-performance indicator because it directly shows whether the available fleet can support the operational schedule. For fleet-sizing analysis, this KPI is especially important because a fleet configuration can only be considered fully sufficient under a strict service-level requirement if all tasks are completed within their service windows.

However, the on-time task completion rate should not be interpreted in isolation. It should be combined with the unserved task rate and the delay characteristics of late tasks. The unserved task rate measures whether tasks remain uncompleted at the end of the simulation. This is important because a task that is completed late and a task that is not completed at all have different operational implications. The average delay of late tasks, the number of late tasks, and the median delay provide additional information on the severity of service failures. An experiment with a small number of tasks that are only one or two minutes late is operationally different from an experiment with many tasks that are substantially delayed.

Vehicle-related KPIs are also needed to understand resource requirements. The overall vehicle utilisation rate indicates how intensively the fleet is used during the simulation period. In addition, operational time shares provide more detail by separating vehicle time into service time, travel time, charging time, depot idle time, and waiting time. These indicators show whether vehicles are mainly occupied by productive service work, movement between locations, charging, or idle time. This distinction is useful because low utilisation does not automatically mean that the fleet is oversized. In airport operations, vehicles may be idle during quiet periods but still be necessary during short demand peaks.

Movement-related KPIs are required because GSE performance depends not only on service time at the aircraft stand, but also on travel between stands, depots, charging points, refill points, and dump points. The total driven distance measures the total operational movement of the fleet. The dispatch-to-task distance measures the distance travelled specifically to reach assigned tasks. These metrics help to evaluate the spatial efficiency of the dispatching logic and provide insight into the operational effort required to perform the task schedule. They are also relevant for estimating energy demand, vehicle wear, and the realism of simulated operations.

Charging-related KPIs are necessary because electrification makes vehicle availability dependent on battery state and charging infrastructure. Relevant charging indicators include the number of charging sessions, average charger occupancy, peak charger occupancy, immediate charging start rate, time with a charging queue, average charging queue length, maximum charging queue length, average charging waiting time, and maximum charging waiting time. These KPIs show whether charging infrastructure is used intensively, whether vehicles must wait before charging, and whether charging creates a bottleneck for operational performance.

For resource-limited vehicle types, additional process-specific indicators may be needed. For water vehicles, water demand, refill frequency, and refill-point use are relevant. For toilet vehicles, waste capacity, dumping frequency, and dump-point use are relevant. These indicators are important because such vehicles are not only constrained by time and battery energy, but also by onboard resource capacity. A vehicle may be available in time and have sufficient battery charge, but still be unable to perform a task if it must first refill or dump.

The answer to SQ5 is therefore that the best KPI set combines a primary service-performance metric with supporting operational and infrastructure metrics. The on-time task completion rate is the central KPI for fleet sizing, because it directly measures whether tasks are completed within their required service windows. However, it must be supported by the unserved task rate, late-task delay characteristics, vehicle utilisation, operational time shares, driven distance, charging sessions, charger occupancy, charging queues, and vehicle-type-specific resource indicators. This combination provides a more complete view of the impact of eGSE implementation on both operational performance and resource requirements. These KPIs should be interpreted as indicators of GSE service reliability and resource adequacy, not as direct measures of aircraft departure performance.

8.1.6. Sub-question (SQ) 6

The required eGSE fleet capacity can be determined by testing different fleet sizes in a simulation model and evaluating each configuration with performance-based KPIs. The main criterion used in this thesis is the on-time task completion rate. A fleet size is considered sufficient under a strict service-level interpretation when it is the smallest tested configuration that completes 100% of tasks within their service windows. This criterion provides a clear and conservative definition of the required operational fleet size.

Scenario analysis is used to determine whether this fleet-size requirement changes under different operating conditions. First, deterministic fleet-sizing experiments are conducted without task-timing uncertainty and travel-time uncertainty. These experiments show the minimum fleet capacity required under nominal operating conditions. Second, stochastic uncertainty experiments are conducted with task-timing uncertainty, travel-time uncertainty, and combined uncertainty. These experiments show whether the deterministic fleet remains sufficient when operational variability is included. Third, a demand-case robustness check can replace the complete operational input day. This is not treated as a standard parameter sensitivity, because it changes the full demand pattern rather than one isolated assumption. Fourth, operational positioning policies, such as idle forward staging, can be evaluated. Fifth, task-selection variants, such as feasibility-aware triage, can be evaluated. Finally, parameter sensitivity analysis can be used to test how selected assumptions, such as service duration, travel speed, battery capacity, energy consumption, and charging speed, affect the robustness of the outcomes.

The fleet-size decision should not be based only on the strict 100% threshold. Near the highest-performing fleet sizes, the marginal improvement of adding one extra vehicle becomes small. Therefore, a more pragmatic interpretation can also be considered. If a very small number of short internal service-window violations is operationally acceptable, lower fleet sizes may be defensible. These lower-bound configurations can achieve very high on-time completion rates, but they do not satisfy the strict 100% criterion.

Demand conditions affect fleet-size requirements. A lower-demand day may require fewer vehicles than the baseline day under strict deterministic or stochastic criteria. However, the relationship is not purely proportional to total task volume. Temporal clustering, peak simultaneous demand, service-window structure, stand distribution, travel requirements, and uncertainty remain relevant. Therefore, fleet sizing should consider not only the number of daily tasks, but also the timing and spatial distribution of these tasks.

Operational positioning also affects service performance. Idle forward staging shows that fleet-capacity requirements can be influenced by where idle vehicles are positioned. By keeping idle vehicles near the VOP when battery state and future demand allow this, unnecessary depot-return movements can be reduced. This can lower driven distance and depot returns and, for selected vehicle groups, reduce fleet-size thresholds. The policy is therefore a practical operational contribution, but not a universal substitute for capacity. It is most effective when lateness is caused by avoidable depot-return movements rather than by unavoidable peak concurrency or travel-time variance.

Task-selection rules affect the distribution of lateness. Feasibility-aware triage can reduce the number of late tasks under scarce-fleet conditions by prioritising tasks that can still be completed on time. However, it may increase the average lateness of the remaining late tasks, because already-infeasible tasks are postponed. It is therefore not a general operational improvement or a direct fleet-size reduction measure. Its value is diagnostic: it shows that the chosen task-selection rule influences whether lateness is spread across many tasks or concentrated in fewer tasks.

The general answer to SQ6 is therefore that required eGSE fleet capacity can be determined by combining fleet-size experiments, scenario analysis, and performance-based KPIs, while explicitly distinguishing strict and pragmatic service-level interpretations. Simulation-based fleet sizing should report not only fleet-size thresholds, but also the demand case, service-level interpretation, operational control assumptions, lateness severity, vehicle utilisation, charging behaviour, and spatial-efficiency indicators that explain why a fleet size is or is not sufficient.

Case-Study Application

For the KLM Schiphol case study, the deterministic experiments show that the minimum fleet sizes required to achieve 100% on-time task completion are 12 water vehicles, 11 toilet vehicles, and 23 loaders. These fleet sizes are sufficient when the system operates under nominal conditions and no stochastic task-timing or travel-time uncertainty is included. When combined task-timing and travel-time uncertainty is included, the strict fleet-size requirements increase to 16 water vehicles, 15 toilet vehicles, and 25 loaders. The difference between deterministic and stochastic fleet-size requirements shows that uncertainty can increase the required fleet capacity, even when the total daily number of tasks remains unchanged.

Under a more pragmatic interpretation, lower fleet sizes may be defensible if a very small number of short internal service-window violations is operationally acceptable. For the KLM Schiphol case study, the lower-bound configurations are at least 11 water vehicles, 9 toilet vehicles, and 23 loaders. These configurations achieve very high on-time completion rates, but they do not satisfy the strict 100% criterion.

The demand-case robustness check shows that required fleet capacity also depends on the operating day. For water vehicles, the average-demand day requires fewer vehicles than the baseline day under the strict deterministic and stochastic fleet-sizing criteria. However, the pragmatic lower-bound interpretation remains unchanged at 11 water vehicles, because both days still satisfy the minimum service-level threshold at this fleet size. The result is not purely proportional to total task volume: temporal clustering, peak simultaneous demand, service-window structure, and uncertainty remain relevant. The average-day check therefore supports interpreting the baseline results as conservative, while also showing that uncertainty does not disappear under lower demand.

Idle forward staging reduces driven distance and depot returns for all vehicle groups and reduces fleet-size thresholds for selected vehicle groups. In particular, toilet vehicles require fewer vehicles under both deterministic and combined-uncertainty interpretations, while water vehicles show improved robustness under combined uncertainty. This shows that operational positioning can be an important lever in the KLM Schiphol case, especially when lateness is partly caused by avoidable depot-return movements.

The feasibility-aware triage variant shows that task-selection rules affect the distribution of lateness in the KLM Schiphol case study. The variant can reduce the number of late tasks under scarce-fleet conditions, but it may also increase the average lateness of the remaining late tasks. Therefore, it should be interpreted as a diagnostic sensitivity, not as a direct operational recommendation.

For the KLM Schiphol case, the answer to SQ6 is therefore that required fleet capacity depends on the service-level interpretation, the operating day, the uncertainty setting, the positioning policy, and the task-selection logic. The case-study results are 12 water vehicles, 11 toilet vehicles, and 23 loaders under deterministic conditions, and 16 water vehicles, 15 toilet vehicles, and 25 loaders under combined operational uncertainty. A more pragmatic interpretation gives lower-bound fleet sizes of at least 11 water vehicles, 9 toilet vehicles, and 23 loaders, provided that a very small number of short internal service-window violations is acceptable.

8.2. Research Question (RQ)

This section answers the central RQ by translating the simulation results into broader implications for eGSE fleet capacity and operational demand in airport turnaround processes.

What is the impact on the required fleet capacity and operational demand of electric Ground Support Equipment (eGSE) when simulating operational requirements, uncertainty and charging in turnaround processes at airports?

In this thesis, demand refers to operational eGSE demand: the time- and location-specific demand for vehicle availability, task execution, charging, and supporting infrastructure.

The main impact of operational requirements, uncertainty, and charging is that eGSE fleet capacity can no longer be interpreted as a static ratio between the number of vehicles and the number of daily tasks. Instead, fleet capacity becomes a dynamic operational requirement. It depends on the timing of tasks, their spatial distribution across stands, service durations, travel times, vehicle availability, charging opportunities, and the amount of operational slack between tasks. A simulation-based approach therefore gives a more realistic estimate of required fleet capacity than a deterministic or aggregate task-count-based assessment.

Simulating operational requirements first changes how eGSE demand is represented. Instead of only counting how many services are needed during a day, the model translates flight schedules into time-window-based service tasks at specific aircraft stands. This makes it possible to identify when and where demand is concentrated, which vehicles are available, and whether service tasks can be completed within operational time windows. The required fleet size is therefore driven primarily by peak-period capacity and vehicle availability, not by average daily demand.

Adding operational uncertainty further changes the required capacity. Even when the total number of tasks remains unchanged, uncertainty can shift tasks into busier periods, reduce slack between consecutive tasks, increase travel or waiting time, and keep vehicles occupied for longer. As a result, deterministic fleet-size estimates can underestimate the capacity needed for robust operation. The broader mechanism is that uncertainty increases the need for buffer capacity because the fleet must remain reliable under variable operating conditions. However, the effect of uncertainty on measured performance can be non-monotonic. Task-timing uncertainty may either worsen or improve the on-time completion rate depending on whether the sampled shifts increase or reduce peak overlap.

Charging introduces an additional operational constraint because electric vehicles must remain available not only in terms of task assignment, but also in terms of battery state and access to charging infrastructure. Charging can affect operational demand by changing charging-session timing, charger occupancy, peak charger use, and infrastructure utilisation. More generally, charging should therefore be interpreted as an infrastructure and availability constraint. It may become binding under specific charger configurations, lower charger power, colder weather, battery degradation, higher demand, or stricter charging strategies.

The results also show that fleet-size recommendations depend strongly on the chosen service-level criterion. If 100% on-time task completion is treated as a hard requirement, additional vehicles are needed to guarantee performance across stochastic operating conditions. If a very small number of short internal service-window violations is operationally acceptable, lower fleet sizes may also be defensible. This distinction is important because an internal service-window violation does not automatically imply an aircraft departure delay, although repeated or poorly timed violations can still reduce turnaround robustness. The confidence-interval analysis shows that pragmatic lower-bound configurations should be interpreted as operationally informative lower bounds, not as statistical guarantees.

The main implication is that eGSE fleet capacity should be interpreted as an operational robustness requirement rather than as a fixed vehicle-to-task ratio. Deterministic fleet sizes indicate nominal sufficiency under fixed assumptions, whereas stochastic fleet sizes indicate the operational buffer needed to maintain service performance under variability. Pragmatic lower-bound fleet sizes provide a possible efficiency-oriented alternative, but only if small and short internal service-window violations are considered acceptable. Therefore, the final fleet-size decision should explicitly state which service-level interpretation is used.

