

Optimizing Air Cargo Operations

An Integrated Model for Fleet Assignment, ULD routing, and Cargo Assignment to ULDs

MSc. Thesis
Luuk Barbian

Optimizing Air Cargo Operations

An Integrated Model for Fleet Assignment, ULD
routing, and Cargo Assignment to ULDs

by

Luuk Barbian

to obtain the degree of Master of Science

at the Delft University of Technology,

to be defended publicly on Friday December 13, 2024 at 09:30 AM.

Student number: 4871979
Project duration: March 1, 2024 – December 13, 2024
Thesis committee: Ir. P.C. Roling, TU Delft, chair
Dr. A. Bombelli, TU Delft, supervisor
Dr. J. Sun, TU Delft, external

This thesis is confidential and cannot be made public until December 31, 2024.

An electronic version of this thesis is available at <http://repository.tudelft.nl/>.

Preface

This thesis marks the conclusion of my time at TU Delft, where I have spent six truly enjoyable and formative years. These years have been a blend of learning, challenges, and opportunities, all of which have shaped me both personally and professionally. Reflecting on this journey, I feel immense gratitude for the experiences that I believe have set me up for success in the years to come.

The research for this thesis has been both interesting and rewarding, and I am truly pleased with the final result. I could elaborate on the technical hurdles I faced, but I trust that the thesis itself will tell that story.

Instead, I want to take this opportunity to express my heartfelt thanks to those who made this journey possible. I would like to thank my supervisors, Alessandro and Felipe, for their invaluable guidance and support throughout the entire process. Your expertise and feedback have been instrumental in shaping this thesis. I also owe much to my friends and family. Whether it was through thoughtful advice, shared moments of laughter, or simply being there when it mattered most, your contributions have meant a great deal to me.

*Luuk Barbian
Delft, December 2024*

Summary

The air cargo market is defined by the transportation of high-value, often specialized goods, requiring precise handling. From a planning perspective, air cargo operations involve complex, nested decision-making across multiple levels. While previous research has addressed fleet assignment, cargo routing, and air cargo load planning as individual problems, few studies attempt a fully integrated approach. This work presents a comprehensive model that jointly optimizes fleet assignment, Unit Load Device (ULD) routing, and cargo allocation to ULDs within a full freighter network. We propose two solution methods: a sequential approach and an integrated approach. The sequential approach first performs a myopic fleet assignment and cargo routing, treating each aircraft as a single bin without assigning cargo to ULDs or routing ULDs. These tasks are then addressed in the second stage, ensuring compliance with bin-packing and compatibility constraints in the final result. In contrast, the integrated approach employs column generation to optimize all routing and packing decisions simultaneously. Both models were tested on synthetic instances based on a major European combination airline's freighter operations. Results indicate that the integrated approach consistently outperforms the sequential method, achieving profit increases of 6-24% across various operational scenarios. This improvement is driven by the ability to yield higher revenue while lowering ULD operating costs.

Contents

| | |
|---|-----------|
| Preface | i |
| Summary | ii |
| Nomenclature | iv |
| Motivation and Research Proposal | v |
| Scientific Paper | 1 |

Nomenclature

Abbreviations

| Abbreviation | Definition |
|--------------|------------------------------------|
| ULD | Unit Load Device |
| ACP | Aircraft Configuration Problem |
| APP | Air Cargo Palletization Problem |
| ALNS | Adaptive Large Neighborhood Search |
| BSP | Build-up Scheduling Problem |
| WBP | Weight and Balance Problem |
| PDP | Pickup and Delivery Problem |
| TSP | Traveling Salesman Problem |
| TSN | Time-Space Network |
| MILP | Mixed Integer Linear Programming |
| CG | Column Generation |
| RMP | Restricted Master Problem |
| AMS | Amsterdam |
| BOG | Bogota |
| CAI | Cairo |
| EZE | Buenos Aires |
| GUA | Guatemala City |
| HRE | Harare |
| JNB | Johannesburg |
| MIA | Miami |
| NBO | Nairobi |
| SCL | Santiago |
| UIO | Quito |
| VCP | Campinas |
| OD | Origin-Destination |
| EU | Europe |
| AF | Africa |
| LA | Latin-America |
| NA | North-America |

Motivation and Research Proposal

The growth in global air cargo, driven by the rise of e-commerce, expanding global trade, and increased demand for timely transportation of high-value or perishable goods, has led to a heightened focus on optimizing air cargo operations. Air cargo represents less than 1% of global trade by volume, while it accounts for approximately 35% of trade value due to its concentration on high-priority commodities, such as perishables, pharmaceuticals, and electronics (Boeing 2022). Airlines transport cargo either in the belly of passenger planes or on dedicated freighter aircraft. While widebody passenger planes significantly supported the growth of cargo operations pre-pandemic, dedicated freighters are expected to continue handling over half of global air cargo volume (Boeing 2022). These freighters are especially suited for specialized goods, such as temperature-sensitive or hazardous items, and operate in dedicated networks that require careful planning of loads and routes.

From a planning perspective, air cargo operations involve a nested structure of decision-making across strategic, tactical, and operational levels. Research in fleet assignment, cargo routing, and air cargo load planning has thus far addressed individual aspects of these decisions but rarely has attempted to integrate them comprehensively. For instance, studies such as Li et al. (2007) and Derigs and Friederichs (2013) highlight the complexity of combining fleet assignment with cargo routing while managing aircraft capacity constraints. Additional work has focused on recovering disrupted cargo schedules (Delgado et al. 2020) and optimizing connection times by integrating aircraft tail assignment with cargo routing (Xiao et al. 2022). However, these studies simplify air cargo load planning by treating the aircraft as a single bin with volume or weight constraints.

Treating the aircraft as a single bin oversimplifies the problem, as in reality, load planning requires consideration of bin-packing and the compatibility constraints of cargo and Unit Load Devices (ULDs). Different types of cargo, such as hazardous materials and perishables, may be incompatible and cannot be packed together in the same ULD. Furthermore, specific cargo types, like temperature-sensitive goods, require specialized temperature-controlled ULDs. Thus, effective load planning should integrate decisions on both the appropriate type of ULDs (addressed in the Aircraft Configuration Problem, or ACP) and the optimal arrangement of cargo shipments within these ULDs (the Air Cargo Palletization Problem, or APP) (Brandt and Nickel 2019).

The APP is a classical bin-packing problem, frequently modeled in two or three dimensions, and takes into account important packing considerations, such as cargo type compatibility and ULD capacity. However, existing studies of the APP typically focus on individual flight legs and fail to incorporate network-level routing or multi-aircraft scenarios (Paquay, Schyns, et al. 2016; Paquay, Limbourg, et al. 2018). The lack of overlap between the integration of network cargo routing and air cargo load planning creates a gap in the literature, as air cargo operations involve complex networks with multiple flight legs and diverse aircraft types, where routing and scheduling decisions should work together with load planning.

This thesis aims to fill this gap by proposing an integrated model that combines fleet assignment, ULD routing, and cargo shipment allocation to ULDs within a full freighter network. The integration of these decisions allows for the consideration of load planning and its associated compatibility constraints directly within the cargo routing process. Without this approach, fleet assignment and cargo routing decisions may allocate a full set of cargo shipments to a flight, only to require offloading at a later stage due to unforeseen compatibility violations. Furthermore, fleet assignment decisions themselves may need to adjust based on compatibility demands, as certain routes with high demand for incompatible or specialized cargo types might require changes in ULD capacity.

Specifically, our work is structured around two research questions:

1. How can we develop an integrated model that effectively combines fleet assignment, ULD routing, and cargo allocation to ULDs, with the goal of maximizing revenue and minimizing associated costs, while ensuring computational efficiency?
2. Under what conditions can we demonstrate that the integrated model outperforms sequentially solving the subproblems in terms of overall solution quality and computational efficiency?

In this context, ULD routing refers to determining optimal paths for ULDs across multiple flight legs, while cargo allocation to ULDs involves assigning cargo shipments to specific ULDs based on compatibility and other bin-packing constraints. To address the research questions, the thesis introduces two formulations: a fully arc-based model and a hybrid model that combines arc-based aircraft routing with path-based ULD and cargo routing. These formulations are evaluated using two solution approaches: sequential and integrated. The sequential approach consists of two stages. The first stage handles fleet assignment and cargo routing while treating the aircraft as a single bin with basic weight and volume constraints. In the second stage, bin-packing and compatibility constraints are introduced, addressing the ULD routing and assigning cargo requests to specific ULDs. This simpler sequential approach allows for an assessment of whether the added complexity of an integrated model is truly necessary. The integrated approach, in contrast, solves these subproblems simultaneously, capturing their interdependencies. Integrating these decisions presents considerable computational challenges, given that a bin-packing problem alone is NP-hard. To maintain computational feasibility, certain aspects of the problem definition will be simplified, and a fast heuristic based on the hybrid arc-path formulation will be implemented.

References

- Boeing (2022). *Boeing website: World Air Cargo Forecast 2022-2041*. https://www.boeing.com/content/dam/boeing/boeingdotcom/market/assets/downloads/Boeing_World_Air_Cargo_Forecast_2022.pdf. Accessed: August 2024.
- Brandt, Felix and Stefan Nickel (2019). "The air cargo load planning problem - a consolidated problem definition and literature review on related problems". In: *European Journal of Operational Research* 275.2, pp. 399–410. ISSN: 0377-2217. DOI: <https://doi.org/10.1016/j.ejor.2018.07.013>.
- Delgado, Felipe et al. (2020). "Recovering from demand disruptions on an air cargo network". In: *Journal of Air Transport Management* 85, p. 101799. DOI: <https://doi.org/10.1016/j.jairtraman.2020.101799>.
- Derigs, Ulrich and Stefan Friederichs (2013). "Air cargo scheduling: integrated models and solution procedures". In: *OR spectrum* 35.2, pp. 325–362. DOI: <https://doi.org/10.1007/s00291-012-0299-y>.
- Li, D et al. (2007). "Simultaneous fleet assignment and cargo routing using benders decomposition". In: *Container Terminals and Cargo Systems: Design, Operations Management, and Logistics Control Issues*, pp. 315–331.
- Paquay, Célia, Sabine Limbourg, et al. (2018). "A tailored two-phase constructive heuristic for the three-dimensional Multiple Bin Size Bin Packing Problem with transportation constraints". In: *European Journal of Operational Research* 267.1, pp. 52–64. DOI: <https://doi.org/10.1016/j.ejor.2017.11.010>.
- Paquay, Célia, Michaël Schyns, et al. (2016). "A mixed integer programming formulation for the three-dimensional bin packing problem deriving from an air cargo application". In: *International Transactions in Operational Research* 23.1-2, pp. 187–213. DOI: <https://doi.org/10.1111/itor.12111>.
- Xiao, Fan et al. (2022). "Integrated aircraft tail assignment and cargo routing problem with through cargo consideration". In: *Transportation Research Part B: Methodological* 162, pp. 328–351.

Scientific Paper

Optimizing Air Cargo Operations: An Integrated Model for Fleet Assignment, ULD Routing, and Cargo Assignment to ULDs

Luuk Barbian^a

^a*MSc. Student, Air Transport and Operations, Faculty of Aerospace Engineering, Delft University of Technology*

Abstract

The air cargo market is defined by the transportation of high-value, often specialized goods, requiring precise handling. From a planning perspective, air cargo operations involve complex, nested decision-making across multiple levels. While previous research has addressed fleet assignment, cargo routing, and air cargo load planning as individual problems, few studies attempt a fully integrated approach. This work presents a comprehensive model that jointly optimizes fleet assignment, Unit Load Device (ULD) routing, and cargo allocation to ULDs within a full freighter network. We propose two solution methods: a sequential approach and an integrated approach. The sequential approach first performs a myopic fleet assignment and cargo routing, treating each aircraft as a single bin without assigning cargo to ULDs or routing ULDs. These tasks are then addressed in the second stage, ensuring compliance with bin-packing and compatibility constraints in the final result. In contrast, the integrated approach employs column generation to optimize all routing and packing decisions simultaneously. Both models were tested on synthetic instances based on a major European combination airline's freighter operations. Results indicate that the integrated approach consistently outperforms the sequential method, achieving profit increases of 6-24% across various operational scenarios. This improvement is driven by the ability to yield higher revenue while lowering ULD operating costs.

Keywords: Air cargo routing; Bin packing; Fleet assignment; ULD routing; Column generation

1 Introduction

The global air cargo market is experiencing significant growth, with traffic expected to increase at an annual rate of 4.1% from 2022 to 2041 (Boeing, 2022). This expansion is driven by e-commerce growth, increased global trade, and the need for reliable transport of high-value, time-sensitive, or specialized goods. Although air cargo comprises less than 2% of global trade by volume, it represents approximately 35% of trade value due to its focus on high-priority commodities, including perishables, pharmaceuticals, and electronics (Boeing, 2022).