Overall, the impact of operational requirements, uncertainty, charging, demand conditions, operational positioning, and task-selection logic is that required eGSE fleet capacity becomes more context-dependent and more sensitive to operational robustness criteria than deterministic analysis alone suggests. For airport turnaround processes in general, eGSE fleet sizing should be based on experiment-based simulation, time-window performance, vehicle availability, charging-infrastructure use, and explicit service-level assumptions. This allows decision-makers to distinguish between nominal fleet sufficiency, robust fleet capacity, and pragmatic lower-bound alternatives.

Case-Study Application

The KLM Schiphol case study demonstrates the general mechanism with concrete fleet-sizing outcomes for water vehicles, toilet vehicles, and loaders. When combined task-timing and travel-time uncertainty is included, the strict fleet-size requirements increase from 12 to 16 water vehicles, from 11 to 15 toilet vehicles, and from 23 to 25 loaders. These exact values are case-specific, but they show that deterministic fleet-size estimates can underestimate the capacity needed for robust operation.

In the tested KLM Schiphol experiments, charging is not the dominant driver of late task completion. The fleet-size increase is mainly caused by vehicle availability during demand peaks rather than by charging queues or insufficient battery capacity. However, charging remains relevant for operational demand because it affects charging-session timing, charger occupancy, peak charger use, and infrastructure utilisation. This means that charging should still be included in the model, even when it is not the binding constraint in the tested configuration.

The distinction between strict and pragmatic service-level interpretations is also visible in the KLM Schiphol case. The strict stochastic fleet sizes are 16 water vehicles, 15 toilet vehicles, and 25 loaders. The pragmatic lower-bound configurations are 11 water vehicles, 9 toilet vehicles, and 23 loaders. These lower-bound configurations achieve mean on-time completion rates around or above 99.90%, but their lower confidence bounds fall below this threshold. They should therefore be interpreted as operationally informative lower bounds, not as strict service-level guarantees.

The KLM Schiphol case study also shows that operational control assumptions matter. Idle forward staging can reduce unnecessary depot-return movements and improve robustness for selected vehicle groups. Feasibility-aware triage changes how lateness is distributed across tasks under scarce capacity. These results show that fleet size should not be evaluated in isolation. It should be evaluated together with positioning policy, task-selection logic, charging assumptions, and the accepted service-level interpretation.

The case-study results should therefore be interpreted as validated evidence within the tested scope, not as universal fleet-size values. The broader conclusion is that eGSE fleet sizing requires dynamic simulation of operational demand, uncertainty, charging, and control rules. The KLM Schiphol case provides concrete evidence for this conclusion, while the exact numerical fleet-size thresholds remain specific to the tested vehicle groups, operating days, assumptions, and validation basis.

8.3. Recommendations

The main recommendation is that eGSE fleet sizing should be treated as a simulation-based decision-support process rather than as a one-time vehicle-count calculation. Fleet-size decisions should not be based only on deterministic schedules, average task volumes, or simple vehicle-to-task ratios, because these approaches do not capture the interaction between service time windows, peak simultaneity, vehicle availability, charging, spatial movement, and operational uncertainty. Fleet sizing should therefore start with an explicit service-level definition. Operational fleet requirements, technical reserves, charging infrastructure, and operational-control measures can subsequently be assessed as connected but distinct decision components.

Define the service-level requirement before sizing the fleet. Airport operators and ground handlers should first determine how service performance is interpreted. A strict interpretation requires 100% on-time task completion under the tested stochastic conditions. A pragmatic interpretation allows a very small number of short internal service-window violations, but only when these violations are operationally acceptable and do not affect aircraft departure punctuality. This choice should be made before simulation outcomes are translated into fleet-size decisions, because the selected service-level interpretation can materially affect the required fleet size.

Separate operational fleet requirements from procurement decisions. The simulated fleet sizes should be interpreted as operational fleet-capacity requirements under the tested assumptions, not as final procurement numbers. Final procurement should follow a two-step approach. First, the operational fleet requirement should be determined through simulation. Second, a technical reserve should be added for maintenance, failures, battery degradation, charger unavailability, cleaning, inspections, and other sources of temporary vehicle downtime. This reserve should be determined separately for each vehicle group, because operational criticality, redundancy options, expected technical availability, fleet size, access to backup vehicles, and downtime patterns differ between vehicle types.

Operational-control improvements should not be used as direct justification for reducing procurement targets before they have been validated in practice. Instead, they provide evidence on how operational robustness can be improved, how the available eGSE fleet can be used more effectively, and which risks require additional fleet or infrastructure capacity.

Evaluate fleet capacity across multiple operating days. Fleet-sizing decisions should not be based on a single operating day. A selected baseline day can function as a conservative stress case, but it cannot represent the full range of operational conditions. Fleet-size conclusions should therefore be tested across a representative portfolio of operating days, including baseline or peak days, average days, weekday and weekend patterns, seasonal variants, winter operations, holiday periods, and disruption-prone days. This allows decision-makers to distinguish structural fleet-capacity requirements from day-specific demand effects and to assess whether the derived fleet-size thresholds remain stable under different demand patterns, aircraft mixes, and operational peak structures.

Prioritise data collection on operational time parameters. Sensitivity analysis should guide further data collection and model validation. Priority should be given to parameters that directly affect task feasibility and vehicle availability, particularly service durations, apron travel times, travel speeds, waiting times, depot dwell times, depot-return behaviour, task-window definitions, and vehicle handover times. More detailed measurement of these parameters would improve the reliability of final fleet-size recommendations and reduce the risk of under- or overestimating the required fleet capacity.

Relevant data can be obtained from vehicle telematics, operational logs, location data, charger data, and dispatcher records. Service-resource processes, vehicle utilisation, charging behaviour, and SOC development should also be validated further. Where detailed operational data remain unavailable, assumptions and simulated behaviour should be reviewed with operational experts. Energy-related parameters should continue to be monitored, but operational time parameters should be prioritised when they have the strongest effect on task punctuality.

Improve operational execution before increasing fleet size. Late task completion can result from peak simultaneity, long service durations, low travel speeds, depot-return behaviour, depot dwell time, and limited vehicle availability during demand peaks. Operational robustness should therefore not automatically be increased by adding vehicles. Airport operators and ground handlers should first investigate measures that make vehicles available more quickly during critical periods. Relevant measures include reducing travel time, shortening depot dwell time, improving vehicle handover procedures, adjusting vehicle starting locations, and limiting unnecessary depot-return movements. These measures may improve service punctuality and operational efficiency without immediately increasing fleet size.

Use the model to test operational-control and disruption scenarios. The simulation model should be used to evaluate operational-control policies and plausible changes in operating conditions before structural fleet or infrastructure decisions are made. Relevant experiments include alternative dispatching rules, different vehicle starting locations, shorter depot dwell times, different charging thresholds, improved coordination of charging sessions, temporary vehicle unavailability, pooling, fewer available chargers, reduced charging power, higher energy consumption, increased task demand, longer service durations, aircraft stand changes, and temporary restrictions at refill or dump locations.

Scenario analysis should also examine combinations of vehicle unavailability, charger outages, higher demand, longer service durations, and reduced infrastructure capacity. Testing combined disruptions is important because several moderate disruptions may have a larger operational effect than one isolated parameter change. These experiments can identify the conditions under which the existing fleet and infrastructure configuration no longer provides sufficient robustness, which processes are most vulnerable, and whether operational improvements, spare vehicles, alternative chargers, or additional structural capacity are required. The model can consequently support both operational improvement and contingency planning.

Idle forward staging is a relevant operational policy to test further. It can reduce unnecessary depot-centred deadheading and improve travel efficiency, vehicle availability, and service robustness. In practice, this rule should only be considered when idle vehicles can safely remain staged at or near their current aircraft stand area and when battery state, expected future demand, charging access, apron safety, parking or VOP capacity, dispatcher visibility, and operator handover requirements allow this.

The policy should be evaluated using service, efficiency, charging, and operational-feasibility indicators. These include on-time task completion, the number of late tasks, average and maximum lateness, driven distance, the number of depot returns, charger occupancy, and practical feasibility on the apron. Idle forward staging should only be adopted structurally when operational pilots demonstrate that it improves robustness without creating apron-space conflicts, unsafe staging behaviour, charging-access problems, reduced dispatcher visibility, or difficulties with operator handovers.

Feasibility-aware triage should be treated differently. It should not be framed as a general operational improvement, but as a contingency-support rule for situations in which fleet capacity is temporarily constrained, such as peak demand, vehicle unavailability, or disruption recovery. In these situations, the rule may help dispatchers prioritise tasks that can still be completed on time and thereby reduce the total number of late tasks.

However, feasibility-aware triage can increase the lateness severity of tasks that are already infeasible. It should therefore not be implemented as a simple automatic priority rule. Further testing should explicitly evaluate the trade-off between the number of late tasks and the severity of lateness. Safeguards should include task-criticality checks, maximum-lateness limits, dispatcher override options, and explicit assessment of aircraft departure impact.

When validated through operational pilots, idle forward staging and feasibility-aware triage can complement fleet-capacity planning and support daily decision-making under both normal and disrupted conditions. Until such validation has taken place, their simulated effects should be interpreted as evidence of operational potential rather than as direct justification for reducing procurement targets.

Plan charging infrastructure as a connected but separate decision. Charging infrastructure should be evaluated separately from fleet size while remaining part of the same simulation framework. Charging may not always be the dominant cause of late task completion, but it affects charging-session timing, charger occupancy, maximum simultaneous charger use, vehicle availability, and infrastructure utilisation. Charger-sizing conclusions should therefore not be based on one tested configuration only.

Charger requirements should be tested across multiple operating days, higher-demand cases, different uncertainty conditions, alternative vehicle starting locations, and infrastructure disruptions such as charger unavailability. Coordinated charging strategies and charger sharing between compatible vehicle groups should also be investigated. Smarter distribution of charging sessions, higher-power chargers at selected locations, and shared charger capacity may reduce infrastructure requirements without reducing service performance. This is particularly relevant at airports where airside space and grid capacity are limited and where unnecessary charging locations could occupy space required for other operational functions.

Interpret utilisation and pooling carefully. Fleet utilisation should be interpreted carefully, especially for equipment with peak-driven demand. Low average utilisation does not automatically imply overcapacity. Some vehicle groups may be idle during quiet periods but remain necessary during short operational peaks with strict service windows. Fleet capacity should therefore be evaluated based on peak availability, service-window feasibility, and robustness rather than on average utilisation alone.

Where operational compatibility allows it, vehicle pooling across areas or departments can be evaluated with the simulation model. Such pooling should only be considered when sufficient vehicle availability during critical handling peaks can still be guaranteed. Pooling experiments should account for additional shared task demand, service windows, aircraft compatibility, operator availability, starting locations, handover procedures, and restrictions on cross-departmental use.

Extend the model gradually to other GSE types. The model can be extended to other GSE types, but each extension should be treated as a new verification and validation step. Additional vehicle groups require vehicle-specific service logic, aircraft compatibility rules, operational locations, energy parameters, charging assumptions, resource constraints where applicable, and separate validation before fleet-size conclusions can be drawn.

A gradual extension can support broader assessment of fleet capacity, charging demand, charger sharing, pooling, and operational robustness across a larger part of the eGSE fleet. It can also support the development of an integrated multi-vehicle and multi-department representation in which vehicle groups compete for shared road space, chargers, staging locations, depots, operators, and operational attention.

The general recommendations can be translated into the following action plan for the KLM Schiphol case. The actions focus on using the simulation results as decision-support input rather than as direct procurement advice, as shown in Table 8.1.

Table 8.1: *Case-specific action plan for KLM GS*

Step	Action	KLM-specific application
1	Select service-level interpretation	Decide whether fleet sizing should follow the strict stochastic criterion or the pragmatic lower-bound interpretation. The tested strict stochastic requirements of 16 water vehicles, 15 toilet vehicles, and 25 loaders, and the pragmatic lower bounds of at least 11 water vehicles, 9 toilet vehicles, and 23 loaders, should initially be treated as reference outcomes rather than final fleet decisions.
2	Improve operational data	Prioritise measurement of service durations, apron travel times, travel speeds, waiting times, depot dwell times, depot-return behaviour, task-window definitions, vehicle handover times, charging behaviour, and SOC development using telematics, operational logs, location data, charger data, and dispatcher records.