Air cargo transport is carried out via dedicated freighters or passenger aircraft belly holds, both essential to global logistics networks. Dedicated freighters are often preferred on high-demand routes, for their palletized capacity and for shipments requiring specific handling capabilities, such as temperature-sensitive or hazardous cargo (Boeing, 2022). While widebody passenger aircraft helped cargo capacity grow before the pandemic, freighters are projected to maintain at least 50% of global air cargo volume, even as passenger traffic returns to pre-pandemic levels (Boeing, 2022). With passenger belly capacity now at 2019 levels, its future growth could influence dedicated freighter demand. However, currently, the growth of belly capacity has slowed, while freighter capacity saw a modest 5.0% year-on-year increase in August (IATA, 2024).

In passenger travel, all passengers occupy a similar amount of space and are singularly assigned to a bin (i.e., the selected seat). In contrast, cargo shipments vary widely in weight, volume, and dimensions, making their handling more complex. In addition, in recent years, the air cargo *cold chain* has gained momentum (Baxter and Kourousis, 2015). This supply chain focuses on transporting temperature-sensitive cargo, such as food or pharmaceutical products, that might need ad-hoc Unit Load Devices (ULDs) (ColdChain, 2021) capable of maintaining low enough temperatures, especially

during transshipment operations on the ground. For airlines operating in the cold chain, this translates into an additional set of requirements on the ULDs to have on-board. Furthermore, airlines must adhere to potential incompatibility requirements among different cargo commodities. For example, perishables and chemicals should not be packed in the same ULD to prevent contamination or deterioration (Brandt and Nickel, 2019).

Moving cargo across an airport network implies a nested and intertwined set of decisions and associated models ranging from long-term strategic decisions (e.g., fleet planning) through medium-term tactical decisions (e.g., route and frequency planning) to operational decisions. Feng et al. (2015) highlight four types of problems a cargo airline faces: revenue management, terminal operations, flight scheduling and fleet routing, and aircraft loading.

In our work, we deal with the last two problems: fleet routing and aircraft loading. In particular, we aim to integrate the following decisions: (1) fleet assignment in a full freighter network, (2) assignment of ULDs, i.e., containers and pallets that contain shipments, to flights, and (3) assignment of shipments to ULDs. Decision 1 belongs to the broader category of a multi-commodity flow problem, while decision 3 is a bin-packing problem in nature. Decision 2 combines aspects of both a multi-commodity flow problem and a bin-packing problem, as it involves routing and capacity considerations simultaneously.

The problem at hand is highly complex and cannot be solved in a fully integrated manner without simplifying certain aspects. For example, the three-dimensional bin packing problem is known to be NP-hard on its own (Brandt and Nickel, 2019). Consequently, prior research has primarily focused on either the routing problem by simplifying the bin-packing component or on the bin-packing problem by relaxing network routing. In the former approach, aircraft are typically treated as a single bin, where weight or volume limits must be met by the loaded shipments (Delgado et al., 2020; Delgado and Mora, 2021). In the latter case, a single flight leg is examined, allowing for a fully three-dimensional problem that is computationally tractable (Paquay et al., 2016, 2018). This paper seeks a balanced trade-off between these two approaches.

We model the assignment of shipments to ULDs in a one-dimensional fashion (using both volume and weight as packing constraints), considering shipment compatibility to restrict items that cannot be assigned to the same ULD and considering ULD compatibility for cargo (e.g., temperature-sensitive cargo can only be transported in temperature-controlled ULDs). Also, we model the assignment of ULD types to flights to satisfy the maximum number of transportable ULDs per type and per aircraft type. Furthermore, we include fleet assignment by designating an aircraft type to each flight leg. The schedule is fixed, and each flight leg is regarded as mandatory in principle. However, in a subsequent sensitivity analysis, this assumption will be relaxed, allowing all flight legs to be optional to evaluate the effects of flight selection decisions. Our overall goal is to maximize the revenue obtained from delivering shipments while minimizing the associated operational costs.

We propose two different formulations for our problem. The first is a fully arc-based formulation, while the second is a hybrid approach where we model aircraft routing with an arc-based formulation and the routing of ULDs and cargo with a path-based formulation. Moreover, we compare two different solution approaches: sequential and integrated. In the sequential approach, we first determine the routing of each aircraft and the routing of all cargo shipments, relaxing the assignment of cargo to ULDs. In the second phase, based on these decisions, shipments are assigned to specific ULDs while considering potential incompatibilities. In the integrated approach, all these decisions are taken simultaneously.

The remainder of this paper is organized as follows. We present a review of relevant works in Section 2. In Section 3 we introduce the problem setting, while Section 4 contains the two mathematical formulations of the problem at hand. Section 5 describes how the mathematical formulations are translated into two different solution approaches. In Section 6 we describe the test instance generation. Section 7 describes our computational experiments addressing accelerating strategies, solution quality, managerial insights, and a sensitivity analysis. Finally, Section 8 presents our conclusions and recommendations for future work.

2 Literature review

This literature review comprises two subsections. The first is pertaining to fleet and cargo routing decisions. In the second part we focus on the air cargo load planning, more specifically the Aircraft Configuration Problem (ACP) and the Air Cargo Palletization Problem (APP) as described by Brandt and Nickel (2019). For literature reviews concerning fleet assignment models we refer readers to Zhou et al. (2020) or Xu et al. (2024). Additionally, for a broader understanding of cargo operations, Feng et al. (2015) offer a comprehensive literature review.

2.1 Fleet assignment and cargo routing

Li et al. (2007) presents a model incorporating fleet assignment and cargo routing within the context of a major Asia-Pacific combination carrier. The model is hybrid since the fleet assignment is modeled using an arc-based formulation, while cargo routing is modeled with a path-based formulation. In this study the authors combine a full freighter and a passenger-cargo combination network, thus considering both passenger and cargo demands. A Benders' decomposition solution approach is proposed.

In the context of freighter airlines, Derigs et al. (2009) formulates two integrated models that combine the three planning steps: flight selection from a list of mandatory and optional flights, aircraft rotation planning, and cargo routing. The first model is hybrid because aircraft routing is arc-based, while cargo routing has a path-based formulation. In the second approach, both stages are modeled based on routes. Previous research predominantly considers aircraft capacity limitations based solely on weight, However, for this work, the consideration extends to encompass both weight and volume as capacity constraints. Finally, they propose a solution method based on column generation. Derigs and Friederichs (2013) extends on this work by including fleet assignment as a fourth planning step alongside addressing several other critical planning considerations for cargo airlines. These include factors such as available capacities on external flights, cargo handling costs and constraints, and adherence to aircraft maintenance regulations. A branch and price and cut approach is employed for solving the mathematical programs.

Delgado et al. (2020) proposes an arc-based formulation for their air cargo schedule recovery problem to react against cargo demand disruptions. This is a short-term planning problem, where decisions regarding flight rescheduling and the rerouting of cargo are made simultaneously. Delgado and Mora (2021) further extended on this work by proposing a matheuristic solution method, based on column generation in which each subproblem is solved using an ad-hoc heuristic.

Xiao et al. (2022) presents an integrated model for tail assignment and cargo routing, focusing particularly on through cargo connections where cargo remains on the same aircraft. This distinction enables shorter connection times for cargo that stays on board, compared to cargo requiring transfers. The authors introduce two models, a string-based and a connection-based one, and develop a column generation framework for the former. The models considers the capacity, in terms of the number of pallets to be transported.

Similarly to Xiao et al. (2022), the paper Zheng et al. (2023) delves into optimizing connection times within air cargo. Their study focuses on developing a model that combines identifying hub location, flight deployment, and flight-sequence-based cargo distribution to minimize the overall stay time of freighter cargoes at hub locations. This is modeled within a time-space network framework. The authors introduce the concept of the capacity matrix to reformulate the model and propose a matrix-based Adaptive Large Neighborhood Search (ALNS) heuristic for efficient solving.

2.2 Air cargo load planning

In a literature study on the air cargo load planning problem Brandt and Nickel (2019) defined it as consisting of four sub-problems: Aircraft Configuration Problem (ACP), Build-up Scheduling Problem (BSP), Air Cargo Palletization Problem (APP), and Weight and Balance Problem (WBP). As aforementioned for this work we are interested in the ACP and APP.

The existing research into the ACP is sparse. In Yan et al. (2006), the authors present a model with the objective to minimize the total ULD handling cost across various cargo origins and a hub within an Asian air express network. The model aims to determine the optimal allocation of cargo

streams into different types of ULDs. They put an emphasis on the early consolidation of cargo destined for the same final destinations to minimize subsequent handling. The model is formulated as a non-linear MIP which is then linearized to solve using a commercial solver. Notably, only two types of ULDs are considered, with one type designated for the main deck and the other for the lower deck positions. This restriction results in a single aircraft configuration where ULD types cannot be swapped.

Additionally, certain studies, such as Mongeau and Bes (2003), Vancroonenburg et al. (2014) and Zhao et al. (2021), primarily focus on the WBP, but also integrate aspects of the ACP. In these problems cargo is already distributed across various ULD types, resulting in the weight of each ULD to be known. The task is to assign these ULDs to loading positions within the aircraft, while satisfying stability and safety requirements. Each predefined position accommodates a specific set of allowable ULD types. Since these studies consider multiple ULD types and more ULDs than loading positions, the models also determine the aircraft configuration. Different configurations are considered for the predefined positions accommodating various types of loaded ULDs.

Mesquita and Sanches (2024) addresses air cargo load planning combined with routing, pickup and delivery. The study particularly focuses on the APP in a one-dimensional context, considering volume and weight constraints alongside the WBP. Additionally, the work incorporates two additional sub-problems: simultaneous pickup and delivery at each node (PDP) and optimizing the route akin to the Traveling Salesman Problem (TSP). The authors devised a solution strategy using historical transport data from Brazilian hub networks and conducted experiments using a commercial solver, five established meta-heuristics, and a novel heuristic developed specifically for this problem. While the model is solved for a single aircraft scenario, it accounts for multiple flight legs.

Given the NP-hard nature of the APP, the literature has mainly relied on developing different heuristics to solve realistic instances. Paquay et al. (2016) propose different heuristics for the three-dimensional bin-packing problem in air cargo. The authors solve instances of up to 80 shipments in less than an hour. Paquay et al. (2018) develop a two-phase heuristic capable of solving instances up to 100 boxes with computational times not exceeding 12 seconds. The first phase deals with packaging the boxes in identical bins, while the second phase generates different loading patterns considering the different types of containers available. The solution with the minimum used volume is selected. Gueret et al. (2003), Heidelberg et al. (1998) and Nance et al. (2011) propose different heuristics for the two-dimensional bin-packing problem in air cargo in the context of military operations. All the works mentioned above have focused on developing reasonable solutions in a short time, considering a single flight leg exclusively.

2.3 Contribution of the work

Our work complements the literature by integrating different planning decisions. In particular, we consider the following to be our main contributions:

- To the best of our knowledge, we integrate for the first time **fleet assignment**, **ULD routing** (assignment of ULDs to flights), and **cargo allocation to ULDs** in a mathematical model.
- we explore the **conditions** under which integrating these decisions yields benefits compared to addressing these problems sequentially. To do this, we develop two solution algorithms, one **integrated** and one **sequential**, and offer **managerial insights** depending on the nature of the cargo transported.

3 Problem setting

The modeled problem is a multi-commodity flow problem where three commodities, namely aircraft types, ULDs, and cargo items, move across a network. We proceed by first introducing the basic properties of the network and the three commodity types. The network, as visualized in Figure 1, is a Time-Space Network (TSN) that is built using a set of flight arcs \mathcal{F} , indexed by f , as the main input. Each flight arc $f \in \mathcal{F}$ is characterized by two nodes in the TSN, representing respectively the origin airport $l_f^{o,\mathcal{F}}$ and departure time DT_f and the destination airport $l_f^{d,\mathcal{F}}$ and

arrival time AT_f . The unique airports across all flight arcs define the set of airports \mathcal{L} , indexed by l , part of the TSN: each $f \in \mathcal{F}$ originates from node $(l_f^{o,\mathcal{F}}, DT_f) \in \mathcal{N}$ and terminates in node $(l_f^{d,\mathcal{F}}, AT_f) \in \mathcal{N}$, where \mathcal{N} is the set of nodes forming the TSN, indexed by n , whose cardinality is $|\mathcal{N}| = 2|\mathcal{F}|$. Two nodes $n_1 = (l_1, t_1), n_2 = (l_2, t_2) \in \mathcal{N}$ such that $l_1 = l_2$ and $t_2 > t_1$ are connected with a ground arc g , with the full set of ground arcs being \mathcal{G} .