Continued on next page

Table 8.1 – continued from previous page

Step	Action	KLM-specific application
3	Strengthen model validation	Use the improved data to further validate service times, travel times, depot dwell assumptions, vehicle utilisation, charging behaviour, SOC development, and service-resource processes. Where detailed data remain unavailable, review the assumptions and simulated behaviour with operational experts.
4	Validate across more days	Test additional Schiphol operating days before final decisions are made, including peak days, average weekdays, weekend days, holiday periods, seasonal schedules, winter operations, different aircraft mixes, and disruption-prone days.
5	Test operational and disruption scenarios	Evaluate shorter depot dwell times, alternative dispatching rules, different vehicle starting locations, charging thresholds, improved charging coordination, temporary vehicle unavailability, pooling, fewer available chargers, reduced charging power, higher energy consumption, higher task demand, longer service durations, aircraft stand changes, and temporary restrictions at refill or dump locations. Also test combinations of disruptions to determine whether operational changes, additional fleet capacity, or additional infrastructure are required.
6	Pilot idle forward staging	Pilot idle forward staging under practical constraints related to apron safety, available VOP space, charging access, battery state, future demand, dispatcher visibility, and operator handover. Evaluate on-time completion, late-task count, average and maximum lateness, driven distance, depot returns, charger occupancy, and practical apron feasibility.
7	Use triage only as contingency support	Treat feasibility-aware triage as a scarce-fleet or disruption-support rule rather than as a standard dispatching improvement. Evaluate the trade-off between late-task count and lateness severity using task-criticality checks, maximum-lateness limits, dispatcher override options, and explicit assessment of aircraft departure impact.
8	Reassess charger assumptions	Evaluate charger sizing separately from fleet sizing while keeping both within the simulation framework. The maximum simultaneous charger use in the tested configurations is 7 water chargers, 7 toilet chargers, and 21 loader chargers, compared with assumed availability of 11, 9, and 28. This indicates that the current charger assumptions may be conservative, but it does not constitute a final charger-sizing conclusion.
9	Test charger coordination and sharing	Before changing charger assumptions, test multiple operating days, higher-demand cases, different uncertainty settings, alternative vehicle starting locations, charger-unavailability scenarios, different charging-power levels, coordinated charging strategies, and charger sharing between compatible vehicle groups. Include airport grid and airside-space constraints in the assessment.
10	Finalise operational fleet and technical reserve	After the additional validation, operating-day tests, scenario analyses, and operational pilots, determine the operational fleet requirement for each vehicle group. Add a separate technical reserve for maintenance, failures, battery degradation, charger unavailability, cleaning, inspections, and other temporary downtime. The reserve should reflect operational criticality, expected technical availability, fleet size, redundancy options, downtime patterns, and access to backup vehicles.

Continued on next page

Table 8.1 – continued from previous page

Step	Action	KLM-specific application
11	Extend model scope	Extend the model to loader pooling and additional GSE types after the current vehicle groups and assumptions have been sufficiently validated. Loader pooling should include additional shared task demand, service windows, aircraft compatibility, operator availability, starting locations, handover procedures, locations, and restrictions on cross-departmental use. Other extensions require vehicle-specific service logic, aircraft compatibility rules, operational data, energy parameters, charging assumptions, resource constraints, and separate verification and validation.

8.4. Future Research

Future research can extend this thesis in two main directions. The first direction concerns the academic development of the simulation framework, including its scientific validity, model structure, methodological capabilities, and generalisability beyond the current case-study setting. The second direction concerns the continued development of the KLM Schiphol application through additional validation, operational data collection, policy testing, charging-infrastructure planning, economic assessment, and extension to additional GSE types and operational departments.

8.4.1. Academic future research

Integrate eGSE service performance with turnaround and departure outcomes. Future research could link task-level service performance more directly to aircraft turnaround performance and departure punctuality. The current model evaluates whether individual tasks are completed within their internal service windows. However, an internal service-window violation does not necessarily cause an aircraft delay when sufficient turnaround slack remains or when the delayed task is not part of the critical path.

A useful extension would therefore be to model the complete set of turnaround precedence relations more explicitly. This could include interactions between water and toilet servicing, baggage handling, catering, cleaning, fuelling, passenger boarding, pushback, and other relevant ground-handling processes. Such an extension would allow the model to distinguish between delayed eGSE tasks that directly affect the turnaround critical path and delays that can be absorbed by slack elsewhere in the process.

The resulting model could connect fleet and task performance to airline-level indicators such as D0 punctuality, critical-path delay, actual off-block delay, and missed-connection risk. Fleet-size decisions could then be evaluated based on their effect on airline and turnaround performance rather than only on internal task punctuality. This would be particularly valuable when the model is used for operational decision support in addition to strategic fleet-capacity planning.

Broaden the uncertainty representation. The uncertainty representation could be extended beyond task-timing and travel-time uncertainty. These two uncertainty sources were included in the current study because they directly affect task release times, vehicle availability, and service-window feasibility. Future research could additionally model uncertainty in service durations, task occurrence, operator behaviour, energy consumption, charger availability, infrastructure availability, weather conditions, aircraft stand changes, and vehicle breakdowns.

These uncertainty sources should be introduced selectively and only when reliable operational data or sufficient expert validation are available. Adding stochastic assumptions without adequate empirical support could reduce model credibility rather than improve it. Future research could also examine correlations between uncertainty sources, differences between operating-day types, and the combined effect of several moderate disruptions.

Develop and compare advanced operational-control strategies. The current model uses transparent rule-based dispatching, which supports verification, interpretation, and communication with operational stakeholders. These rules may, however, not generate the best possible operational decisions. Future research could compare the current approach with optimisation-based dispatching, rolling-horizon planning, model predictive control, or hybrid simulation-optimisation methods.

These approaches could make greater use of expected task demand, future vehicle availability, charging opportunities, service criticality, and spatial vehicle positioning. The comparison should not focus only on on-time performance. Computational complexity, transparency, robustness, explainability, data requirements, dispatcher override possibilities, and practical usability should also be assessed.

Include economic and environmental performance. Future research could combine the operational simulation outcomes with economic and environmental indicators. Relevant inputs include vehicle investment costs, maintenance costs, electricity costs, battery replacement costs, charger installation costs, grid-connection costs, and emission reductions.

This would allow fleet-size, charging-infrastructure, and operational-control alternatives to be compared in terms of total cost of ownership, operational resilience, and environmental impact. Such an extension is relevant because the smallest operational fleet or infrastructure configuration is not necessarily the most cost-effective, sustainable, or robust long-term investment.

Test the generalisability of the framework. The simulation framework is designed to be generic, but the current validation is specific to KLM operations at Schiphol and to the selected vehicle groups. Future research could apply the framework to other airports, ground handlers, apron layouts, hub structures, fleet compositions, and operating concepts.

Comparative applications would show which model components are broadly transferable and which require case-specific adaptation. This would strengthen the scientific contribution of the framework and clarify its applicability beyond the KLM Schiphol case.

Develop the model towards real-time decision support. The current model is primarily used for offline scenario analysis and strategic fleet-capacity assessment. With live operational data, the modelling structure could be developed into a day-of-operations decision-support system for vehicle assignment, charging prioritisation, idle staging, disruption recovery, and contingency planning.

Such a system would require integration with real-time flight information, vehicle-location data, battery and SOC data, charger-status data, operational-control systems, and dispatcher interfaces. Further development would also be required regarding computational speed, data reliability, system integration, user interaction, human oversight, and decision explainability. If these requirements are met, the model could develop from an offline simulation study into an operational decision-support tool for electrified ground-support operations.

8.4.2. Case-study continuation

Validate the case study across more operating days. For the KLM Schiphol application, the first priority is broader validation across more operating days. This thesis uses one selected baseline day as the main reference case and an average-day robustness check for water vehicles. This is sufficient to demonstrate the modelling approach and test the main operational mechanisms, but it does not establish whether the fleet-size outcomes remain stable across the complete range of KLM operating conditions.

Future work should therefore apply the model to a representative portfolio of KLM operating days. This portfolio should include additional peak days, average weekdays, weekend days, holiday periods, seasonal schedules, winter operations, different aircraft mixes, and disruption-prone days. This would make it possible to distinguish structural fleet-capacity requirements from day-specific demand effects and determine whether the strict and pragmatic fleet-size thresholds remain stable under different demand patterns and operational peak structures.

Improve case-specific calibration and validation data. KLM should collect more detailed operational data for the parameters that have the strongest effect on task feasibility and vehicle availability. The sensitivity analysis indicates that service duration, travel speed, depot dwell time, task-window definitions, waiting time, depot-return behaviour, and vehicle handover time are particularly important when interpreting fleet-size outcomes.

Measurements from vehicle telematics, operational logs, location data, charger data, and dispatcher records would improve both calibration and validation. Charging behaviour, SOC development, vehicle utilisation, and service-resource processes should also be examined. This is especially important for fleet configurations close to the pragmatic lower bound, where relatively small changes in timing assumptions can determine whether a configuration appears feasible or infeasible.

Validate operational-control policies under Schiphol conditions. Idle forward staging should be evaluated across a broader set of Schiphol-specific operating days, vehicle groups, apron layouts, parking-space and VOP constraints, operator rules, charging-access conditions, safety requirements, handover procedures, and real-time disruption conditions. This would clarify whether the improvements observed in the current study remain robust in practice or are partly caused by the selected demand pattern and modelling assumptions.

Feasibility-aware triage should be studied separately as a contingency-support rule. Future work should explicitly evaluate the trade-off between reducing the number of late tasks and increasing the lateness severity of tasks that are already infeasible. The assessment should include task criticality, aircraft departure impact, maximum acceptable lateness, dispatcher override options, and practical implementation safeguards.

Both policies should be validated through operational pilots before they are incorporated into structural fleet-capacity or procurement decisions. When validated, they may support daily decision-making under both normal and disrupted conditions.

Develop charging-infrastructure planning for Schiphol. KLM should further investigate charging-infrastructure planning at Schiphol. Charging is not the dominant cause of late task completion in the tested configurations, but it remains important for vehicle availability, infrastructure utilisation, and long-term electrification planning.

Future work could analyse the required number and spatial distribution of chargers, the effects of different charging-power levels, charging thresholds, coordinated charging strategies, charger sharing between compatible vehicle groups, and the operational consequences of charger unavailability. These analyses should cover multiple operating days, higher-demand conditions, alternative vehicle starting locations, different uncertainty settings, and future fleet-growth scenarios.

The assessment should also consider airport grid constraints, airside-space restrictions, and competition between charging infrastructure and other operational functions. Higher-power chargers at selected locations and shared charger capacity may reduce the total number of required chargers, but only when service performance and operational robustness remain sufficient.

Develop an integrated multi-fleet and multi-department model. The model should be extended gradually to additional GSE types within the KLM operation. The current case study focuses on water vehicles, toilet vehicles, and loaders. These vehicle groups test the main mechanisms included in the model, including depot-based resource-constrained operations, charging, refill and dump processes, stand-side service operations, and time-window-based dispatching. However, they do not represent the full operational diversity of the GSE fleet.

Future work could include baggage tractors, belt loaders, transporters, loader-transporters, passenger stairs, catering vehicles, dispensers, pushback tractors, and other relevant equipment groups. Each extension requires vehicle-specific service rules, aircraft compatibility logic, operational locations, energy parameters, charging assumptions, resource constraints where applicable, and separate verification and validation.

The extended model should also account for interactions between vehicle groups, operational departments, and shared infrastructure. The current experiments evaluate vehicle groups separately, which is appropriate for isolating vehicle-specific fleet-size requirements. In practice, however, different GSE types and operational departments may compete for shared road space, chargers, staging locations, depots, operators, and operational attention.

An integrated multi-vehicle and multi-department simulation would allow KLM to determine whether conclusions for individual vehicle groups remain valid when several eGSE fleets and operational task pools are represented simultaneously. It would also enable assessment of charger sharing, apron congestion, shared staging areas, cross-fleet dependencies, and operational trade-offs.

This integrated perspective is particularly relevant for loader operations. The results show that loaders have relatively low average utilisation, but their demand is concentrated during peak periods and constrained by strict service windows. Low average utilisation alone is therefore not sufficient evidence for reducing loader capacity or reallocating loaders to other operations.

Future research could add loader tasks from other operational areas, departments, or handling processes to assess whether a shared loader pool can be used without reducing availability during critical peaks. Such an extension should include the additional task demand, service windows, aircraft compatibility rules, operator availability, starting locations, handover procedures, operational locations, and restrictions on cross-departmental use. This would allow KLM to determine whether loader pooling can improve asset utilisation while maintaining sufficient operational robustness.

Add a KLM-specific economic and environmental assessment. The extended model could support procurement and infrastructure decisions by incorporating KLM-specific economic and environmental data. Relevant inputs include vehicle investment costs, maintenance costs, electricity costs, battery replacement costs, charger installation costs, grid-connection costs, and emission reductions.

Combining these inputs with the operational simulation outcomes would allow KLM to compare fleet-size, charging-infrastructure, pooling, and operational-control alternatives in terms of total cost of ownership, operational resilience, and sustainability impact. This would support a more complete assessment than fleet size or on-time performance alone.

Bibliography

[1] Justine Marie E. Abarro et al. "A Tale of Nickel-Iron Batteries: Its Resurgence in the Age of Modern Batteries". In: *Batteries* 9.7 (July 2023), p. 383. ISSN: 2313-0105. DOI: 10.3390/batteries9070383. URL: http://dx.doi.org/10.3390/batteries9070383.

[2] Aero Specialties, Inc. "Aero Boeing B777 Towbar". In: *Aero Specialties Ground Support Equipment Product Information* (Jan. 2025). URL: https://www.aerospecialties.com/aviation-ground-support-equipment-gse-products/towbar-head-equipment/commercial-towbars/aero-boeing-b777-towbar/.

[3] Air France-KLM. *Annual Report 2007–08*. Annual report. Air France-KLM, 2008. URL: https://www.airfranceklm.com/sites/default/files/publications_en/document/annual-report_2007-08_en.pdf (visited on 06/19/2026).

[4] Air Transport Action Group. *Aviation and Climate Change: Fact Sheet #2*. Fact sheet, Enviro.aero / Air Transport Action Group. Dec. 2024. URL: https://atag.org/media/gw5cgzzh/fact-sheet_2_aviation-and-climate-change.pdf (visited on 06/19/2026).

[5] Airbus S.A.S. "A320 Family Aircraft Characteristics: Airport and Maintenance Planning". In: *Airbus Aircraft Characteristics Series* AC_A320_0624 (June 2024). URL: https://aircraft.airbus.com/sites/g/files/jlcbta126/files/2025-01/AC_A320_0624.pdf.

[6] Airbus S.A.S. "A320neo". In: *Airbus Commercial Aircraft Product Information* (Jan. 2025). URL: https://www.aircraft.airbus.com/en/aircraft/a320-family/a320neo.

[7] KLM Royal Dutch Airlines. *Company Profile*. Web page, Corporate information on the KLM website. Accessed 2025-12-03. Dec. 2025. URL: https://www.klm.nl/en/information/corporate/company-profile.

[8] KLM Royal Dutch Airlines. *Sustainability*. Web page, KLM website – Sustainability section. Accessed 2025-12-03. Dec. 2025. URL: https://www.klm.nl/en/information/sustainability.

[9] Airport Carbon Accreditation Programme. *Airport Carbon Accreditation: Annual Report 2013–2014*. Annual report Year 5. Airports Council International Europe, 2014. URL: https://www.airportcarbonaccreditation.org/wp-content/uploads/2023/11/ACA-Report_WEB.pdf (visited on 06/19/2026).

[10] Ahmet Aktaş and Yağmur Kirçiçek. "Solar Hybrid Systems and Energy Storage Systems". In: *Solar Hybrid Systems*. Elsevier, 2021, pp. 87–125. ISBN: 9780323884990. DOI: 10.1016/b978-0-323-88499-0.00005-7. URL: http://dx.doi.org/10.1016/B978-0-323-88499-0.00005-7.

[11] Mohammed Alruwaili and Liana Cipcigan. "Airport electrified ground support equipment for providing ancillary services to the grid". In: *Electric Power Systems Research* 211 (Oct. 2022), p. 108242. ISSN: 0378-7796. DOI: 10.1016/j.epsr.2022.108242. URL: http://dx.doi.org/10.1016/j.epsr.2022.108242.

[12] Calvin A. Arter et al. "Air quality and health-related impacts of traditional and alternate jet fuels from airport aircraft operations in the U.S." In: *Environment International* 158 (Jan. 2022), p. 106958. ISSN: 0160-4120. DOI: 10.1016/j.envint.2021.106958. URL: http://dx.doi.org/10.1016/j.envint.2021.106958.

[13] Aviation Ground Equipment. "Military Air Conditioning and Heating Equipment". In: *Aviation Ground Equipment Product Catalog* (Jan. 2025). URL: https://www.aviationgroundequip.com/category_military/air-conditioning-and-heating/.

[14] E.V.M. van Baaren. "The feasibility of a fully electric aircraft towing system". Graduation date: 24-05-2019. Accessed 2025-12-02. Master thesis. TU Delft, May 2019. URL: https://resolver.tudelft.nl/uuid:c47a1e3e-8d3b-4eda-8b3b-134ae29f6af9.

[15] Dan-wen Bao, Zhuo Chen, and Di Kang. "Proactive real-time scheduling method for apron service vehicles based on mixed strategies". In: *Computers & Industrial Engineering* 192 (June 2024), p. 110182. ISSN: 0360-8352. DOI: 10.1016/j.cie.2024.110182. URL: http://dx.doi.org/10.1016/j.cie.2024.110182.

[16] Dan-Wen Bao et al. "Mixed fleet scheduling method for airport ground service vehicles under the trend of electrification". In: *Journal of Air Transport Management* 108 (May 2023), p. 102379. ISSN: 0969-6997. DOI: 10.1016/j.jairtraman.2023.102379. URL: http://dx.doi.org/10.1016/j.jairtraman.2023.102379.

[17] Ranjan Kumar Bose et al. "Airport Ground Support Equipment Infrastructure & Logistics Electrification Assessment Tool: 2025 Data Development, Modeling and Analysis for DFW". In: *Oak Ridge National Laboratory Technical Report* ORNL/TM-2025/3999 (July 2025). DOI: 10.2172/2586884. URL: https://doi.org/10.2172/2586884.

[18] Ranjan Kumar Bose et al. "Electrifying Airport GSE: Monte Carlo Grid Impacts". In: *2025 IEEE/AIAA Transportation Electrification Conference and Electric Aircraft Technologies Symposium (ITEC+EATS)*. IEEE, June 2025, pp. 1–6. DOI: 10.1109/itec63604.2025.11097948. URL: http://dx.doi.org/10.1109/ITEC63604.2025.11097948.

[19] T.P.G. Bosma. "Design of an Airport Service Infrastructure for Sustainable Refueling of Aircraft". In: *Delft University of Technology* MSc Thesis (June 2022).

[20] S.-T. Chen. "Current Aircraft Ground Handling Operations". In: *Delft University of Technology* PhD Thesis (Jan. 2022).

[21] Shaohua Cui, Hui Zhao, and Cuiping Zhang. "Locating Charging Stations of Various Sizes with Different Numbers of Chargers for Battery Electric Vehicles". In: *Energies* 11.11 (Nov. 2018), p. 3056. ISSN: 1996-1073. DOI: 10.3390/en11113056. URL: http://dx.doi.org/10.3390/en11113056.

[22] H.S. Das et al. "Electric vehicles standards, charging infrastructure, and impact on grid integration: A technological review". In: *Renewable and Sustainable Energy Reviews* 120 (Mar. 2020), p. 109618. ISSN: 1364-0321. DOI: 10.1016/j.rser.2019.109618. URL: http://dx.doi.org/10.1016/j.rser.2019.109618.

[23] Armen Der Kiureghian and Ove Ditlevsen. "Aleatory or Epistemic? Does It Matter?" In: *Structural Safety* 31.2 (Mar. 2009), pp. 105–112. ISSN: 0167-4730. DOI: 10.1016/j.strusafe.2008.06.020. URL: http://dx.doi.org/10.1016/j.strusafe.2008.06.020.

[24] Expert. *eGSE operations, turnaround constraints, and battery/charging considerations*. Personal communication (conversation). Role: team member of the GSE team. Topics: baggage handling constraints during missing passenger/bag removal; variability of turnaround service sets vs manufacturer guidelines; omission of water/lavatory services at outstations for short-haul flights; suitability of lithium-based batteries (LFP, NMC) for eGSE due to cycle life and energy density vs lead-acid; safety considerations in KLM decision-making; on-board vs off-board charging depends on vehicle type, duty cycle and airport operational requirements. Notes held by the author. Dec. 2025.

[25] Pablo Fleurquin, José J. Ramasco, and Victor M. Eguíluz. "Characterization of delay propagation in the US air transportation network". In: *arXiv preprint arXiv:1304.2528* (Apr. 2013). URL: https://arxiv.org/abs/1304.2528.

[27] Weigang Fu et al. “A bi-objective optimization approach for scheduling electric ground-handling vehicles in an airport”. In: *Complex &; Intelligent Systems* 11.4 (Mar. 2025). ISSN: 2198-6053. DOI: 10.1007/s40747-025-01815-x. URL: http://dx.doi.org/10.1007/s40747-025-01815-x.

[28] Yagmur S. Gök et al. “Constraint-based robust planning and scheduling of airport apron operations through simheuristics”. In: *Annals of Operations Research* 320.2 (Feb. 2022), pp. 795–830. ISSN: 1572-9338. DOI: 10.1007/s10479-022-04547-0. URL: http://dx.doi.org/10.1007/s10479-022-04547-0.

[29] Philipp Grosche et al. *Accelerating airport decarbonization*. Insights / Publication, Roland Berger. Accessed 2025-12-02. Aug. 2023. URL: https://www.rolandberger.com/en/Insights/Publications/Accelerating-airport-decarbonization.html.

[30] Kyle Gulan, Eduardo Cotilla-Sanchez, and Yue Cao. “Charging Analysis of Ground Support Vehicles in an Electrified Airport”. In: *2019 IEEE Transportation Electrification Conference and Expo (ITEC)*. IEEE, June 2019, pp. 1–6. DOI: 10.1109/ITEC.2019.8790550. URL: https://doi.org/10.1109/ITEC.2019.8790550.

[31] M.A. Hannan et al. “A review of lithium-ion battery state of charge estimation and management system in electric vehicle applications: Challenges and recommendations”. In: *Renewable and Sustainable Energy Reviews* 78 (Oct. 2017), pp. 834–854. ISSN: 1364-0321. DOI: 10.1016/j.rser.2017.05.001. URL: http://dx.doi.org/10.1016/j.rser.2017.05.001.

[32] Samuel Hasselwander, Markus Meyer, and Ines Österle. “Techno-Economic Analysis of Different Battery Cell Chemistries for the Passenger Vehicle Market”. In: *Batteries* 9.7 (July 2023), p. 379. ISSN: 2313-0105. DOI: 10.3390/batteries9070379. URL: http://dx.doi.org/10.3390/batteries9070379.

[33] Kathryn Hoad, Stewart Robinson, and Ruth Davies. “Automating DES Output Analysis: How Many Replications to Run”. In: *Proceedings of the 2007 Winter Simulation Conference*. 2007, pp. 505–512. URL: https://www.informs-sim.org/wsc07papers/060.pdf.

[34] Pamela G. Horkos, Emile Yammine, and Nabil Karami. “Review on Different Charging Techniques of Lead-Acid Batteries”. In: *2015 3rd International Conference on Technological Advances in Electrical, Electronics and Computer Engineering (TAEECE)*. IEEE, Apr. 2015, pp. 27–32. DOI: 10.1109/TAEECE.2015.7113595. URL: https://doi.org/10.1109/TAEECE.2015.7113595.

[36] Fayaz Hussain et al. “Energy storage technologies”. In: *Energy for Sustainable Development*. Elsevier, 2020, pp. 125–165. ISBN: 9780128146453. DOI: 10.1016/b978-0-12-814645-3.00006-7. URL: http://dx.doi.org/10.1016/B978-0-12-814645-3.00006-7.

[37] International Air Transport Association. *IATA Ground Operations Manual (IGOM)*. Used for industry-standard ground handling process definitions. International Air Transport Association. n.d. URL: https://www.iata.org/en/publications/manuals/iata-ground-operations-manual/ (visited on 06/19/2026).

[38] International Civil Aviation Organization. *Airport Air Quality Manual (Doc 9889)*. 2nd ed. International Civil Aviation Organization. 2020. URL: https://www.icao.int/sites/default/files/sp-files/environmental-protection/Documents/Doc%209889.SGAR.WG2.Initial%20Update.pdf (visited on 06/19/2026).

[39] Mehmet Cagin Kirca, Andrew McGordon, and Truong Quang Dinh. “Multi-Input Multi-Output Model of Airport Infrastructure for Reducing CO2 Emissions”. In: *Proceedings of the 2020 IEEE Vehicle Power and Propulsion Conference (VPPC)* (Dec. 2020). DOI: 10.1109/VPPC49601.2020.9330949. URL: https://ieeexplore.ieee.org/document/9330949.

[40] KLM Ground Services. *Energy-consumption data for water, toilet and loader GSE*. Internal unpublished dataset. 2026.

[41] KLM Royal Dutch Airlines. "Careers at KLM". In: *KLM Corporate Website* (Jan. 2025). URL: https://careers.klm.com/nl/.

[42] KLM Royal Dutch Airlines. *Read all about the history of KLM*. Used for KLM Cityhopper's feeder role within the Schiphol hub structure. n.d. URL: https://www.klm.com/information/corporate/history (visited on 04/14/2026).