The fleet is composed of a set of aircraft types \mathcal{K} , indexed by k . The number of aircraft of type k available is defined as N_k , and the model can be flexible in the way aircraft are initially positioned in the TSN. To achieve this goal, a source node s and a sink node t are added “before” and “after” the TSN, so that the full set of nodes where aircraft can be dispatched is $\mathcal{N}_{\mathcal{K}} = \mathcal{N} \cup \{s\} \cup \{t\}$. s is connected to the TSN with a set of arcs in the form $a = (l, \hat{t}_l)$, where (l, \hat{t}_l) defines the first node $n \in \mathcal{N}$ available, temporally, in airport l . The set containing such arcs is \mathcal{E} . Similarly, the TSN reconnects to t with a set of arcs in the form $a = (l, \bar{t}_l)$, where (l, \bar{t}_l) defines the last node $n \in \mathcal{N}$ available, temporally, in airport l . The set containing such arcs is \mathcal{O} . Finally, s and t are directly connected with a by-pass arc a_{bp} .

The model can be configured with full flexibility, creating a source and sink connection for every airport. With such an approach, we allow the model to position and mix aircraft in the most suitable way in each airport $l \in \mathcal{L}$ at the start of the planning period by “injecting” them into the TSN via arcs $a \in \mathcal{E}$ and to perform fleet assignment to ensure that all aircraft eventually reach t via arcs $a \in \mathcal{O}$. Alternatively, the model can be set up to allow aircraft to enter or exit the TSN only through specific airports. This might be used to model operational constraints or to reflect strategic decisions. Additionally, we model the possibility of not using part of the fleet with the by-pass arc a_{bp} . Note that without a_{bp} , the model could still allow unused aircraft by assigning them to a specific airport and forcing them to use a sequence of ground arcs $g \in \mathcal{G}$ from s to t : this option entails sets of symmetrical solutions that the by-pass arc prevents instead. Finally, because of aircraft range limitations, not all aircraft types might be able to fly every flight arc (even without any payload). Hence, we define \mathcal{K}_f as the subset of aircraft types capable of operating flight arc $f \in \mathcal{F}$.

We consider a set of ULD types \mathcal{V} , indexed by v , each with distinct characteristics. The overall set of ULDs is \mathcal{U} , whose cardinality is $|\mathcal{U}| = \sum_{v \in \mathcal{V}} N_v$, with N_v being the number of ULDs of type v available. Defining \mathcal{U}_v the subset of ULDs of type v , it follows that $\mathcal{U} = \mathcal{U}_1 \cup \dots \cup \mathcal{U}_{|\mathcal{V}|}$. Each ULD $u \in \mathcal{U}$ is characterized by a release time $RT_u^{\mathcal{U}}$ in the origin airport $l_u^{o,\mathcal{U}}$ and a delivery deadline $DD_u^{\mathcal{U}}$ in the destination airport $l_u^{d,\mathcal{U}}$. We translate such an information into two nodes, respectively $s_u^{\mathcal{U}} = (l_u^{o,\mathcal{U}}, RT_u^{\mathcal{U}})$ and $t_u^{\mathcal{U}} = (l_u^{d,\mathcal{U}}, DD_u^{\mathcal{U}})$, mapping the origin and destination of each ULD. We define the two sets $\mathcal{S}^{\mathcal{U}}$ and $\mathcal{T}^{\mathcal{U}}$ containing, respectively, the origin and destination nodes of all ULDs. We connect each $s_u^{\mathcal{U}}$ with the node $n = (l, t) \in \mathcal{N}$ such that $l = l_u^{o,\mathcal{U}}$ and t is the minimum time greater or equal to $RT_u^{\mathcal{U}}$ via arc a_u^+ . The set containing such arcs is $\mathcal{A}^{\mathcal{U}^+}$. We connect each node $n = (l, t) \in \mathcal{N}$ such that $l = l_u^{d,\mathcal{U}}$ and t is the maximum time smaller or equal to $DD_u^{\mathcal{U}}$ with $t_u^{\mathcal{U}}$ via arc a_u^- instead. The set containing such arcs is $\mathcal{A}^{\mathcal{U}^-}$.

Finally, we consider a set of cargo requests \mathcal{R} , indexed by r . Their definition is akin to ULDs, with a source node and a destination node $s_r^{\mathcal{R}} = (l_r^{o,\mathcal{R}}, RT_r^{\mathcal{R}})$ and $t_r^{\mathcal{R}} = (l_r^{d,\mathcal{R}}, DD_r^{\mathcal{R}})$. Differently from ULDs, these nodes are not explicitly part of the network but are used to ensure cargo requests are assigned to ULDs in a feasible fashion. To this avail, we assume that cargo requests are assigned to a ULD in the origin airport and that each ULD travels without being partially unpacked or repacked until its destination airport, where all cargo requests in the ULD are unpacked. To this end, subsets \mathcal{F}_{ru+} and \mathcal{F}_{ru-} are defined. \mathcal{F}_{ru+} contains the flight arcs f such that $l_f^{o,\mathcal{F}} = l_r^{o,\mathcal{R}} = l_r^{o,\mathcal{U}}$, $DT_f \geq RT_r^{\mathcal{R}}$, and $DT_f \geq RT_u^{\mathcal{U}}$, i.e., the flights where request r and ULD u can feasibly leave the origin airport (both in terms of airport compatibility and time restrictions). \mathcal{F}_{ru-} contains the flight arcs f such that $l_f^{d,\mathcal{F}} = l_r^{d,\mathcal{R}} = l_r^{d,\mathcal{U}}$, $AT_f \leq DD_r^{\mathcal{R}}$, and $AT_f \leq DD_u^{\mathcal{U}}$, i.e., the flights where request r and ULD u can feasibly reach the destination airport (both in terms of airport compatibility and time restrictions).

The set $\mathcal{N}^{\mathcal{K}}$ of nodes where aircraft can be routed is $\mathcal{N}^{\mathcal{K}} = \mathcal{N} \cup \{s\} \cup \{t\}$, while the set $\mathcal{N}^{\mathcal{U}}$ of nodes where ULDs can be routed is $\mathcal{N}^{\mathcal{U}} = \mathcal{N} \cup \mathcal{S}^{\mathcal{U}} \cup \mathcal{T}^{\mathcal{U}}$.

The full set of arcs characterizing the problem is $\mathcal{A} = \mathcal{G} \cup \mathcal{F} \cup \mathcal{E} \cup \mathcal{O} \cup \{a_{bp}\} \cup \mathcal{A}^{\mathcal{U}^+} \cup \mathcal{A}^{\mathcal{U}^-}$. The subset $\mathcal{A}^{\mathcal{K}}$ of arcs where aircraft can be routed is $\mathcal{A}^{\mathcal{K}} \subset \mathcal{A} = \mathcal{G} \cup \mathcal{F} \cup \mathcal{E} \cup \mathcal{O} \cup \{a_{bp}\}$. The subset $\mathcal{A}^{\mathcal{U}}$

of arcs where ULDs can be routed is $\mathcal{A}^{\mathcal{U}} \subset \mathcal{A} = \mathcal{G} \cup \mathcal{F} \cup \mathcal{A}^{\mathcal{U}^+} \cup \mathcal{A}^{\mathcal{U}^-}$.

Both node and arc sets can be tailored to be made aircraft type- and ULD-specific via, respectively, subsets $\mathcal{N}_k^{\mathcal{K}}$, $\mathcal{N}_u^{\mathcal{U}}$, $\mathcal{A}_k^{\mathcal{K}}$, and $\mathcal{A}_u^{\mathcal{U}}$. For example, an ULD $u \in \mathcal{U}$ can only use its specific entry arc a_u^+ to access the TSN and exit arc a_u^- to egress the TSN. Additionally, a ULD cannot use ground or flight arcs defined before its release time $RT_u^{\mathcal{U}}$ or after its delivery deadline $DD_u^{\mathcal{U}}$.

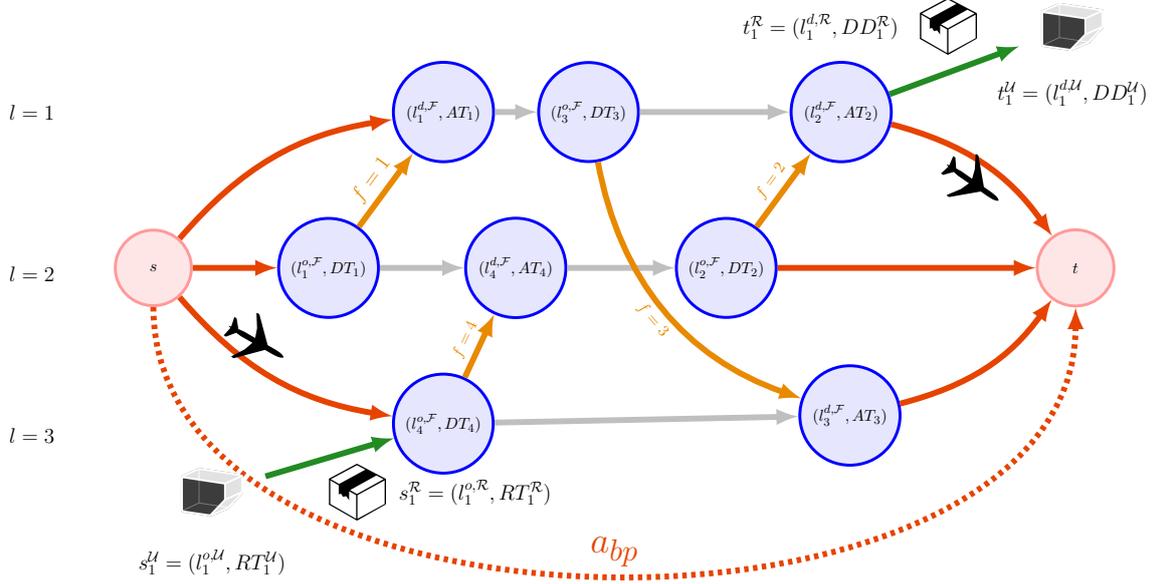


Figure 1: Representation of the network where the routing of aircraft, ULDs, and cargo requests takes place.

The goal of the problem is to simultaneously route in the network the aircraft fleet, the available ULDs, and the cargo requests in a profit-maximizing way. In Figure 1 we provide a visual representation of the network, comprising the original TSN and the auxiliary nodes and arcs. For arcs, we use the following colors: $\mathcal{G} \rightarrow$ gray, $\mathcal{F} \rightarrow$ orange, \mathcal{E} and $\mathcal{O} \rightarrow$ dark orange, $a_{bp} \rightarrow$ dashed dark orange, $\mathcal{A}^{\mathcal{U}^+}$ and $\mathcal{A}^{\mathcal{U}^-} \rightarrow$ green. In the example, $|\mathcal{L}| = 3$, $|\mathcal{F}| = 4$, $|\mathcal{K}| = 1$ and $N_1^{\mathcal{K}} = 1$ (hence only one aircraft is available), $|\mathcal{V}| = 1$ and $N_1^{\mathcal{V}} = 1$ (hence only one ULD is available $\rightarrow |\mathcal{U}| = 1$), and $|\mathcal{R}| = 1$. Note that ULD $u = 1$ and cargo request $r = 1$ are compatible both location-wise ($l_1^{o,\mathcal{U}} = l_1^{o,\mathcal{R}} = 3$ and $l_1^{d,\mathcal{U}} = l_1^{d,\mathcal{R}} = 1$) and time-wise. Time compatibility arises because the flight arc departing the origin ($f = 4$) is in subset $\mathcal{F}_{1,1+}$, and $\mathcal{F}_{1,1-}$ includes flight arc $f = 2$.

Hence, pending weight and volume limitations, cargo request $r = 1$ can be packed in ULD $u = 1$. To route ULD $u = 1$ with cargo request $r = 1$ packed inside, the model can inject one aircraft into the system from node s to the first available node $n \in \mathcal{N}$ of airport $l = 3$ (shown in Figure 1 with the contour of an aircraft to the side of the used dark orange arc $a \in \mathcal{E}$). Packed ULD $u = 1$ accesses the TSN as well via the same node $n \in \mathcal{N}$ by using its specific green arc $a \in \mathcal{A}^{\mathcal{U}^+}$. The aircraft then performs flights $f = 4$ and $f = 2 \in \mathcal{F}$ (connected via a ground arc $g \in \mathcal{G}$) and ends its “physical” journey in airport $l = 1$ and its “mathematical” journey in node t via the dark orange arc $a \in \mathcal{O}$ highlighted with the same aircraft contour. The ULD $u = 1$ uses its specific green arc $a \in \mathcal{A}^{\mathcal{U}^-}$ to reach its sink node $t_1^{\mathcal{U}}$ instead.

4 Mathematical formulations

In this section, we present two Mixed Integer Linear Programming (MILP) models that we employ to solve the integrated routing decisions (i.e., aircraft and cargo routing) with air cargo loading decisions (i.e., assignment of shipments to ULDs and ULDs to aircraft).

4.1 Arc-based formulation

The first mathematical model presented is an arc-based formulation where routing of aircraft and ULDs is modeled as a connected sequence of arcs. Cargo requests, when transported, are assigned to a specific ULD $u \in \mathcal{U}$. We build on the notation introduced in Section 3 to define the parameters and the decision variables characterizing the model.

We assume that each aircraft type k is characterized by a fixed aircraft configuration and can allocate at most N_{kv} ULDs of type v . Because of payload-range diagram considerations, we define W_{fk} the maximum cargo payload that aircraft of type k can carry if using flight arc $f \in \mathcal{F}$. Concerning ULD types, we define respectively $W_v^{\mathcal{V}}$ and $V_v^{\mathcal{V}}$ the maximum weight and volume that a ULD of type v can accommodate. Hence, every ULD $u \in \mathcal{U}_v$ inherits such weight and volume capacity ($W_u^{\mathcal{U}}$ and $V_u^{\mathcal{U}}$). Each cargo request is characterized by a weight $W_r^{\mathcal{R}}$ and a volume $V_r^{\mathcal{R}}$. We define \mathcal{U}_r as the subset of ULDs that can accommodate cargo request r . Conversely, \mathcal{R}_u is the subset of requests that can be transported by ULD u , thus representing the reverse mapping of \mathcal{U}_r . Because of cargo incompatibility restrictions, we define a set of incompatible item pairs $\mathcal{R}_{inc} \subset \mathcal{R} \times \mathcal{R}$, where each $(r_1, r_2) \in \mathcal{R}_{inc}$ defines two cargo requests that cannot be placed in the same ULD u . Regarding revenue and costs, we define R_r the revenue obtained if the delivery of cargo request r is performed, C_{fk} the operational fixed cost associated with operating aircraft type k for flight f , and C_{fu} the operational fixed cost of transporting ULD u on flight f . The decision variables are as follows: $x_{ak} \in \mathbb{N}_0$ defines the number of aircraft of type k using arc a , $y_{au} \in \{0, 1\}$ is unitary if ULD u uses arc a , $z_{ru} \in \{0, 1\}$ is unitary if cargo request r is assigned to ULD u , and $w_{ruf} \in \{0, 1\}$ is unitary if cargo request r is packed in ULD u and travels along flight arc f . A full summary of the notation used in the arc-based formulation is provided in Table A.13.

The resulting mathematical formulation is as follows:

$$\max \sum_{r \in \mathcal{R}} \sum_{u \in \mathcal{U}_r} R_r z_{ru} - \left(\sum_{f \in \mathcal{F}} \sum_{k \in \mathcal{K}_f} C_{fk} x_{fk} + \sum_{f \in \mathcal{F}} \sum_{u \in \mathcal{U}_f} C_{fu} y_{fu} + \sum_{e \in \mathcal{E}} C_e x_{ek} \right) \quad (1)$$

s.t.:

$$\sum_{k \in \mathcal{K}_f} x_{fk} = 1 \quad \forall f \in \mathcal{F} \quad (2)$$

$$\sum_{u \in \mathcal{U}_v} y_{fu} \leq \sum_{k \in \mathcal{K}_f} N_{kv} x_{fk} \quad \forall f \in \mathcal{F}, v \in \mathcal{V} \quad (3)$$

$$\sum_{r \in \mathcal{R}_u} W_r^{\mathcal{R}} z_{ru} \leq W_u^{\mathcal{U}} \quad \forall u \in \mathcal{U} \quad (4)$$

$$\sum_{r \in \mathcal{R}_u} V_r^{\mathcal{R}} z_{ru} \leq V_u^{\mathcal{U}} \quad \forall u \in \mathcal{U} \quad (5)$$

$$\sum_{u \in \mathcal{U}_r} z_{ru} \leq 1 \quad \forall r \in \mathcal{R} \quad (6)$$

$$\sum_{a \in \mathcal{E}} x_{ak} + x_{a_{bp}k} = N_k \quad \forall k \in \mathcal{K} \quad (7)$$

$$\sum_{a \in \mathcal{O}} x_{ak} + x_{a_{bp}k} = N_k \quad \forall k \in \mathcal{K} \quad (8)$$

$$\sum_{a \in \mathcal{A}_{k,n}^{\mathcal{K}+}} x_{ak} - \sum_{a \in \mathcal{A}_{k,n}^{\mathcal{K}-}} x_{ak} = 0 \quad \forall k \in \mathcal{K}, n \in \mathcal{N} \quad (9)$$

$$\sum_{a \in \mathcal{A}_{u,n}^{\mathcal{U}+}} y_{au} - \sum_{a \in \mathcal{A}_{u,n}^{\mathcal{U}-}} y_{au} = 0 \quad \forall u \in \mathcal{U}, n \in \mathcal{N} \quad (10)$$

$$z_{ru} \leq \sum_{f \in \mathcal{F}_{ru}^+} y_{fu} \quad \forall r \in \mathcal{R}, u \in \mathcal{U}_r \quad (11)$$

$$z_{ru} \leq \sum_{f \in \mathcal{F}_{ru}^-} y_{fu} \quad \forall r \in \mathcal{R}, u \in \mathcal{U}_r \quad (12)$$

$$z_{r_1, u} + z_{r_2, u} \leq 1 \quad \forall (r_1, r_2) \in \mathcal{R}_{inc}, u \in \mathcal{U}_{r_1} \cap \mathcal{U}_{r_2} \quad (13)$$

$$x_{ak} \in \mathbb{N}_0 \quad \forall k \in \mathcal{K}, a \in \mathcal{A}_k^{\mathcal{K}} \quad (14)$$

$$y_{au} \in \{0, 1\} \quad \forall u \in \mathcal{U}, a \in \mathcal{A}_u^{\mathcal{U}} \quad (15)$$

$$z_{ru} \in \{0, 1\} \quad \forall r \in \mathcal{R}, u \in \mathcal{U}_r \quad (16)$$

$$w_{ruf} \in \{0, 1\} \quad \forall r \in \mathcal{R}, u \in \mathcal{U}_r, f \in \mathcal{F}_u \quad (17)$$

The objective function (1) aims at maximizing the profit of the solution by accounting for the revenue generated from the transported cargo requests and the costs incurred by the dispatching of the aircraft fleet and ULDs. C_e is a fictitious small cost ($C_e \simeq 0$) used to encourage the model to use the by-pass arc a_{bp} for unused aircraft rather than randomly assigning them to an airport and ground them. While $C_e \simeq 0$ ensures already that such a term has a negligible effect on the final profit, the term $\sum_{e \in \mathcal{E}} C_e x_{ek}$ should be subtracted from the final objective to ensure a faithful representation of the profit. Constraint set (2) ensures that each flight leg is flown by exactly one aircraft. Constraint set (3) ensures that if aircraft type k flies flight leg f , no more ULDs per type v are carried on-board than available by the configuration. Constraints (4)-(5) ensure that each ULD is packed within its weight and volume limits respectively. Constraint set (6) ensures that each cargo request is assigned to one ULD at most. Constraint set (7) ensures that every aircraft of type k leaves the source node s either via the TSN (hence it is being used) or via the by-pass arc a_{bp} . Constraint set (8) is the counterpart for the sink node t . Constraint set (9) imposes conservation of flow per aircraft type k in every node of the TSN. Constraint set (10) imposes conservation of flow for each ULD u in every node of the TSN. Note that if a ULD $u \in \mathcal{U}$ is used, then constraint (10) will ensure such a ULD enters the TSN in node $s_u^{\mathcal{U}}$ and exits the TSN in node $t_u^{\mathcal{U}}$. Constraint set (11) ensures that a cargo request r can leave the origin airport only if a ULD that can transport such a request leaves the origin airport later than $RT_r^{\mathcal{R}}$. Constraint set (12) ensures that a cargo request r can reach the destination airport only if a ULD that can transport such a request arrives there sooner than $DD_r^{\mathcal{R}}$. Constraint set (13) ensures that no incompatible cargo requests are assigned to the same ULD. Finally, (14)-(17) define the nature of the decision variables.

It should be noted that decision variables w_{ruf} have been defined but never appear in constraints (2)-(12). This is not an oversight, because they are needed only if an integer solution is obtained where a flight leg f is flown by aircraft type k with an overall payload that is greater than W_{fk} . In fact, the formulation above can ensure that no more ULDs than allowed by N_{kv} restrictions are loaded on-board and that each ULD is loaded within weight and volume restrictions, but it cannot ensure that no more payload than what is allowed by payload-range diagram considerations is carried. To ensure this, a constraint in the form

$$\sum_{u \in \mathcal{U}_f} \sum_{r \in \mathcal{R}_u} W_r^{\mathcal{R}} w_{ruf} \leq \sum_{k \in \mathcal{K}_f} W_{fk} x_{fk} \quad \forall f \in \mathcal{F} \quad (18)$$

and the linearization constraints

$$w_{ruf} \geq y_{fu} + z_{ru} - 1 \quad \forall f \in \mathcal{F}, u \in \mathcal{U}_f, r \in \mathcal{R}_u \quad (19)$$

$$w_{ruf} \leq y_{fu} \quad \forall f \in \mathcal{F}, u \in \mathcal{U}_f, r \in \mathcal{R}_u \quad (20)$$

$$w_{ruf} \leq z_{ru} \quad \forall f \in \mathcal{F}, u \in \mathcal{U}_f, r \in \mathcal{R}_u \quad (21)$$

which are needed to enforce $w_{ruf} = 1$ if and only if both $y_{fu} = 1$ and $z_{ru} = 1$ hold. Because of the gargantuan size of such constraint sets for large problems, the approach followed here is not to

define any constraint (18)-(21) upfront. For any integer solution that is found during the branch-and-bound process, the payload transported along each flight arc is checked with Equation (18). If no infeasibility is found, no action is taken. Conversely, for every flight leg where the payload is exceeded, the associated constraints (18)-(21) are added ad-hoc so that feasibility is restored. This *lazy constraint* approach ensures that constraints are only introduced when necessary, avoiding the computational burden of modeling them upfront.

4.2 Hybrid arc- and path-based formulation

In the problem at hand, there is a large disparity cardinality-wise between the commodities involved. In fact, the number of aircraft types involved is much smaller than the number of ULDs and cargo requests: $|\mathcal{K}| \ll |\mathcal{U}| \simeq |\mathcal{R}|$. While only a fraction of y_{au} decision variables will be non-zero in the optimal solution, a fully arc-based formulation requires the definition of all such routing decision variables beforehand. To alleviate this algorithmic shortcoming, a common workaround is to employ a Column Generation (CG) algorithm. In CG, the routing of a commodity is represented as a path (a sequence of connected arcs) from its origin to its destination, rather than as a sequence of individual arcs. Hence, to implement a CG algorithm for this problem, we must revise our mathematical formulation. In this hybrid arc- and path-based formulation the aircraft types are still routed in an arc-based fashion. However, routing of ULDs and cargo requests is carried out using a path-based perspective.

The first step to fully characterize the hybrid formulation is to define ULD paths. Consider a set of ULD paths \mathcal{P} where each path $p \in \mathcal{P}$ is characterized by a routing component and a packing component. The routing component is a sequence of connected arcs starting from $s_u^{\mathcal{U}}$ and ending in $t_u^{\mathcal{U}}$ for any $u \in \mathcal{U}$. This means that a path p is specific to a certain ULD and can only be used by that ULD. Furthermore, the packing component defines the cargo requests $r \in \mathcal{R}$ packed in the ULD. We then define \mathcal{P}_u as the subset of paths of ULD u such that the full set of ULD paths is given by: $\mathcal{P} = \mathcal{P}_1 \cup \dots \cup \mathcal{P}_{|\mathcal{U}|}$. Additionally, with \mathcal{P}_u^f and \mathcal{P}_u^r we indicate, respectively, the subset of paths of ULD u containing flight arc $f \in \mathcal{F}$ and cargo request $r \in \mathcal{R}$. In this formulation, a path already carries information regarding ULD routing from its origin to its destination and cargo request packing inside it, making decision variables y_{au} and z_{ru} redundant. They are replaced by the set of binary decision variables z_{up} , where each z_{up} is unitary if ULD $u \in \mathcal{U}$ is routed in the network via path $p \in \mathcal{P}_u$ (and carrying the cargo requests associated with path p). Referring back to Figure 1, ULD $u = 1$ was characterized by a routing of 5 connected arcs and was transporting cargo request $r = 1$. This full set of information defines a potential path p for ULD $u = 1$ in this formulation. For every path, the following parameters are known. W_{up} and R_{up} are, respectively, the cumulative weight and revenue of the cargo requests carried by the considered ULD along path p , while C_{up} is the overall operational cost due to the transportation of ULD u along path p . A full summary of the additional notation used for the hybrid formulation is provided in Table B.14.