[43] KLM Royal Dutch Airlines. *Transfer to a connecting flight with KLM*. Used for Schiphol minimum transfer times and the transfer function of the KLM network. n.d. URL: https://www.klm.com/information/airport/transfers (visited on 04/14/2026).

[44] Vincent B.F. Kuipers. "Simulation and Control Approaches for Enhancing Efficiency of Electric Ground Fleets in Airport Operations". In: *Eindhoven University of Technology* MSc Thesis (Sept. 2024). URL: https://research.tue.nl/en/studentTheses/simulation-and-control-approaches-for-enhancing-efficiency.

[45] Averill M. Law. *Simulation Modeling and Analysis*. 5th ed. McGraw-Hill Education, 2015. ISBN: 9780073401324.

[46] Zhenhe Li, Amir Khajepour, and Jinchun Song. "A comprehensive review of the key technologies for pure electric vehicles". In: *Energy* 182 (Sept. 2019), pp. 824–839. ISSN: 0360-5442. DOI: 10.1016/j.energy.2019.06.077. URL: http://dx.doi.org/10.1016/j.energy.2019.06.077.

[47] Eckhard Limpert, Werner A. Stahel, and Markus Abbt. "Log-normal Distributions across the Sciences: Keys and Clues". In: *BioScience* 51.5 (May 2001), pp. 341–352. ISSN: 0006-3568. DOI: 10.1641/0006-3568(2001)051[0341:LNDATS]2.0.CO;2. URL: http://dx.doi.org/10.1641/0006-3568(2001)051%5B0341:LNDATS%5D2.0.CO;2.

[48] Xiaochen Liu et al. "Quantifying energy flexibility potential of ground electric vehicles in an airport with real behavior data". In: *Sustainable Cities and Society* 105 (June 2024), p. 105331. ISSN: 2210-6707. DOI: 10.1016/j.scs.2024.105331. URL: http://dx.doi.org/10.1016/j.scs.2024.105331.

[49] Qian Luo et al. "Multi-strategy cooperative scheduling for airport specialized vehicles based on digital twins". In: *Scientific Reports* 14.1 (July 2024). ISSN: 2045-2322. DOI: 10.1038/s41598-024-66350-0. URL: http://dx.doi.org/10.1038/s41598-024-66350-0.

[50] Evangelos Mitsokapas et al. "Statistical Characterization of Airplane Delays". In: *arXiv preprint arXiv:2101.10789* (Jan. 2021). URL: https://arxiv.org/abs/2101.10789.

[51] Muhamad Aiman Nazri et al. "Screen-Printed Nickel–Zinc Batteries: A Review of Additive Manufacturing and Evaluation Methods". In: *3D Printing and Additive Manufacturing* 8.3 (June 2021), pp. 176–192. ISSN: 2329-7670. DOI: 10.1089/3dp.2020.0095. URL: http://dx.doi.org/10.1089/3dp.2020.0095.

[52] New York State Energy Research and Development Authority. *Energy Storage System Performance Impact Evaluation*. Impact evaluation report. New York State Energy Research and Development Authority, June 2024. URL: https://www.nyserda.ny.gov/-/media/Project/Nyserda/Files/Publications/PPSER/NYSERDA/2024-Energy-Storage-Impact-Evaluation-Report.pdf (visited on 06/19/2026).

[53] Simon van Oosterom and Mihaela Mitici. "An environmentally-aware dynamic planning of electric vehicles for aircraft towing considering stochastic aircraft arrival and departure times". In: *Transportation Research Part C: Emerging Technologies* 169 (Dec. 2024), p. 104857. ISSN: 0968-090X. DOI: 10.1016/j.trc.2024.104857. URL: http://dx.doi.org/10.1016/j.trc.2024.104857.

[54] 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 (Feb. 2023), p. 103995. ISSN: 0968-090X. DOI: 10.1016/j.trc.2022.103995. URL: http://dx.doi.org/10.1016/j.trc.2022.103995.

[56] Silvia Padrón et al. "A bi-objective approach for scheduling ground-handling vehicles in airports". In: *Computers &; Operations Research* 71 (July 2016), pp. 34–53. ISSN: 0305-0548. DOI: 10.1016/j.cor.2015.12.010. URL: http://dx.doi.org/10.1016/j.cor.2015.12.010.

[58] Habiballah Rahimi-Eichi et al. "Battery Management System: An Overview of Its Application in the Smart Grid and Electric Vehicles". In: *IEEE Industrial Electronics Magazine* 7.2 (June 2013), pp. 4–15. ISSN: 1932-4529. DOI: 10.1109/MIE.2013.2250351. URL: https://ieeexplore.ieee.org/document/6515963.

[59] Satish Rajagopalan et al. "Power Quality Impacts of Airport Ground Support Equipment Charging Systems". In: *2003 IEEE Power Engineering Society General Meeting*. IEEE, July 2003, pp. 1226–1231. DOI: 10.1109/PES.2003.1270504. URL: https://doi.org/10.1109/PES.2003.1270504.

[60] Thomas B. Reddy. *Linden's Handbook of Batteries*. 4th ed. McGraw-Hill, 2010. ISBN: 9780071624213. URL: https://books.google.com/books/about/Linden_s_Handbook_of_Batteries_4th_Editi.html?id=bVNCAQAACAAJ.

[63] Sebastian Rivera et al. "Electric Vehicle Charging Infrastructure: From Grid to Battery". In: *IEEE Industrial Electronics Magazine* 15.2 (June 2021), pp. 37–51. ISSN: 1932-4529. DOI: 10.1109/MIE.2020.3039039. URL: https://ieeexplore.ieee.org/document/9352049.

[64] Manuel D. Rossetti. *Simulation Modeling and Arena*. Open Educational Resources. University of Arkansas, 2021. URL: https://scholarworks.uark.edu/oer/6/.

[65] Royal Schiphol Group. *Airport Industry Connectivity Report*. Accessed 14 April 2026; page states Schiphol ranks first in European direct connectivity and third in European hub connectivity. 2025. URL: https://www.schiphol.nl/en/advertising/news/airport-industry-connectivity-report (visited on 04/14/2026).

[66] Royal Schiphol Group. *Aviation Solutions | Deep Turnaround*. Accessed 14 April 2026; page describes Deep Turnaround as detecting over 70 unique turnaround events across 30 turnaround processes. 2026. URL: https://www.schiphol.nl/en/aviation-solutions/deep-turnaround/ (visited on 04/14/2026).

[67] Royal Schiphol Group. *Emission free by 2030*. Web page, Amsterdam Airport Schiphol – Schiphol as a neighbour. Accessed 2025-12-02. Dec. 2025. URL: https://www.schiphol.nl/en/schiphol-as-a-neighbour/emission-free-by-2030/.

[68] Royal Schiphol Group. *Government commits to fewer flights at Schiphol*. Accessed 14 April 2026; page states a maximum of 478,000 flight movements per year. Dec. 2024. URL: https://www.schiphol.nl/en/you-and-schiphol/news/government-commits-to-fewer-flights-at-schiphol/ (visited on 04/14/2026).

[69] Royal Schiphol Group. *Royal Schiphol Group to significantly expand number of electric charging points in coming years*. Accessed 14 April 2026; page states Schiphol plans up to 10,000 charging stations by 2030 and links this to emission-free airports by 2030. Nov. 2022. URL: https://news.schiphol.com/royal-schiphol-group-to-significantly-expand-number-of-electric-charging-points-in-coming-years/ (visited on 04/14/2026).

[70] Royal Schiphol Group. *Schiphol | Annual Traffic Reviews*. Accessed 14 April 2026; page reports Schiphol's 2024 traffic figures including 66.8 million passengers and 473,815 air transport movements. 2025. URL: https://www.schiphol.nl/en/schiphol-group/traffic-review (visited on 04/14/2026).

[71] Royal Schiphol Group. *Schiphol | Valuable Assets*. Accessed 14 April 2026; page states Schiphol has 6 runways and 245 aircraft stands. Sept. 2022. URL: https://www.schiphol.nl/en/blog/valuable-assets/ (visited on 04/14/2026).

[72] Royal Schiphol Group. *The Schiphol Tower(s)*. Accessed 14 April 2026; page states Schiphol uses two towers regularly and that the Polderbaan is too far from the main tower. Feb. 2022. URL: https://www.schiphol.nl/en/blog/the-schiphol-tower/ (visited on 04/14/2026).

[73] Royal Schiphol Group N.V. and Schiphol Nederland B.V. *Prospectus dated 8 May 2024: Euro Medium Term Note Programme*. Used for facts and figures on Amsterdam Airport Schiphol, including the role of Air France-KLM and transfer dependence. May 2024. URL: https://assets.ctfassets.net/biom0eqyyi6b/5bVzojicFIaXytQKtvuUoH/1ff96539a26481a3ba0e1fc6c6906f05/EMTN_Prospectus_2024.pdf (visited on 04/14/2026).

[74] Michael Schmidt. "A review of aircraft turnaround operations and simulations". In: *Progress in Aerospace Sciences* 92 (July 2017), pp. 25–38. ISSN: 0376-0421. DOI: 10.1016/j.paerosci.2017.05.002. URL: http://dx.doi.org/10.1016/j.paerosci.2017.05.002.

[75] Weixiang Shen, Thanh Tu Vo, and Ajay Kapoor. "Charging Algorithms of Lithium-Ion Batteries: An Overview". In: *2012 7th IEEE Conference on Industrial Electronics and Applications (ICIEA)*. IEEE, July 2012, pp. 1567–1572. DOI: 10.1109/ICIEA.2012.6360973. URL: https://doi.org/10.1109/ICIEA.2012.6360973.

[76] Sony Corporation. "Sony US18650 Lithium Ion Battery Specifications". In: *Sony Lithium Ion Battery Catalog* SPEC_SONY_US18650 (June 2001). URL: https://www.megacellmonitor.com/pdf/vendor_specs/SPEC_SONY_US18650.pdf.

[77] Yuxiang Tang et al. "Cooperative scheduling of airport ground electric service vehicles considering workload balance: A column generation approach". In: *Computers &; Industrial Engineering* 200 (Feb. 2025), p. 110773. ISSN: 0360-8352. DOI: 10.1016/j.cie.2024.110773. URL: http://dx.doi.org/10.1016/j.cie.2024.110773.

[78] Koen Timmermans. "Electric Ground Support Equipment at Airports". In: *Delft University of Technology* MSc Thesis (Dec. 2023).

[79] Koen Timmermans et al. "The impact of transitioning to electric Ground Support Equipment on the fleet capacity and energy demand at airports". In: *Case Studies on Transport Policy* 21 (Sept. 2025), p. 101498. ISSN: 2213-624X. DOI: 10.1016/j.cstp.2025.101498. URL: http://dx.doi.org/10.1016/j.cstp.2025.101498.

[80] Koen Timmermans et al. "The impact of transitioning to electric Ground Support Equipment on the fleet capacity and energy demand at airports". In: *Case Studies on Transport Policy* 21 (Sept. 2025), p. 101498. ISSN: 2213-624X. DOI: 10.1016/j.cstp.2025.101498. URL: http://dx.doi.org/10.1016/j.cstp.2025.101498.

[81] TLD Group. "Products". In: *TLD Group Product Portfolio* (Jan. 2025). URL: https://www.tld-group.com/products/.

[84] Fengde Wang et al. "A Deep Reinforcement Learning Framework for Multi-Fleet Scheduling and Optimization of Hybrid Ground Support Equipment Vehicles in Airport Operations". In: *Applied Sciences* 15.17 (Sept. 2025), p. 9777. ISSN: 2076-3417. DOI: 10.3390/app15179777. URL: http://dx.doi.org/10.3390/app15179777.

[85] Shixiong Wang et al. "Accurate Tracking, Collision Detection, and Optimal Scheduling of Airport Ground Support Equipment". In: *IEEE Internet of Things Journal* 8.1 (Jan. 2021), pp. 572–584. ISSN: 2372-2541. DOI: 10.1109/jiot.2020.3004874. URL: http://dx.doi.org/10.1109/JIOT.2020.3004874.

[86] Han Xiao et al. "A Survey of Influence of Electric Vehicle Charging on Power Grid". In: *2014 IEEE 9th Conference on Industrial Electronics and Applications (ICIEA)*. IEEE, June 2014, pp. 121–126. DOI: 10.1109/ICIEA.2014.6931143. URL: https://doi.org/10.1109/ICIEA.2014.6931143.

[87] Wei Xie, Barry L. Nelson, and Russell R. Barton. "Statistical Uncertainty Analysis for Stochastic Simulation". In: *arXiv preprint arXiv:2011.04207* (Nov. 2020). URL: https://arxiv.org/abs/2011.04207.

[88] Guoning Xu et al. "Ferry Scheduling Optimization Considering Arrival Time Uncertainty and In-Place Time Differences". In: *Applied Sciences* 13.20 (Oct. 2023), p. 11174. ISSN: 2076-3417. DOI: 10.3390/app132011174. URL: http://dx.doi.org/10.3390/app132011174.

[89] Mengyun Yao et al. "A Two-Stage Optimization Model for Airport Stand Allocation and Ground Support Vehicle Scheduling". In: *Applied Sciences* 14.23 (Dec. 2024), p. 11407. ISSN: 2076-3417. DOI: 10.3390/app142311407. URL: http://dx.doi.org/10.3390/app142311407.

[90] Jia Ying Yong et al. "A review on the state-of-the-art technologies of electric vehicle, its impacts and prospects". In: *Renewable and Sustainable Energy Reviews* 49 (Sept. 2015), pp. 365–385. ISSN: 1364-0321. DOI: 10.1016/j.rser.2015.04.130. URL: http://dx.doi.org/10.1016/j.rser.2015.04.130.

[91] Ziyu Zhang et al. "A review of technologies and applications on versatile energy storage systems". In: *Renewable and Sustainable Energy Reviews* 148 (Sept. 2021), p. 111263. ISSN: 1364-0321. DOI: 10.1016/j.rser.2021.111263. URL: http://dx.doi.org/10.1016/j.rser.2021.111263.

[92] Peng Zhou et al. "A comprehensive review of ground support equipment scheduling for aircraft ground handling services". In: *Transportation Research Part E: Logistics and Transportation Review* 203 (Nov. 2025), p. 104341. ISSN: 1366-5545. DOI: 10.1016/j.tre.2025.104341. URL: http://dx.doi.org/10.1016/j.tre.2025.104341.

[93] Bin Zhu et al. "Research on data mining model of fault operation and maintenance based on electric vehicle charging behavior". In: *Frontiers in Energy Research* 10 (Jan. 2023). ISSN: 2296-598X. DOI: 10.3389/fenrg.2022.1044379. URL: http://dx.doi.org/10.3389/fenrg.2022.1044379.

[94] M. Zoutendijk and M. Mitici. "Fleet scheduling for electric towing of aircraft under limited airport energy capacity". In: *Energy* 294 (May 2024), p. 130924. ISSN: 0360-5442. DOI: 10.1016/j.energy.2024.130924. URL: http://dx.doi.org/10.1016/j.energy.2024.130924.

[95] M. Zoutendijk, M. Mitici, and J.M. Hoekstra. "An investigation of operational management solutions and challenges for electric taxiing of aircraft". In: *Research in Transportation Business &; Management* 49 (Aug. 2023), p. 101019. ISSN: 2210-5395. DOI: 10.1016/j.rtbm.2023.101019. URL: http://dx.doi.org/10.1016/j.rtbm.2023.101019.

A

Ground Support Equipment

This information was taken from Timmermans (2023) [78].

A.1. Ground Support Equipment

eGSE plays a critical role in ensuring efficient freight and passenger movement at airports. Such freight facility equipment facilitates aircraft servicing between flights, including refuelling, towing, luggage and freight handling, passenger transport, de-icing, catering and sewage removal. GSE supports, services, and maintains aircraft operations on the ground, ensuring efficiency, safety, and timely departures. This includes belt and deck loaders, highloaders, de-icer, potable-water trucks, fuel trucks, baggage trucks, cargo transporters, lavatory/catering/cleaning trucks, disabled-passenger lift trucks, towing tractors, passenger stairs, and delivering GPU, ASU, and PCA. Classified into 5 categories: aircraft handling equipment, cargo and baggage handling equipement, passenger handling equipment, aircraft servicing equipment, maintenance and safety equipment [17].

eGSE is increasingly being adopted due to its operational benefits, including high torque, and lower maintenance costs. Airlines, contractors, and airports benefit from centralized procurement and maintenance, making eGSE a viable alternative to traditional diesel-powered equipment. Additionally, electric chargers can be more widely distributed across airports compared to diesel refuelling stations, minimizing unnecessary equipment movement [17].

A.1.1. Aircraft Support Equipment

- **Ground Power Unit (GPU)**

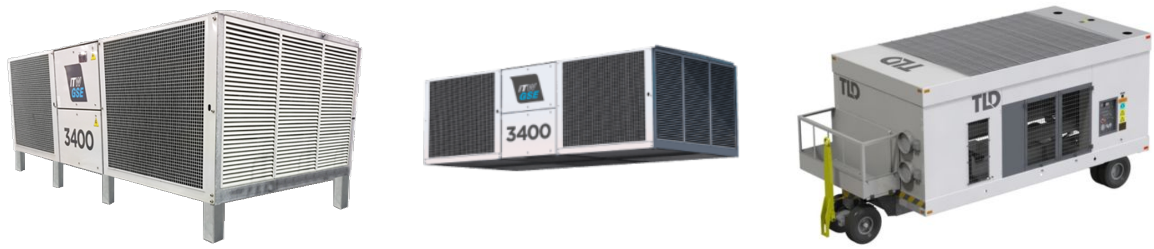
  A GPU is designed to provide electricity to an aircraft when it is parked on the ground and its main engines and on-board APU are off. Aircraft require power at 115 V and 400 Hz. Depending on the size of the aircraft, it may be necessary to connect multiple GPUs. They can also be used to start aircraft engines. There are three types of GPUs with respect to their location and mobility: Fixed GPU; Bridge-mounted GPU; Mobile GPU.

(a) *An example of a fixed GPU [81].* **(b)** *An example of a bridge-mounted GPU [81].* **(c)** *An example of a mobile GPU [81].*

Figure A.1: *Three different types of GPUs.*

- **Pre-Conditioned Air Unit (PCA)**

  Also referred to as air carts, these units provide conditioned (i.e., cooled and heated) air to parked aircraft when their main engines and APU are off. As with the GPUs, the following three types of PCA units can be distinguished: Fixed PCA unit; Bridge-mounted PCA unit; Mobile PCA unit.



(a) *An example of a fixed PCA [13].* **(b)** *An example of a bridge-mounted PCA [81].* **(c)** *An example of a mobile PCA [81]*

Figure A.2: *Three different types of PCAs.*

- **Air Starter Unit (ASU)**
  An Air Starter Unit (ASU) delivers compressed air for the initial rotation of the engine of an aircraft to help start the engines during instances when the on-board APU is not operational or there are no other means for starting the engines. A towing tractor is often used to transport the ASU, but there are also "truck-mounted" ASUs. In addition to starting engines, an ASU can also be used for the air supply of PCA. The ASU is then connected to an air conditioning "pack" on the aircraft.

Figure A.3: *An example of an ASU [81].*

A.1.2. Aircraft Movement Equipment

- **Aircraft Tractor**
  Aircraft tractors (i.e., pushback and/or tow tugs) assist with aircraft movements when an aircraft cannot operate its engines, or does not have sufficient manoeuvrability to move and turn under its own engine power. It is most frequently used to push an aircraft back from the gate/-stand and onto the aircraft movement area (e.g., taxiway). It may also be used to move an aircraft to other locations on an airport (e.g., maintenance hangars and parking bays).There are two main types of aircraft tractors: Conventional aircraft tractor, Conventional aircraft tractors use tow-bars that are connected to an aircraft's nose wheel; Towbarless aircraft tractors, Towbarless aircraft tractors scoop up the aircraft nose wheel and lift it off the ground before moving.

(a) *An example of a conventional aircraft tractor [81].* **(b)** *An example of a towbar [2].* **(c)** *An example of a towbarless aircraft tractor [81].*

Figure A.4: *Two different types of aircraft tractors and an example of a towbar.*

A.1.3. Aircraft Servicing Equipment

- **Catering Truck**
  Catering trucks are used to deliver beverage and food trolleys but they are also used to remove trash and other used materials. Catering trucks use a scissor lift elevation system to raise their "container" to the level of an aircraft door. A bridge connection is created through an adjustable platform to allow the staff to replace trolleys.

Figure A.5: *An example of a catering truck [35].*

- **Lavatory Service**
  A vehicle that is designed for, and equipped with, a collection system and a storage tank for the recovery and transport of aircraft lavatory wastes. It is also used for rinsing the aircraft's waste tank and to replenish the disinfecting chemicals into the aircraft lavatory system. Fuel and electric powered models are available.

Figure A.6: *An example of a lavatory service truck [81].*

- **Water Truck**
  This vehicle is equipped with a transfer system (i.e., pipes/hoses, filters, pumps) and storage tank for the provision of (drinkable) water to an aircraft. Fuel and electric powered models are available.

Figure A.7: *An example of a water truck [81].*

- **Refueler**
  A self-contained refueler with pumps, filters, hoses and valves can be used to dispense fuel directly to an aircraft's tank(s). The refueler needs to be refilled at an airport fuel storage facility. Some refuelers have an elevation platform to allow refuelling staff to connect fuel hoses several feet above the ground. Fuel and electric powered models?

Figure A.8: *An example of a fuel truck [61].*

- **Dispenser Truck**
  A dispenser truck can be used to refuel an aircraft when an airport has an underground fueling system and employs hydrant trucks as "connectors" between the underground refuelling system and aircraft. Some dispenser trucks have an elevation platform to allow refuelling staff to connect fuel hoses several feet above the ground. Fuel and electric powered models?

Figure A.9: *An example of a hydrant dispenser truck [62].*

- **Maintenance vehicles**
  Various types of maintenance vehicles are used to provide aircraft maintenance service. The

vehicles are used by airport and/or airline employees to travel to/from maintenance facilities and an aircraft in need of maintenance and repairs.

Figure A.10: *An example of a maintenance vehicle [26].*

- **Deicer**
  Deicers are specially designed vehicles transport that heat, and spray deicing fluid on an aircraft exterior to reduce and/or eliminate ice and snow prior to departure. They are equipped with a hydraulic boom and bucket system to allow the operator to reach different parts of an aircraft. The vehicles sometimes contain two engines: one to power the vehicle and one to heat the deicing fluids.

Figure A.11: *An example of a deicer vehicle [82].*

A.1.4. Passenger Handling Equipment

- **Boarding Stairs**
  Boarding stairs allow the loading and unloading passengers at remote aircraft stands and in the absence of PBBs. They can be moved by towing, pushing, or they can be fixed to a truck / driving system. The height can be adjusted to align the stairs with the aircraft door. A disadvantage of the boarding stairs is that it is not possible to facilitate deplaning of passengers that need special care (especially passengers in a wheelchair).

Figure A.12: *An example of mobile boarding stairs [81].*

- **Disabled Passenger Lift Truck**
  Disabled passenger lift trucks allow the loading and unloading disabled passengers (especially passengers in a wheelchair) at remote aircraft stands and in the absence of PBBs. Just like a catering truck, the vehicle uses a scissor lift elevation system to raise its "container" to the level of an aircraft door. At the rear of the vehicle is a loading ramp to get in and out of the "container" from the apron.

Figure A.13: *An example of disabled passenger lift truck [55].*

- **Bus**
  Buses may be used to transport passengers and crew from a remote stand to the terminal. Buses can also be used to transport passengers from the main terminal to remote "satellites" on the apron of a large airport.

Figure A.14: *An example of a bus [83].*

A.1.5. Baggage and Cargo Handling Equipment

- **Baggage/cargo tug**
  Baggage/cargo tugs are used to transport baggage and cargo between an aircraft and the airport terminal and/or the processing/sorting facilities. Depending on the amount of baggage and cargo, several tugs are needed to handle an aircraft. Furthermore, depending on the weight to be towed, a distinction is made between baggage and cargo tugs. Baggage tugs can

handle baggage and lightweight cargo, whereas cargo tugs can also be used for heavy cargo handling and sometimes even "light" aircraft pushback operations.

The baggage tug is one type of GSE that is already electrically executed. In this way, exhaust fumes in the "basement" of the airport terminal, where the baggage handling facility is often located, are avoided. For the cargo tug, many diesel models are still available.

(a) *An example of a baggage tug [81].* **(b)** *An example of a cargo tug [81].*

Figure A.15: *Examples of a baggage and cargo tug. The main difference is in the maximum weight that can be towed.*

The baggage/cargo tugs can pull several dollies connected together. A distinction is made between three types of dollies: Baggage dolly; Unit Load Device (ULD) dolly; Pallet dollies.

(a) *An example of a baggage dolly [81].* **(b)** *An example of a ULD dolly [81].* **(c)** *An example of a pallet dolly [81].*

Figure A.16: *Three different types of dollies that can be towed by baggage/cargo tug.*

- **Cargo transporter**
  Cargo transporters can transport pallets and containers between warehouses, aircraft loading equipment, dollies and similar equipment. They use rollers to get the cargo from or onto the vehicle.

Figure A.17: *An example of a cargo transporter [81].*

- **Belt loader**
  Belt loaders are used to load and unload baggage and cargo intro/from an aircraft. They have

a conveyor belt installed on a boom whose inclination can be adjusted to match the height of the aircraft cargo compartment. There are fuel and electric powered models available.