The mathematical hybrid arc- and path-based formulation is as follows:

$$\max \sum_{u \in \mathcal{U}} \sum_{p \in \mathcal{P}_u} R_{up} z_{up} - \left(\sum_{f \in \mathcal{F}} \sum_{k \in \mathcal{K}_f} C_{fk} x_{fk} + \sum_{u \in \mathcal{U}} \sum_{p \in \mathcal{P}_u} C_{up} z_{up} + \sum_{e \in \mathcal{E}} C_e x_{ek} \right) \quad (22)$$

s.t.:

$$\sum_{k \in \mathcal{K}_f} x_{fk} = 1 \quad \forall f \in \mathcal{F} \quad (23)$$

$$\sum_{a \in \mathcal{E}} x_{ak} + x_{a_{bp}k} = N_k \quad \forall k \in \mathcal{K} \quad (24)$$

$$\sum_{a \in \mathcal{O}} x_{ak} + x_{a_{bp}k} = N_k \quad \forall k \in \mathcal{K} \quad (25)$$

$$\sum_{a \in \mathcal{A}_{k,n+}^{\mathcal{K}}} x_{ak} - \sum_{a \in \mathcal{A}_{k,n-}^{\mathcal{K}}} x_{ak} = 0 \quad \forall k \in \mathcal{K}, n \in \mathcal{N} \quad (26)$$

$$\sum_{u \in \mathcal{U}_v^f} \sum_{p \in \mathcal{P}_u^f} z_{up} \leq \sum_{k \in \mathcal{K}_f} N_{kv} x_{fk} \quad \forall f \in \mathcal{F}, v \in \mathcal{V} \quad (27)$$

$$\sum_{u \in \mathcal{U}_f} \sum_{p \in \mathcal{P}_u^f} W_{up} z_{up} \leq \sum_{k \in \mathcal{K}_f} W_{fk} x_{fk} \quad \forall f \in \mathcal{F} \quad (28)$$

$$\sum_{p \in \mathcal{P}_u} z_{up} \leq 1 \quad \forall u \in \mathcal{U} \quad (29)$$

$$\sum_{u \in \mathcal{U}_r} \sum_{p \in \mathcal{P}_u^r} z_{up} \leq 1 \quad \forall r \in \mathcal{R} \quad (30)$$

$$x_{ak} \in \mathbb{N}_0 \quad \forall k \in \mathcal{K}, a \in \mathcal{A}_k^{\mathcal{K}} \quad (31)$$

$$z_{up} \in \{0, 1\} \quad \forall u \in \mathcal{U}, p \in \mathcal{P}_u \quad (32)$$

The objective function (22) is the counterpart of objective function (1), where the revenue and ULD-specific cost terms are modified to account for the path-based decision variables. Constraints (23)-(26) are inherited directly from the arc-based formulation, as the aircraft routing follows the same principles. Constraint set (27) is a modified version of (3) that ensures no more ULDs of type v than allowed are transported per flight leg. Constraint set (28) ensures that the cargo payload for each flight leg satisfies payload-range limitations. Since in this formulation a path already integrates information on both ULD routing and cargo requests contents, this constraint can be directly imposed with the available decision variables without resorting to additional ones as seen in Section 4.1. Constraint set (29) enforces that, for each ULD $u \in \mathcal{U}$, at most one available path $p \in \mathcal{P}_u$ can be selected. In a similar fashion, constraint set (30) enforces that, for each request $r \in \mathcal{R}$, at most one path containing it can be selected. Finally, (31)-(32) define the nature of the decision variables.

5 Solution approaches

This section describes how the two MILPs presented in Section 4 are used within three distinct solution approaches designed to solve the problem defined in Section 3.

5.1 Sequential approach

Solving the fully arc-based model is computationally challenging given the large number of variables associated with ULD routing and assignment of requests to ULD. Therefore, we employ a sequential approach where we simplify packing constraints in the first stage and then reinstate them in the second stage. We consider this approach a myopic bin-packing model, as simplifying packing decisions in the initial phase might lead to a poor assignment of requests to ULDs in the second step, especially when many incompatibility constraints between requests occur. In the remainder of the paper, we will refer to this first solution approach as sequential approach.

5.1.1 Sequential approach: First stage

To simplify the assignment of cargo requests to ULDs in the first stage, we introduce a set of dummy ULDs, denoted by $\bar{\mathcal{U}}$. This set has the same cardinality as the set of cargo requests, i.e., $|\bar{\mathcal{U}}| = |\mathcal{R}|$. Each dummy ULD $\bar{u} \in \bar{\mathcal{U}}$ is uniquely associated with a specific request $r \in \mathcal{R}$, inheriting its volume, weight, and required features. To ensure a one-to-one correspondence, we define $\bar{\mathcal{U}}_r = \{\bar{u}_r\}$, a singleton set containing only the dummy ULD associated with request r . Consequently, each request r is constrained to be assigned exclusively to its associated dummy ULD, \bar{u}_r . With this setup, we effectively bypass the intermediate step of assigning requests to physical ULDs. Instead, the problem reduces to directly assigning cargo requests to aircraft types. The dummy ULDs serve as proxies for requests, allowing the fully arc-based formulation to directly assign requests to flight legs while ensuring compliance with operational constraints.

The mathematical formulation of the first stage of the sequential approach is as follows:

$$\max \sum_{r \in \mathcal{R}} R_r z_{r\bar{u}_r} - \left(\sum_{f \in \mathcal{F}} \sum_{k \in \mathcal{K}_f} C_{fk} x_{fk} + \sum_{e \in \mathcal{E}} C_e x_{ek} \right) \quad (33)$$

s.t.: (2), (7)-(9)

$$\sum_{a \in \mathcal{A}_{\bar{u},n}^{\bar{U}+}} y_{a\bar{u}} - \sum_{a \in \mathcal{A}_{\bar{u},n}^{\bar{U}-}} y_{a\bar{u}} = 0 \quad \forall \bar{u} \in \bar{U}, n \in \mathcal{N} \quad (34)$$

$$z_{r\bar{u}_r} \leq \sum_{f \in \mathcal{F}_{r\bar{u}_r+}} y_{f\bar{u}} \quad \forall r \in \mathcal{R} \quad (35)$$

$$z_{r\bar{u}_r} \leq \sum_{f \in \mathcal{F}_{r\bar{u}_r-}} y_{f\bar{u}} \quad \forall r \in \mathcal{R} \quad (36)$$

$$\sum_{\bar{u} \in \bar{U}_f} W_{\bar{u}}^{\bar{U}} y_{f\bar{u}} \leq \sum_{v \in \mathcal{V}_f} \sum_{k \in \mathcal{K}_f} N_{kv} W_v^{\mathcal{V}} x_{fk} \quad \forall f \in \mathcal{F} \quad (37)$$

$$\sum_{\bar{u} \in \bar{U}_f} V_{\bar{u}}^{\bar{U}} y_{f\bar{u}} \leq \sum_{v \in \mathcal{V}_f} \sum_{k \in \mathcal{K}_f} N_{kv} V_v^{\mathcal{V}} x_{fk} \quad \forall f \in \mathcal{F} \quad (38)$$

$$\sum_{\bar{u} \in \bar{U}_f} W_{\bar{u}}^{\bar{U}} y_{f\bar{u}} \leq \sum_{k \in \mathcal{K}_f} W_{fk} x_{fk} \quad \forall f \in \mathcal{F} \quad (39)$$

$$x_{ak} \in \mathbb{N}_0 \quad \forall k \in \mathcal{K}, a \in \mathcal{A}_k^{\mathcal{K}} \quad (40)$$

$$y_{a\bar{u}} \in \{0, 1\} \quad \forall \bar{u} \in \bar{U}, a \in \mathcal{A}_{\bar{u}}^{\bar{U}} \quad (41)$$

$$z_{r\bar{u}_r} \in \{0, 1\} \quad \forall r \in \mathcal{R} \quad (42)$$

The objective function (33) is nearly identical to objective function (1), with the only difference being that we omit operational costs due to routing ULDs as we are only dealing with dummy ULDs. The decision variables are now applied to the dummy ULD set \bar{U} . Constraints (34), (35) and (36) are also very similar to their arc-based counterparts (10), (11), and (12). They ensure conservation of flow in every node for dummy ULDs, feasible request departure from the origin airport, and feasible request arrival at the destination airport respectively. Constraints (37) and (38) enforce cumulative weight and volume limitations on the number of dummy ULDs that can be transported along each flight leg. The cumulative weight and volume transported by the dummy ULDs must not exceed the available cumulative weight and volume capacity of the original ULD types for the specific aircraft configuration. Furthermore constraint set (39) follows from the *lazy constraint* set (18), ensuring that the cumulative weight of each flight leg does not exceed the maximum payload restrictions according to the payload-range diagram. Because now we are directly assigning requests (via dummy ULDs) to flights, such a constraint can be directly implemented without auxiliary decision variables as in the fully arc-based approach. Finally, (40)-(42) define the nature of the decision variables.

5.1.2 Sequential approach: Second stage

The myopic solution obtained in the first stage provides an upper bound on the actual solution. Since we neglected bin-packing and incompatibility constraints between requests, it should result in a deterioration (decrease) in the objective value once those constraints are reinstated. Additionally, reintroducing the ULD operational costs might also potentially reduce the objective value. The goal of the second stage in our sequential approach is to reinstate the bin-packing constraints while fixing the aircraft type and request routing based on the first-stage solution.

Let us define \mathcal{X} , the set of tuples (a, k) with $a \in \mathcal{A}_k^{\mathcal{K}}$ and $k \in \mathcal{K}$. A tuple (a, k) belongs to \mathcal{X} if in the first stage solution $x_{ak} = 1$, meaning \mathcal{X} defines the set of active routing decision variables for the aircraft fleet. By defining \mathcal{X} in this way, the second stage no longer requires the fleet routing decision variables to be explicitly considered. Additionally, we define \mathcal{F}_r as the subset of flight arcs

that are active for a given request. These are directly derived from the first stage solution where $y_{f\bar{u}} = 1$. Each request will have to follow its fixed path.

The second stage of the sequential approach is mathematically formulated as follows:

$$\max \sum_{r \in \mathcal{R}} \sum_{u \in \mathcal{U}_r} R_r z_{ru} - \left(C^{\mathcal{FK}} + \sum_{f \in \mathcal{F}} \sum_{u \in \mathcal{U}_f} C_{fu} y_{fu} \right) \quad (43)$$

s.t.: (3)-(5), (10)-(13),

$$\sum_{u \in \mathcal{U}_v} y_{fu} \leq N_{fv} \quad \forall f \in \mathcal{F}, v \in \mathcal{V} \quad (44)$$

$$z_{ru} \leq y_{fu} \quad \forall r \in \mathcal{R}, u \in \mathcal{U}_r, f \in \mathcal{F}_r \cap \mathcal{F}_u \quad (45)$$

$$z_{ru} + y_{fu} \leq 1 \quad \forall r \in \mathcal{R}, u \in \mathcal{U}_r, f \in \mathcal{F}_u \setminus \mathcal{F}_r \quad (46)$$

$$y_{au} \in \{0, 1\} \quad \forall u \in \mathcal{U}, a \in \mathcal{A}_u^{\mathcal{U}} \quad (47)$$

$$z_{ru} \in \{0, 1\} \quad \forall r \in \mathcal{R}, u \in \mathcal{U}_r \quad (48)$$

The objective function stems from objective function (1), with the notable change that the aircraft fleet operating cost $C^{\mathcal{FK}}$ is now treated as a constant (and is computed using set \mathcal{X}), and the fictitious term involving C_e is omitted. This is because the aircraft routing has already been determined in the first stage. Constraints (44) are a modification of (3) since the aircraft type, and therefore the maximum number of ULD slots available per flight arc, is known from the first stage. Constraint set (45) enforces that if a request r is assigned to a ULD u , the ULD must traverse all flight arcs in the union of \mathcal{F}_r and \mathcal{F}_u . Similarly, due to constraints (46), a request r cannot be assigned to a ULD u if that ULD utilizes flight arcs that are in \mathcal{F}_u but not in \mathcal{F}_r . These constraints maintain the request routing decisions made in the first stage while allowing flexibility in the assignment of requests to ULDs. Finally in (47) and (48) the decision variables are defined.

5.2 Integrated approach

Our second solution approach, referred to as the integrated approach, is based on the hybrid formulation and performs both the request-to-ULD and ULD-to-aircraft assignments simultaneously. Section 5.2.1 describes how an iterative Column Generation (CG) procedure is carried out that adds promising paths as needed.

5.2.1 Integrated approach: Column generation

In classic CG fashion, we first solve the linear relaxation of the hybrid arc- and path-based formulation, i.e., our Restricted Master Problem (RMP). To do so, we need an initial set of paths. Using a simple heuristic, a single path is created for each ULD $u \in \mathcal{U}$. Each path is assigned the shortest arc path between the origin and destination nodes as its routing component. Subsequently, for the packing component, the ULD is filled with appropriate requests from \mathcal{R}_u in random order until the ULD is full.

After solving the linear relaxation, we use the dual values associated with the optimal solution to solve a pricing problem for each ULD $u \in \mathcal{U}$, which is described in Section 5.2.2. The goal of this pricing problem is to find additional paths, also named columns, that could be beneficial for the solution of the RMP, i.e., columns with a positive reduced cost since we are dealing with a maximization problem. For each instance of the pricing problem, we save the 20 best solutions. If the objective value of a solution is strictly positive and the path is not a duplicate of one already considered, we add it to the RMP as a new potential path. Once the pricing problem is solved for all $u \in \mathcal{U}$, a large number of new paths (columns) are added to the RMP, which is then solved again. This yields new dual values, altering the pricing problem and allowing the entire process to repeat. This iterative procedure continues until no new columns can be added or until a predetermined number of full iterations is reached. Unless mentioned otherwise, we set the limit to 5 iterations.

Once the CG process is complete, the integrality constraints of the decision variables in the RMP are reinstated, and it is solved again. This will result in a drop in the objective value, but the solution will be feasible.

5.2.2 Integrated approach: Pricing problem

Considering the formulation (22)-(32), let us define the dual values γ_{fv} , δ_f , η_u , and λ_r for constraints (27), (28), (29), and (30), respectively. Because dual values can only be computed for formulations where decision variables are continuous, we must consider the relaxed version of (22)-(32), where $x_{ak} \in \mathbb{R}_0$ and $z_{up} \in [0, 1]$.

The pricing problem is set up as a maximization problem of the reduced cost vector. This problem is solved for every ULD $u \in \mathcal{U}$ to find promising paths that should be added to the RMP. Since we are dealing with a maximization problem, a path should be added if the reduced costs are positive, i.e. $c - z > 0$. Let us define the two decision variables: $y_a \in \{0, 1\}$ is unitary if ULD u uses arc a , and $z_r \in \{0, 1\}$ is unitary if cargo request r is assigned to ULD u .

The mathematical formulation of the pricing problem for a specific ULD $u \in \mathcal{U}$ is as follows:

$$\max \sum_{r \in \mathcal{R}_u} (R_r - \lambda_r) z_r - \sum_{f \in \mathcal{F}_u} (C_{fu} + \gamma_{fv} + W_{up} \delta_f) y_f - \eta_u \quad (49)$$

s.t.:

$$\sum_{r \in \mathcal{R}_u} W_r^{\mathcal{R}} z_r \leq W_u^{\mathcal{U}} \quad (50)$$

$$\sum_{r \in \mathcal{R}_u} V_r^{\mathcal{R}} z_r \leq V_u^{\mathcal{U}} \quad (51)$$

$$\sum_{a \in \mathcal{A}_{u,n}^{\mathcal{U}+}} y_a - \sum_{a \in \mathcal{A}_{u,n}^{\mathcal{U}-}} y_a = 0 \quad \forall n \in \mathcal{N} \quad (52)$$

$$z_r \leq \sum_{f \in \mathcal{F}_{ru}^+} y_f \quad \forall r \in \mathcal{R}_u \quad (53)$$

$$z_r \leq \sum_{f \in \mathcal{F}_{ru}^-} y_f \quad \forall r \in \mathcal{R}_u \quad (54)$$

$$z_{r_1} + z_{r_2} \leq 1 \quad \forall (r_1, r_2) \in \mathcal{R}_{inc} \quad (55)$$

$$z_r \in \{0, 1\} \quad \forall r \in \mathcal{R}_u \quad (56)$$

$$y_a \in \{0, 1\} \quad \forall a \in \mathcal{A}_u^{\mathcal{U}} \quad (57)$$

The objective function (49) represents the reduced costs of the possible path. The first component reflects the revenue gained by transporting a request through the network, while the second component maps the costs associated with transporting the ULD across the active flight arcs. The last components is a constant in the objective function.

Constraints (50) and (51) follow from (4) and (5) and ensure the ULD considered adheres to its maximum weight and volume capacity $W_u^{\mathcal{U}}$, and $V_u^{\mathcal{U}}$ respectively. Additionally, constraint set (53) mandates that requests can only be assigned to a ULD if the ULD departs from the origin airport after the request's release time. Similarly, constraint set (54) dictates that requests can only be assigned to a ULD if the ULD arrives at the destination airport before the request's delivery deadline. These are modifications of (11) and (12) respectively. Following from (13), constraint set (55) ensures that incompatible requests cannot be assigned to the same ULD. Finally, (56)-(57) define the nature of the decision variables.

The formulation of objective function (49) is nonlinear as decision variable y_f is multiplied with W_{up} , i.e., the cumulative weight of the cargo requests carried along path p . W_{up} depends on the request assigned to the ULD which we do not know because they are part of the decision-making

process. We therefore need a linearization decision variable in the form of w_{uf} which modifies the mathematical formulation of the pricing problem to:

$$\max \sum_{r \in \mathcal{R}_u} (R_r - \lambda_r) z_r - \sum_{f \in \mathcal{F}_u} (C_{fu} + \gamma_{fv}) y_f - \sum_{f \in \mathcal{F}_u} \delta_f w_{uf} - \eta_u \quad (58)$$

s.t.: (50)-(57),

$$w_{uf} \in \mathbb{N}_0 \quad \forall f \in \mathcal{F}_u \quad (59)$$

$$w_{uf} \geq \sum_{r \in \mathcal{R}_u} W_r^{\mathcal{R}} z_r - M(1 - y_f) \quad \forall f \in \mathcal{F}_u \quad (60)$$

$$w_{uf} \leq \sum_{r \in \mathcal{R}_u} W_r^{\mathcal{R}} z_r \quad \forall f \in \mathcal{F}_u \quad (61)$$

$$w_{uf} \leq M y_f \quad \forall f \in \mathcal{F}_u \quad (62)$$

The new objective function (58) retains the structure of the previous objective function while incorporating the linearization decision variable. Constraint sets (59)-(62) define w_{uf} such that $w_{uf} = W_{up}$ when $y_f = 1$, and $w_{uf} = 0$ when $y_f = 0$.

Introducing this linearization variable can significantly increase memory usage, especially in larger instances where the pricing problem has to be solved numerous times. To address this issue, for larger instances, we simplify the linearization as follows:

$$\max \sum_{r \in \mathcal{R}_u} (R_r - \lambda_r) z_r - \sum_{f \in \mathcal{F}_u} (C_{fu} + \gamma_{fv} + W_{max} \delta_f) y_f - \eta_u \quad (63)$$

s.t.: (50)-(57).

In this simplified formulation (63), W_{up} always takes the maximum weight that could be placed in the ULD, based on its type. While this assumption is somewhat optimistic, it aligns with the pricing problem's objective of maximizing revenue, which is driven by the total cargo being transported. As a result of this estimation, the second component of objective function (63) takes on a larger negative value than necessary, which might lead to some beneficial columns being excluded from the RMP. However, this is not considered a significant drawback. The model already selects the top 20 solutions from the pricing problem, even though many other solutions may still have strictly positive reduced costs. While some potentially beneficial columns may be excluded due to these simplifications, limiting the number of added columns is a common and practical strategy to manage the size of the RMP and reduce computational time. Moreover, simplifying the linearization process helps optimize memory usage while preserving the core functionality of the model.

5.3 Original fully arc-based formulation

Together with the two solution approaches presented in Section 5.1 and Section 5.2, we will also run every instance using the fully arc-based formulation for benchmarking purposes. In fact, this model is the only one that captures all the original complexity of our problem, while the integrated approach is a heuristic solution to it, and the sequential approach a two-stage decomposition of it. However, the fully arc-based formulation will also struggle the most with computational time and convergence as the problem size grows.

6 Instance description

The instances used for our computational experiments are primarily constructed using synthetic data, as finding good-quality, publicly available, data for cargo operations is extremely challenging, if not impossible (van Bockstaele et al., 2023). This section describes in sequence the input parameters in terms of flight schedules, aircraft types, ULDs, and cargo requests. Then we provide some

insights into the size and characteristics of our instances. Our input flight schedules mirror parts of the summer 2024 freighter timetable of a major European combination airline and is described in Section 6.1. Their network comprising Europe, Africa, and the Americas contains the following twelve airports: Amsterdam (AMS), Bogota (BOG), Cairo (CAI), Buenos Aires (EZE), Guatemala City (GUA), Harare (HRE), Johannesburg (JNB), Miami (MIA), Nairobi (NBO), Santiago (SCL), Quito (UIO), and Campinas (VCP). The location of the airports is shown in Figure 2.

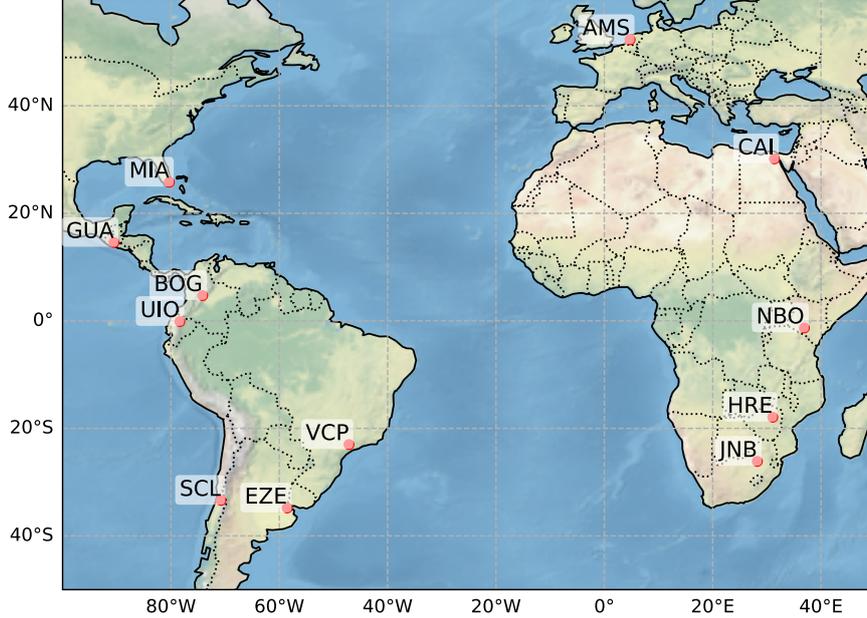


Figure 2: Set of airports included.

6.1 Schedule

In our instances, we consider two schedules: a *full* schedule resembling a week of the 2024 summer timetable of the reference carrier, and a *reduced* schedule resembling only three days of the same week schedule. The full schedule contains 47 flight legs between all 12 airports of the network. The reduced schedule has 16 flight legs but still services 10 airports of the network. A visualization of both schedules in a TSN form can be found in Appendix C. In principle, each flight arc is mandatory to be flown due to constraints (2) and their derivatives. We label such a scenario the *restricted* scenario. We also consider a *free* scenario where all flight arcs become optional. In this scenario the equality constraints (2) and their derivatives are relaxed to inequality constraints, allowing the model to determine, based on projected revenue and operational costs, whether a flight arc is worth operating.

6.2 Aircraft types

The main characteristics of the reference airline fleet are presented in Table 1. This fleet consists of a single B747-400BCF and three B747-400ERF aircraft. There are slight differences in their payload-range characteristics (Boeing, 2008, 2010), and we assume different operating costs for the two aircraft types: 16.03 €/km for the extended-range B747-400ERF (van der Meulen et al., 2023), and a 5% lower cost of 15.23 €/km for the B747-400BCF. We assume an identical ULD configuration for both aircraft types. We acknowledge that using such a homogeneous fleet may not yield very interesting insights into a fleet assignment problem. However, we chose to stick with the reference airline fleet to maintain realism in our test case, as fleet assignment is not the primary focus of this research. Additionally, in the restricted scenario we limit the source arcs $a \in \mathcal{E}$ and sink arcs $a \in \mathcal{O}$ to only be established at the hub airport in Amsterdam, thus $|\mathcal{E}| = |\mathcal{O}| = 1$. In the free scenario, the model has full flexibility in determining where to inject aircraft into the TSN, establishing source and sink connections at every airport $l \in \mathcal{L}$, meaning $|\mathcal{E}| = |\mathcal{O}| = |\mathcal{L}|$.

Table 1: Main characteristics of the two aircraft types.

| Fleet type | # | Max. Payload [t] | Max. Range [km] | Op. Cost [€/km] | LD3 | LD7 |
|-------------|---|------------------|-----------------|-----------------|-----|-----|
| B747-400BCF | 1 | 107 | 7500 | 15.23 | 32 | 30 |
| B747-400ERF | 3 | 112 | 9200 | 16.03 | 32 | 30 |

6.3 ULD types

This work considers four ULD types: LD3, LD3+, LD7, and LD7+. The LD3 and LD7 are standard container and pallet types, respectively, while the "+" versions offer advanced storage capabilities. The LD3 is a container and is hence characterized by a closed and rigid contour, with a cut on the lower part to fit inside the lower part of the fuselage. The LD7 is a pallet with an 88" \times 125" base, with a maximum vertical extension up to 96" (variable due to the curvature of the fuselage). Both the LD3 and LD7 share fixed volume and weight capacities with their respective "+" versions. The ULD types are illustrated in Figure 3.