Belt loaders can be equipped with a power stow, which is an extension to facilitate the loading and offloading of passenger baggage and cargo in the aircraft cargo hold.

(a) *An example of a belt loader [81].*

(b) *An example of a belt loader equipped with a power stow [57].*

Figure A.18: *A belt loader and an example of a "power stow".*

- **Deck loaders**
  Deck loaders are used to load and unload the cargo that is within a ULD on an aircraft. A deck loader consists of a bridge and a loading platform. Both bridges use a scissor lift to adjust their height. The bridge platform acts as a bridge between the aircraft cargo compartment and the loading platform. The ULDs are moved between the ULD dollies and the loading platform. The loading platform is raised to the same height as the bridge platform, to enable the ULDs to be transferred on the level of the aircraft cargo compartment.

  The platforms of a deck loader are equipped with powered rollers to move the Unit Load Devices (ULDs). By using these rollers a ULD can move in two dimensions without human intervention.

  A distinction is made between lower deck loaders and main deck loaders. The basic difference is that the main deck loader is more versatile. The main deck loader can load cargo into the main (top) deck and the lower deck of a NABO/WIBO aircraft. A lower deck loader is used to load cargo into both decks of a NABO aircraft and the lower deck of a WIBO aircraft. Main deck loaders usually have a greater load capacity than lower-deck loaders

(a) *An example of a lower deck loader [81].*

(b) *An example of a main deck loader [81].*

Figure A.19: *Two different types of cargo loaders.*

B

Technical Background on Batteries and Charging

This appendix provides additional technical background on battery technologies and charging systems. The main text only uses the parts that are directly relevant for the simulation model: SOC, battery capacity, charging thresholds, charging power, charging duration, and charger capacity. The additional technical material in this appendix supports the modelling choices but is not part of the active simulation logic.

B.1. Battery Technologies

A battery is one of the most widely used energy storage devices in electric vehicle and power system applications. It stores chemical energy and converts it into electrical energy when required. Several battery chemistries have been developed for transport applications. According to Li, three main battery families are commonly used in road vehicles: lead-acid batteries, nickel-based batteries, and lithium-based batteries. Less common technologies include sodium-sulphur, metal-air, and flow batteries [46]. Each battery technology has different advantages, limitations, and cost implications. A technical comparison is provided in Table B.1 and Table B.2, while Figure B.1 shows the discharge voltage behaviour of several rechargeable battery chemistries.

Lead-acid batteries are among the oldest rechargeable electrochemical batteries and are used in both household and commercial applications. Their main advantages are low capital cost, relatively high efficiency, fast response, and low self-discharge, typically around 2% of rated capacity per month. However, lead-acid batteries have low specific energy density, relatively limited cycle life, and environmental disadvantages during production and disposal. These limitations restrict their wider use in modern electric transport applications [46].

Nickel-based batteries include nickel-iron, Nickel–Cadmium (NiCd), Nickel–Metal Hydride (NiMH), and Nickel–Zinc (NiZn) batteries. Nickel-iron batteries are stable, durable, and relatively low-cost, but they have low power density, high self-discharge, high weight, and relatively high maintenance requirements. NiCd batteries can charge quickly, tolerate overcharge and deep discharge, and operate over a wide temperature range. However, they suffer from memory effect and contain toxic materials, which has strongly reduced their use. NiMH batteries have higher energy density and faster discharge capability than several other nickel-based batteries, but they generate considerable heat during charging. NiZn batteries are relatively safe and environmentally friendly, but their short cycle life limits commercial application. Overall, nickel-based batteries generally provide higher energy density and longer cycle life than lead-acid batteries, but they are more expensive and often have lower efficiencies [46].

Lithium-based batteries are currently the most relevant battery technologies for electric transport applications. They are attractive because of their high energy density, low weight, low self-discharge, absence of memory effect, long lifetime, and ability to support fast charging [90]. Common lithium-based batteries include lithium-ion, Lithium-Polymer (LiPo), Lithium Nickel Manganese Cobalt Oxide (NMC), and Lithium Iron Phosphate (LFP). LFP batteries are relatively expensive, but they have high power density and a long cycle life of more than 3000 cycles. NMC batteries combine high energy density, relatively high cycle life, and comparatively low cost. LiPo batteries have good reliability and robustness, but their conductivity and power density are lower. Lithium-ion batteries provide strong overall cost-performance because they combine high specific energy density, high power density, high efficiency, low self-discharge, and moderate cost. However, their lifetime can decrease under high temperatures and deep discharge, and protection systems are required to ensure safe operation [46, 32, 52].

The efficiencies reported in Table B.1, Table B.2, and Figure B.1 refer to battery round-trip efficiency. This represents the percentage of charging energy that can later be discharged. Round-trip efficiency is important because it determines how effectively a battery can shift energy demand over time and influences the economic viability of the storage system [52]. When the efficiencies of the battery, electric motor, recharging system, and regenerative braking are considered together, the total energy efficiency of BEVs is approximately 60% to 70% [46].

Table B.1: *Comparison of what battery technologies can deliver [90, 31, 46, 91, 10, 1, 51, 36, 60, 32, 52]*

Battery type	Nominal voltage (V)	Energy density (Wh/kg)	Volumetric energy density (Wh/L)	Specific power (W/kg)	Operating temperature (°C)
Lead acid (Pb-acid)	2.0	35	100	180	−15 to +50
Nickel-iron	1.2	40–60	60–110	20–100	−10 to +45
Nickel-cadmium (Ni-Cd)	1.2	50–80	300	200	−20 to +50
Nickel-metal hydride (Ni-MH)	1.2	70–95	180–220	200–300	−20 to +60
Nickel-zinc (Ni-Zn)	1.65	60	150	100–300	−20 to +60
Lithium-ion (Li-ion)	3.6	118–250	200–400	500–2000	−20 to +60
Lithium-ion polymer (LiPo)	3.7	130–225	200–250	260–450	−20 to +60
Lithium-iron phosphate ($LiFePO_4$)	3.2	100–140	220	2000–4500	−45 to +70
Lithium-nickel manganese cobalt ($LiNi_xMn_yCo_zO_2$)	3.6	150–280	300–700	1000–3000	−20 to +60

Table B.2: *Comparison of performance of battery technologies [90, 31, 46, 91, 10, 1, 51, 36, 60, 32, 52]*

Battery type	Efficiency (%)	Life cycle (# cycles)	Self discharge (%/month)	Memory effect	Production cost ($/kWh)
Lead acid (Pb-acid)	85	1000	< 5	No	60
Nickel-iron	80	> 3000	20	No	< 200
Nickel-cadmium (Ni-Cd)	75	2000	10	Yes	250–300
Nickel-metal hydride (Ni-MH)	80	< 3000	20	Rarely	200–250
Nickel-zinc (Ni-Zn)	70	< 1000	< 40	No	100–200
Lithium-ion (Li-ion)	93	> 2000	< 5	No	200
Lithium-ion polymer (LiPo)	92	> 1200	< 5	No	150
Lithium-iron phosphate ($LiFePO_4$)	92	> 3000	< 5	No	350
Lithium-nickel manganese cobalt ($LiNi_xMn_yCo_zO_2$)	81	< 3500	< 5	No	80–145

Each battery type has its own discharge voltage curve. Figure B.1 shows the discharge voltage curves of several rechargeable battery chemistries as a function of the SOC. Lithium-ion batteries operate at a relatively high voltage, while lead-acid and nickel-based batteries generally have lower nominal voltages. Several battery types also show a long, nearly horizontal voltage plateau, meaning that voltage remains almost constant during a large part of the discharge process. A higher voltage is beneficial because, for the same power demand, it results in a lower current. This reduces electrical losses and heat generation. It also means that fewer cells are required in series to reach the desired system voltage, which can simplify the system design. Therefore, high-voltage battery chemistries are generally more suitable for high-power electric vehicle applications [63, 31].

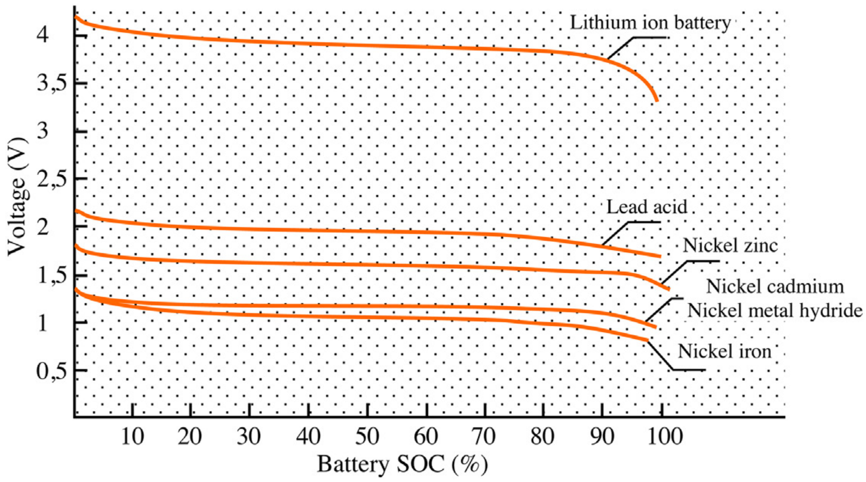


Figure B.1: *Discharge voltage graphs of rechargeable battery cells [10].*

B.2. Internal Structure of LFP Batteries

An LFP battery operates through the movement of lithium ions between the positive and negative electrodes. During charging, lithium ions move from the positive electrode to the negative electrode, while electrons move through the external circuit. During discharging, the process is reversed, allowing electrical energy to be supplied to the vehicle.

The main components of an LFP cell are the positive electrode, negative electrode, current collectors, electrolyte, and separator. The positive electrode is based on LFP, while the negative electrode is usually made of graphite. The electrolyte allows lithium ions to move between the electrodes. The separator prevents direct contact between the electrodes while still allowing ion transport. Current collectors transfer the generated electrical current to the external circuit [93].

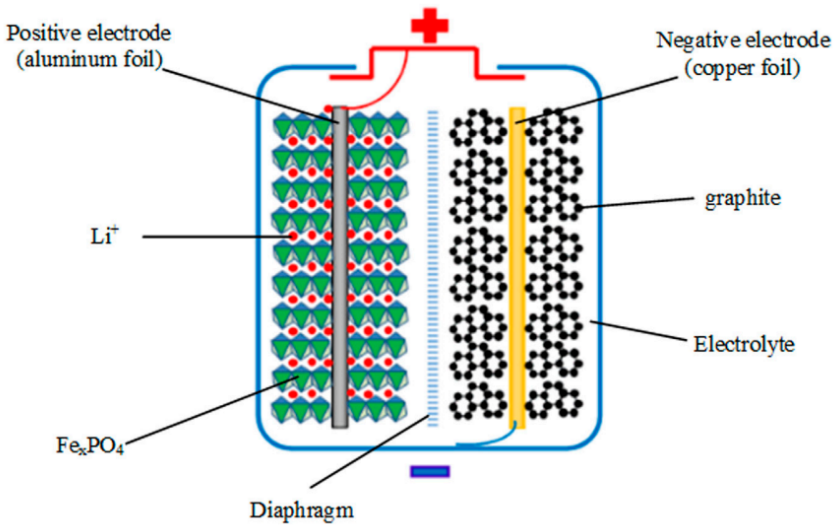


Figure B.2: *Internal structure of an LFP battery [93].*

A Battery Management System (BMS) is required for the safe and reliable operation of lithium-based batteries. It monitors battery states using measurements such as voltage, current, and temperature. It also controls protection mechanisms to avoid overcharging, deep discharge, overheating, and cell imbalance. In multi-cell battery packs, the BMS supports cell balancing and thermal management, both of which are important for safety, efficiency, and battery lifetime [58, 31].