Certain cargo requests might require specific handling and storing restrictions in order to comply with the nature of the shipment. We introduce two key features: accommodation of *hazardous materials* (Feature I) and a *temperature-controlled environment* (Feature II). For each ULD type, we define a tuple (\cdot, \cdot) where the i -th element is unitary if such ULD type can accommodate cargo requiring feature i and zero otherwise. The containers LD3 and LD3+ can handle hazardous materials, but only the LD3+ offers temperature control. The LD7 is a basic pallet with no special features, whereas the LD7+ is an enclosed pallet that supports both hazardous and temperature-sensitive cargo. Each ULD type has a specific operating cost, reflecting the fact that more advanced ULDs incur higher usage expenses. The characteristics of all ULD types are presented in Table 2.

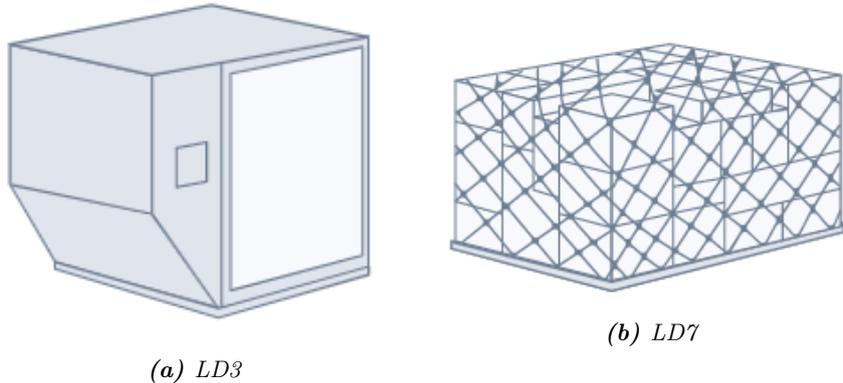


Figure 3: ULD types (from www.searates.com).

Table 2: Main characteristics of the ULD types.

| ULD type | Max. Weight [t] | Max. Vol. [m^3] | Haz. | Temp. | Op. Cost [€/km] |
|----------|-----------------|---------------------|------|-------|-----------------|
| LD3 | 1.5 | 4.5 | 1 | 0 | 0.105 |
| LD3+ | 1.5 | 4.5 | 1 | 1 | 0.115 |
| LD7 | 4.6 | 10.7 | 0 | 0 | 0.1 |
| LD7+ | 4.6 | 10.7 | 1 | 1 | 0.135 |

6.4 Origin-Destination demand matrix

An origin-destination (OD) demand matrix is a critical input for our model, as it represents the flow of cargo requests between airports. Generating such a matrix presents significant challenges due to the highly competitive and confidential nature of relevant data. While some sources, like van Bockstaele et al. (2023) and Boeing (2022), provide aggregated global demand data between regions, this information is often incomplete, unavailable for specific region pairs, or lacks detailed

breakdowns of country-level contributions within regions. Additionally, these sources do not offer insights into how our reference airline compares with competitors in specific regions, making a market share estimation speculative at best. Given these limitations, our demand matrix is primarily derived from flight schedules, with adjustments based on global cargo demand data from Boeing (2022) where possible. The subsequent explanation of OD demand generation considers the full flight schedule; however, a similar methodology can be applied to find an OD demand matrix for a reduced schedule.

The first step is to examine the number of *distinct* flight paths available between each OD pair. By distinct we mean that the paths do not share any common flight arcs. For example, consider the OD pair CAI-JNB. The sequence of airports required to reach this destination would be CAI-NBO-AMS-JNB. Looking at the full schedule in Figure C.5, we find *some* of the possible paths to achieve this are: 60-61-62-63-70-71; 60-61-62-63-78-79; or 60-61-76-77-78-79. However, there is overlap in the flight arcs used by each path: all three paths make use of the leg 60-61, while 62-63 and 78-79 also occur more than once. We only consider one distinct path between CAI-JNB since all paths we have found make use of 60-61. In fact, this is the only flight leaving CAI, so in principle, we can at most have one distinct flight path to every destination, as all possible paths will use this critical flight arc. Consider another example with the OD pair MIA-SCL. To reach this destination, the sequence would be MIA-AMS-VCP-SCL and *some* of the possible paths include: 28-29-10-11-12-13; 56-57-10-11-12-13; 56-57-0-1-2-3; or 48-49-0-1-2-3. Again, there is overlap in the flight arcs being used. The bottleneck here is the number of flights from AMS to SCL (via VCP), which is limited to two. Therefore, the number of distinct paths for this OD pair is also limited to two. For the full distinct path matrix refer to Table D.15.

Generally, more distinct paths should indicate higher demand for an OD pair, however, adjustments are made based on global cargo data. The objective is to incorporate known global imbalances on certain routes while still keeping the total demand matrix relatively balanced. Specifically, the import/export ratio per airport should remain within the 50% to 150% range.

We can assign ratings to each OD pair based on available cargo data or characteristics of their respective paths or the overall flight schedule; per origin, we rate each of the destinations using a relative ranking system A to E informed by cargo data from Boeing (2022). For origins like AMS and MIA, where sufficient data points exist among the possible destinations, a K-means clustering analysis with 5 clusters is performed. Consequently, AMS-CAI and AMS-VCP are both rated A to reflect the high cargo flows from the EU to North Africa and from the EU to Brazil, respectively. Conversely, AMS-GUA is rated E since the Guatemala only accounts for a very small share of the total Latin America-Europe trade. For most origins not enough data is available to objectively assign rankings. For those origins we pick one destination with available cargo data as a baseline with rating C. Data to other destinations that significantly differ from this baseline may be rated higher or lower. For many OD pairs, no data is available at all so there we need to make an educated guess for the rating. This relative approach means that there is no clear cutoff for what constitutes an A, B, or any other rating. Selecting a rating can be an iterative process in order to balance out the final demand matrix. The full rating matrix is available in Table D.16.

As aforementioned, the characteristics of the paths between an OD pair can also have an influence on its rating as some indirect flight paths follow a less than optimal route. For example the SCL-VCP route that goes via AMS, or the HRE-JNB route going via AMS. In these cases a reduced demand can be expected. Therefore we reduce the ranking by one, for example in the demand from SCL to VCP from a C to a D or the demand from HRE to JNB from a B to a C.

Alternatively, in some scenarios, we want to artificially boost demand between a certain OD pair. For example, EZE can only be reached via AMS and VCP early in the schedule, limiting access from other airports. To mimic demand for this destination that stems from last weeks schedule we increase the ranking by two levels. As an example, the demand from AMS to EZE increases from a D to a B rating. To be able to boost an OD pair already rated A or B, we introduce the rankings AAA, and AA. For example, the pair AMS-CAI is increased from an A to an AAA.

Once the ratings are established, they are paired with numerical values: 2.33 for AAA, 2 for AA, 1.67 for A, 1.33 for B, 1 for C, 0.67 for D, and 0.33 for E. These values are then multiplied by the number of distinct paths to calculate the adjusted scores. Based on these adjusted scores,

we determine the percentage of total demand allocated to each destination from a given origin. For instance, originating from AMS, 24.4% of the demand will be allocated to MIA and 7.3% will be directed to BOG. The full OD demand matrix is given in Table D.17.

We acknowledge that this approach to generating the OD demand matrix is somewhat artificial and includes subjective elements. However, given that the primary objective of this research is not the creation of synthetic data, and due to the unavailability of real data, we consider this the most appropriate modeling approach.

6.5 Cargo commodity types

We identified 4 main cargo commodity types: chemicals, perishables, heavy, and others. With chemicals (Lufthansa, 2024), we mean a wide variety of products such as fertilizers, raw materials for the pharmaceutical and cosmetics industries, flavoring agents, adhesives, and sealants, as well as reference materials for laboratory testing. With perishables (IATA, 2021), we mean time and/or temperature-sensitive items that require fast and safe delivery to maintain quality and effectiveness. Some perishable goods require lower temperatures to slow down respiration, which helps delay the ripening process and minimize any negative effects on the perishable cargo. With heavy, we mean commodities not belonging to the previous two categories that are characterized by a higher density, such as metals, machinery, or computers. Finally, with other, we mean generic requests that do not belong to any of the previous three categories.

In Boeing (2022), air cargo commodity flows between various regions are reported as percentages. We used these percentages to determine the probability that a cargo request between those regions belongs to one of the four categories we introduced. When data for a specific OD pair was not available, we replicated the percentages of what we deemed the most similar available OD pair. In Table 3 all commodity type percentages are reported.

Table 3: Commodity flow percentages (weight-wise) in decimal form between different continent OD pairs.

| O-D | Chemical [%] | Perishable [%] | Heavy [%] | Other [%] |
|-------|--------------|----------------|-----------|-----------|
| AF-EU | 1 | 83 | 10 | 6 |
| EU-AF | 18 | 12 | 35 | 35 |
| NA-EU | 20 | 6 | 40 | 34 |
| EU-NA | 13 | 8 | 35 | 44 |
| LA-EU | 2 | 78 | 15 | 5 |
| EU-LA | 15 | 6 | 40 | 39 |
| LA-NA | 2 | 70 | 11 | 17 |
| NA-LA | 18 | 5 | 54 | 23 |
| LA-AF | 2 | 74 | 13 | 11 |
| AF-LA | 1 | 83 | 10 | 6 |
| NA-AF | 19 | 0.5 | 47 | 28.5 |
| AF-NA | 1 | 83 | 10 | 6 |
| AF-AF | 1 | 83 | 10 | 6 |
| LA-LA | 2 | 74 | 13 | 11 |

6.6 Request generation

The OD demand matrix provides the percentage distribution of cargo demand directed toward each destination from a specified origin. To accurately model the actual cargo demand, we must incorporate a payload capacity for each outgoing flight. For example, assuming a payload capacity of 50 tonnes per outgoing flight, and 11 flights departing from AMS, the total export capacity would be $11 \times 50 = 550$ tonnes. Consequently, for AMS-MIA, this would translate to 134.2 tonnes of payload, while for AMS-BOG, it would result in 40.15 tonnes.

For each OD pair, we generate cargo requests until their combined weight matches the total weight specified for that pair, resulting in the set of cargo requests \mathcal{R} indexed by r . We assume some pre-consolidation between single items occurred (Brandt and Nickel, 2019). As such each cargo request

$r \in \mathcal{R}$ is characterized by a weight uniformly sampled from the range [0.4, 0.8] tonnes regardless of the category. In order to determine volume, we define density ranges per commodity type and use a uniformly sampled value within its range to translate a weight into a volume. We assume an average revenue of 4,000 €/ton, which is multiplied by a strategic factor uniformly sampled from a commodity-specific interval, as seen in Table 4, to generate higher- and lower-revenue cargo requests. Each cargo request $r \in \mathcal{R}$ is randomly assigned a release time $RT_r^{\mathcal{R}}$, which corresponds to a node at the origin airport, and a delivery deadline $DD_r^{\mathcal{R}}$, which corresponds to a node at the destination airport. A feasible path must exist between the release time node and the delivery deadline node; otherwise, we continue exploring different delivery deadline nodes until we find one that results in a acceptable combination.

Furthermore, we define three different compatibility scenarios for our model: high- (\mathbb{H}), medium- (\mathbb{M}), and low-compatibility (\mathbb{L}). In scenarios \mathbb{M} and \mathbb{L} , we introduce incompatibility between the cargo types "perishable" and "heavy," as well as "perishable" and "chemical." In the high-compatibility scenario (\mathbb{H}), no incompatibility such as constraint set (13) are imposed. For each combination of compatibility scenario and commodity type, we define a tuple (p_I, p_{II}) , where both p_I and p_{II} range between 0 and 1, representing the probability (in decimal form) that a cargo request requires Feature I (hazardous materials) or Feature II (temperature control), respectively. All cargo request characteristics per commodity types are listed in Table 4. For a given commodity type we have that $p_I(\mathbb{L}) \geq p_I(\mathbb{M}) \geq p_I(\mathbb{H})$ and $p_{II}(\mathbb{L}) \geq p_{II}(\mathbb{M}) \geq p_{II}(\mathbb{H})$ as we decrease the complexity of the packing following the $\mathbb{L} \rightarrow \mathbb{M} \rightarrow \mathbb{H}$ sequence. Note that these restrictions only apply to cargo that was labeled as hazardous or temperature-sensitive. For example, pending that incompatibility restrictions are satisfied, a non-temperature-sensitive cargo can be assigned to a temperature-controlled ULD.