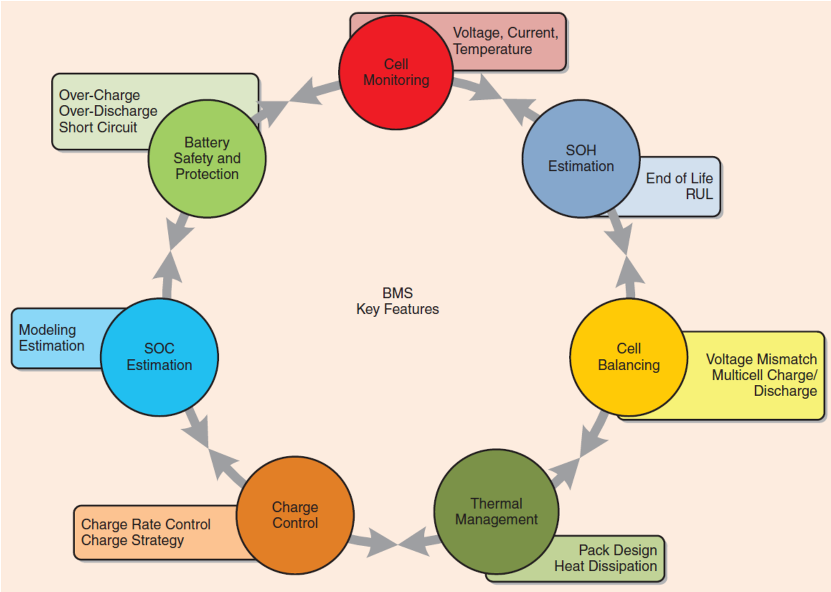


Figure B.3: *BMS features [58].*

The State of Charge (SOC) is one of the most important battery indicators for operational use. It indicates the remaining available energy in the battery. In a simplified form, SOC can be written as [31]:

$$SOC = \frac{C_{res}}{C_n} = 1 - \frac{C_d}{C_n} = 1 - \frac{\int_t I dt}{C_n}, \tag{B.1}$$

where C_{res} is the residual capacity, C_n is the nominal capacity, C_d is the drawn capacity, and I is the battery current. In practice, SOC cannot be measured directly because battery behaviour is nonlinear and affected by ageing, temperature, and charge–discharge history [31].

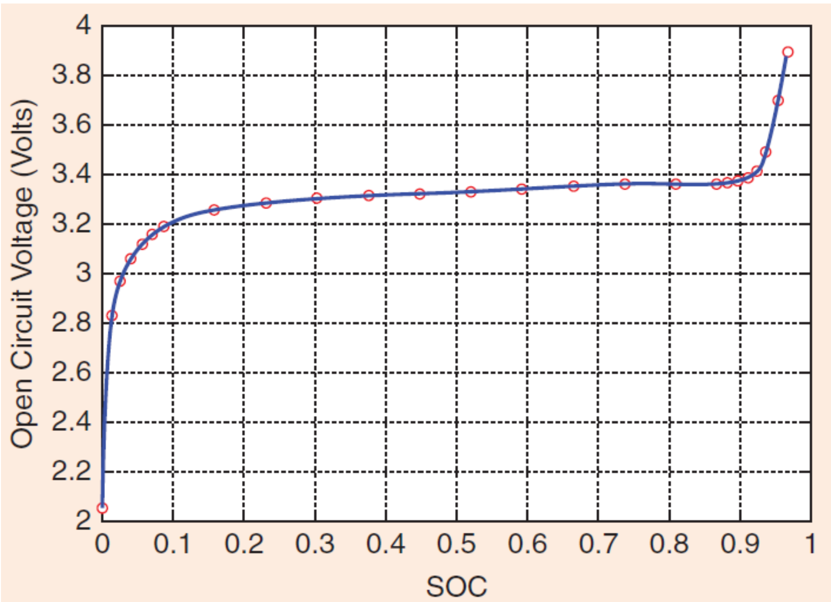


Figure B.4: *Open-Circuit Voltage (OCV)–SOC curve of LFP [58].*

The State of Health (SOH) describes the condition and ageing of the battery. It is often related to the remaining usable capacity of the battery. A common simplified definition is:

$$SOH = \frac{Q_{act}}{Q_R} \cdot 100 \tag{B.2}$$

where Q_R is the rated capacity and Q_{act} is the actual available capacity after degradation. This definition is useful, but incomplete, because battery health also depends on operating history, temperature exposure, charging behaviour, and the specific application [31]. For this reason, SOH is relevant for long-term asset management, but it is not explicitly modelled in the daily operational simulation in this thesis.

B.3. Charging Technologies

BEVs, including eGSE, are charged from the electrical grid. The grid supplies Alternating Current (AC), while batteries store energy as Direct Current (DC). A charger therefore converts AC into DC before the energy is stored in the battery. In fast-charging systems, additional power electronics may be used to control the voltage and current supplied to the battery [90].

Different charging methods can be used. The most common methods are Constant Current (CC), Constant Voltage (CV), Constant Power (CP), taper charging, trickle charging, and combined CC/CV charging. In CC charging, the current is kept constant while the battery voltage increases. In CV charging, the voltage is kept constant while the current decreases. In CP charging, the charging power is held approximately constant. Taper charging and trickle charging are simpler approaches, but they are less central for modern fast charging of lithium-based batteries [90].

For lithium-based batteries, CC/CV charging is commonly used. The process starts with a CC phase, where the battery receives a constant current until a voltage limit is reached. It then switches to a CV phase, where the voltage is kept constant and the current gradually decreases. This means that charging from a low SOC is usually faster than charging near full capacity [90, 75, 34].

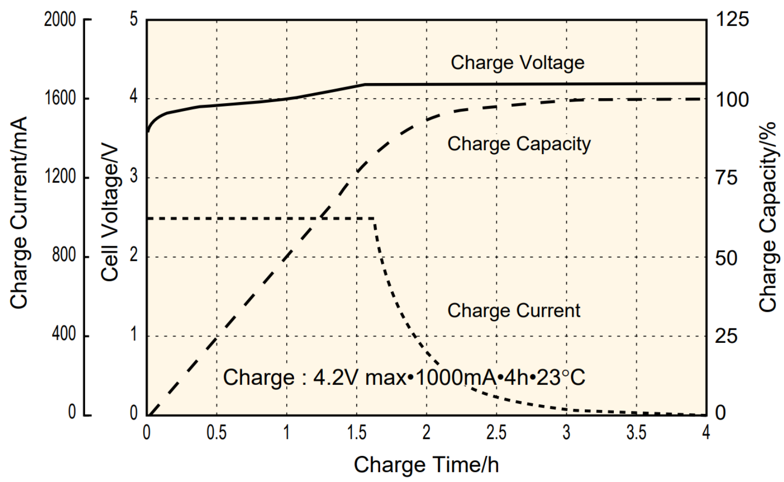


Figure B.5: *CC–CV charging profile [76].*

Charging systems can also be classified as on-board or off-board. In on-board charging, the charger is located inside the vehicle. This provides flexibility, but the charging power is limited by vehicle size, weight, and cost constraints. In off-board charging, the charger is located in the charging infrastructure. This allows higher charging power, but requires more expensive and more complex infrastructure [22].

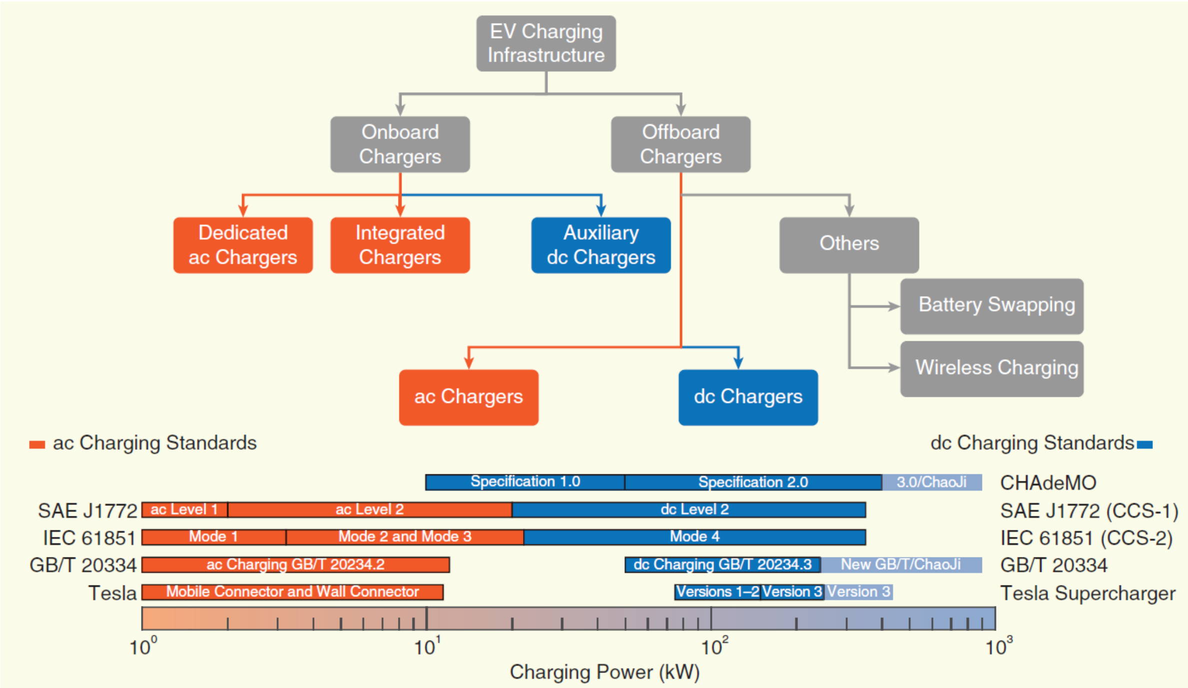


Figure B.6: *Charging standards and power ranges [63].*

Charging systems are commonly grouped into AC and DC charging levels. AC charging relies on the vehicle's on-board charger to convert grid power to DC. DC charging uses off-board conversion and supplies DC power directly to the battery. DC charging can therefore reach higher charging powers than most AC charging systems. However, actual charging time always depends on the battery size, initial and final SOC, allowable charging power, charging efficiency, and the charging curve [22, 63].

Table B.3: *Overview of EV charging levels and indicative power ratings*

Charging level[a]	Indicative rating	Charger type	Typical interpretation
AC charging			
AC level 1	120 V up to 1.9 kW	On-board	Low-power charging, mainly suitable for long parking periods or overnight charging.
AC level 2	240 V up to 19.2 kW	On-board	Medium-power charging. The effective charging power is limited by both the charging point and the vehicle's on-board charger.
AC level 3[b]	>20 kW	On-board	High-power AC charging category proposed in some classifications, but not consistently standardised.
DC charging			
DC level 1	200–450 V_{DC} up to 36 kW	Off-board	Lower-power DC charging. Grid power is converted to DC outside the vehicle.
DC level 2	200–450 V_{DC} up to 90 kW	Off-board	Fast charging. Suitable for applications where vehicle idle time is limited.
DC level 3	200–600 V_{DC} up to 240 kW	Off-board	High-power fast charging. Charging time depends strongly on battery size, allowable charging power, and the charging curve.

Note. [a] Charging level definitions may vary by region, standard, and source. [b] AC level 3 is not consistently finalised or applied across standards.

B.4. Charging Strategies

Charging strategies can be classified by timing, coordination, and grid interaction. In terms of timing, overnight charging uses long idle periods outside operational hours, while opportunity charging uses shorter charging sessions during idle periods between tasks. Delayed charging postpones charging to reduce peak electricity demand and can also result from limited charging infrastructure. Full charging keeps the vehicle connected until the battery is full, while partial charging allows shorter and more flexible charging sessions [80, 19, 39, 92, 54].

In terms of coordination, uncoordinated charging starts immediately when a vehicle is connected to a charger. This is simple, but can lead to high peak loads and congestion at chargers. Coordinated charging actively manages charging to reduce grid impact, operational conflicts, or electricity costs. Priority charging ranks vehicles based on operational criteria such as SOC, availability, vehicle type, or expected task demand [30, 22, 90].

One example of priority charging ranks vehicles based on SOC and availability [30]:

$$\text{Rank} = (1 - SOC) \cdot W_{SOC} + (1 - A) \cdot W_{Avail}, \tag{B.3}$$

where W denotes weighting factors and A denotes vehicle availability. Availability can be expressed as:

$$A = \frac{\text{Number of vehicles in class } X \text{ with } SOC > 0.5}{\text{Total number of vehicles in class } X}. \tag{B.4}$$

This type of rule can help prioritise vehicles with low SOC or high operational relevance. However, it does not necessarily account for future tasks unless these are explicitly included in the control logic.

From a grid perspective, charging can be unidirectional or bidirectional. In Unidirectional Charging

(V1G), the charging power can be controlled, but energy only flows from the grid to the vehicle. In Bidirectional Charging/Vehicle-to-Grid (V2G), vehicles can also discharge electricity back to the grid. V2G can support grid services such as peak shaving, load levelling, or frequency regulation, but it also creates additional efficiency losses and may increase battery cycling [48, 90, 22, 86, 94]. Because this thesis focuses on operational fleet sizing and not on grid-service optimisation, V2G is not included in the simulation model.

C
AI Statement

C.1. AI Statement

For this report for the course ME54035, I have used Generative AI to:

- Obtain inspiration for the overall structure of the report.
- Improve the grammar, style, layout, and/or spelling of the text.
- Generative AI was used as a programming support tool during the development of the simulation model, including assistance with writing, debugging, restructuring, and documenting Python code. All generated suggestions and code fragments were reviewed, tested, and adapted by the author, who remains fully responsible for the final model implementation and results.

In all cases, I have critically reviewed and corrected the AI-generated content and remain fully responsible for the content of this report.