By exploring and experimenting with various combinations of payloads and compatibility levels, we can gain valuable insights into the results and their implications for cargo distribution and management strategies. We will consider payloads of 35, 50, 65, and 80 tonnes per outgoing flight, alongside three compatibility levels: high (\mathbb{H}), medium (\mathbb{M}), and low (\mathbb{L}).

Table 4: Cargo request characteristics per commodity type.

| | Chemical | Perishable | Heavy | Other |
|--------------------------------|--------------|-------------|------------|--------------|
| Weight [t] | [0.4, 0.8] | [0.4, 0.8] | [0.4, 0.8] | [0.4, 0.8] |
| Density [t/m^3] | [0.15, 0.25] | [0.15, 0.2] | [0.5, 0.7] | [0.2, 0.3] |
| Strategic factor | [0.7, 1.3] | [1, 1.3] | [0.5, 0.8] | [0.9, 1.1] |
| \mathbb{H} (p_I, p_{II}) | (0.25, 0.25) | (0, 0.25) | (0.1, 0.1) | (0, 0) |
| \mathbb{M} (p_I, p_{II}) | (0.5, 0.6) | (0, 0.5) | (0.2, 0.3) | (0, 0) |
| \mathbb{L} (p_I, p_{II}) | (1, 0.6) | (0.3, 0.6) | (0.4, 0.8) | (0.25, 0.25) |

6.7 ULD generation

For each cargo request $r \in \mathcal{R}$ a corresponding ULD is generated, forming the set \mathcal{U} indexed by u . Each ULD $u \in \mathcal{U}$ inherits the OD pair, release time, and delivery deadline from the associated cargo request, meaning $RT_u^{\mathcal{U}} = RT_r^{\mathcal{R}}$ and $DD_u^{\mathcal{U}} = DD_r^{\mathcal{R}}$. In addition, each ULD $u \in \mathcal{U}$ is assigned a ULD type that is capable of accommodating the specific features required by the associated cargo request. If multiple ULD types are capable of meeting these requirements, a random selection is made among the available options. The chosen ULD type must also be able to fit the cargo request from a volume perspective.

7 Computational experiments

This section contains the computational experiments conducted to test the different solution approaches. All instances were coded in Python (version 3.12) and solved using Gurobi (version 11.0.1) on a laptop running Windows 11, equipped with an Intel Core i7-8750H CPU and 16 GB of RAM. First, in Section 7.1 we discuss a number of strategies to accelerate our solution approaches.

Then the main numerical results are listed in Section 7.2. Subsequently, Section 7.3 builds on the numerical results to provide managerial insights, and finally Section 7.4 contains a sensitivity analysis.

7.1 Algorithm accelerating strategies

Algorithm acceleration techniques can play a crucial role in enhancing the performance and efficiency of computational tasks. By reducing the time complexity and resource usage of algorithms, these techniques enable the handling of larger datasets and more complex problems within practical timeframes. Drawing on the work of Desaulniers et al. (2002), which presents various acceleration strategies for column generation (CG) methods, we apply similar techniques throughout our solution approach. This section presents key strategies utilized across different phases of our solution process. Section 7.1.1 will discuss pre-processing strategies related to instance generation, specifically how sets and parameters are constructed, which has an impact on the entire solution process. These strategies can be applied to all three solution methods: the integrated, sequential and fully arc-based approaches. In Section 7.1.2 we will describe strategies specific to the CG procedure.

7.1.1 Accelerating strategies: Pre-processing

The computational time of all approaches is highly sensitive to the number of variables and constraints in the mathematical formulation, often increasing exponentially with these factors. The most significant contributors to this complexity are the number of ULDs $|\mathcal{U}|$ and cargo requests $|\mathcal{R}|$, which are typically equal since a ULD is generated for each request. Additionally, the number of ULDs available for a given cargo request $|\mathcal{U}_r|$ and the number of cargo requests that can be transported by a given ULD $|\mathcal{R}_u|$ are critical too, as these also limit the number of variables in the formulation. Since $|\mathcal{R}_u|$ is the reverse mapping of $|\mathcal{U}_r|$, reducing $|\mathcal{U}_r|$ directly decreases $|\mathcal{R}_u|$. A key measure in this context is the average cardinality of \mathcal{R}_u over all $u \in \mathcal{U}$, expressed as: $|\mathcal{R}_{u,\text{avg}}| = \frac{1}{|\mathcal{U}|} \sum_{u \in \mathcal{U}} |\mathcal{R}_u|$. Similarly we define the average cardinality of \mathcal{U}_r over all $r \in \mathcal{R}$ as $|\mathcal{U}_{r,\text{avg}}| = \frac{1}{|\mathcal{R}|} \sum_{r \in \mathcal{R}} |\mathcal{U}_r|$. Minimizing these averages can substantially improve computational efficiency, especially in larger instances. However, it is essential to carefully consider the potential impact on the solution quality when making such reductions.

The first strategy involves **clustering** similar requests based on specific criteria. Clustering requests leads to a reduction in $|\mathcal{R}|$ and therefore also in $|\mathcal{U}|$, $|\mathcal{R}_{u,\text{avg}}|$, and $|\mathcal{U}_{r,\text{avg}}|$. The criteria for clustering include:

1. The OD pair, commodity type, hazardous classification, and temperature control requirements must be the same among the clustered requests.
2. The combined weight and volume of the clustered requests should not exceed a specified maximum. For example the maximum allowable weight and volume of a single ULD.
3. The release times and delivery deadlines of the clustered requests must be within 3 hours of each other. The newly defined cluster will inherit the earliest delivery deadline and the latest release time from its members.
4. A maximum of three requests can be clustered together.

The second strategy focuses on considering a **subset** of \mathcal{U}_r for each request. We limit the size of $|\mathcal{U}_r|$ for every $r \in \mathcal{R}$ by selecting a fixed number of the best ULDs based on the smallest absolute difference in release time and delivery deadline. Specifically, we choose ULDs that minimize the quantity $|RT_r^{\mathcal{R}} - RT_u^{\mathcal{U}}| + |DD_r^{\mathcal{R}} - DD_u^{\mathcal{U}}|$. While this procedure does not lower $|\mathcal{R}|$ or $|\mathcal{U}|$, it does effectively lower $|\mathcal{U}_{r,\text{avg}}|$ and $|\mathcal{R}_{u,\text{avg}}|$, allowing for better control over the problem's complexity.

7.1.2 Accelerating strategies: Column generation

Specifically for CG some additional accelerating strategies are considered. CG relies on iteratively adding columns to improve solutions, but this process can be computationally intense, especially as problem size grows. Effective acceleration strategies are essential for maintaining feasible solution times and ensuring high-quality results. These enhancements are not just improvements but can be vital for the practical application of CG methods.

Our first CG accelerating strategy focuses on **diversification**. Adding multiple columns to the RMP from each pricing problem instance can reduce the number of iterations needed. However, this approach can significantly expand the RMP, even though the optimal solution ultimately includes

only a few of these columns. The selection of columns added to the RMP greatly influences the total number of variables and computing time required. To prevent the RMP from growing excessively, it is crucial to generate a set of the best and most diverse solutions. This can be achieved by only accepting new columns from the same pricing problem instance if they contribute to a diverse set of constraints. Two possible approaches for achieving this are:

1. Each new path must include at least one new request in its packing component. For instance, a path with packing component [a, b, c] will not be accepted if there already exist paths from the current pricing problem instance containing **all** of [a, b, c].
2. There must be no overlap in packing components. For example, a path with packing component [a, b, c] will not be accepted if there are already paths from the current pricing problem instance containing **any** of [a, b, c].

Additionally, we can enforce diversification by setting a limit on the number of paths generated for each request overall. We can remove the request from \mathcal{R}_u for all $u \in \mathcal{U}$ once the threshold is reached, thereby forcing the pricing problem to search for paths with different requests. This procedure is hereafter referred to as **cap**.

A different CG accelerating strategy is **pricing problem aggregation**. While the ULDs and their attributes are generated randomly it is common for many ULDs to have exact duplicates. In fact, depending on the test instance, 30-55% of ULDs have at least one exact duplicate. Instead of recalculating the pricing problem for each duplicate ULD, we can streamline the process by copying previously computed paths for these identical ULDs. Since the dual values are consistent across all ULDs (with the exception of the constant η_u) and \mathcal{R}_u is the same for all identical ULDs, this approach is viable. We simply need to modify the paths to reflect the correct source and sink arcs for each specific ULD. By doing this, we can significantly reduce computational effort while maintaining accuracy the solution.

7.1.3 Accelerating strategies: Experimental results

This subsection presents the experimental results of various acceleration strategies applied to our CG algorithm. We tested different pre-processing techniques and CG-specific enhancements to improve computational efficiency while maintaining solution quality. All instances use the reduced schedule in the *restricted* scenario, perform 5 full CG iterations, have compatibility level \mathbb{L} , and assume a payload of 50 tonnes per outgoing flight. To ensure the problem can scale to larger instances, such as the full-sized schedule, we focus on limiting computational time in the algorithmic tests, while also ensuring that solution quality is not compromised.

Table 5: Comparison between different pre-processing strategies.

| Instances: | P [M€] | BB [M€] | Gap [%] | $ \mathcal{R}_{u,avg} $ | $ \mathcal{U} $ | $ \mathcal{P} $ | T_{CG} [s] | T_{RIC} [s] |
|-------------------|--------|---------|---------|-------------------------|-----------------|-----------------|--------------|---------------|
| No pre-processing | 1.4395 | 1.4501 | 0.7364 | 50.95 | 1342 | 58230 | 240.90 | 3350 |
| Clustering LD7 | 1.2906 | 1.2964 | 0.4532 | 22.33 | 731 | 18174 | 70.00 | 3 |
| Clustering LD3 | 1.4500 | 1.4572 | 0.4999 | 40.56 | 1183 | 65941 | 197.01 | 3146 |
| Subset 30 | 1.4598 | 1.4686 | 0.6036 | 30.39 | 1342 | 54640 | 161.57 | 3438 |
| Subset 15 | 1.4525 | 1.4597 | 0.4990 | 17.58 | 1342 | 47871 | 93.14 | 3321 |

We present the results of various pre-processing strategies across five different instances in Table 5. The first instance serves as a benchmark with no pre-processing applied. The subsequent instances explore clustering methods, with the maximum combined weight and volume of the clusters set to the capacities of different ULD types. Specifically, we analyze clustering where the limits match the capacities of the larger LD7 ULD type and the smaller LD3 ULD type, respectively. Additionally, we implement a subset strategy for \mathcal{U}_r testing limits of 30 and 15. For each method, we report the final objective value, i.e. profit, after one hour of computation, denoted as P (in million euros), the best bound resulting from the branch-and-bound (BB), and the optimality gap. The best bound represents the upper limit of the objective value found during branch-and-bound. It provides an indication that no integer feasible solution can exceed this value, helping gauge how close the current solution is to the optimal one. Furthermore we report the number of ULDs $|\mathcal{U}|$, number of paths $|\mathcal{P}|$ and the average number of requests available per ULD $|\mathcal{R}_{u,avg}|$. The final columns contain the

computational time for 5 full iterations of the CG algorithm (T_{CG}), and the time taken to find the optimal integer solution after reinstating the integrality constraints (T_{RIC}). Their combined time should not exceed one hour.

While clustering reduces $|\mathcal{U}|$, $|\mathcal{R}|$, and $|\mathcal{R}_{u,avg}|$, resulting in faster computational time, it also incurs significant losses in flexibility and efficiency, ultimately harming the objective value. This is particularly evident in the clustering LD7 instance, where the profit drops considerably. The clustering process makes it challenging to fully utilize ULD capacity, as most clusters are characterized by larger weights and volumes. This is especially problematic for smaller ULDs. Conversely, the subset method results in a significantly lower $|\mathcal{R}_{u,avg}|$ without a notable detriment to the profit. In fact, focusing on a limited number of ULDs per request, specifically those with the most similar release times and delivery deadlines, seems to slightly improve the quality of the generated paths. Additionally, a clear relationship between $|\mathcal{U}|$, $|\mathcal{R}_{u,avg}|$, and T_{CG} is evident across all instances. Limiting the computational time for the CG algorithm is particularly important for the scalability of the solution method to larger instances.

Based on these results, we adopt the subset pre-processing step for all subsequent instances, as it offers the best trade-off between computational time and solution quality.

Table 6: Comparison between different CG acceleration strategies.

| Instances: | P [M€] | BB [M€] | Gap [%] | $ \mathcal{P} $ | T_{CG} [s] |
|--------------------------|---------------|---------------|-------------|-----------------|--------------|
| Subset 30 (base, 5 it) | 1.4598 | 1.4686 | 0.60 | 54640 | 161.57 |
| Subset 30 cap 50 | 1.4473 | 1.4552 | 0.54 | 22521 | 66.85 |
| Subset 30 cap 200 | 1.4619 | 1.4690 | 0.49 | 53870 | 114.53 |
| Subset 30 cap 400 | 1.4627 | 1.4705 | 0.53 | 71034 | 159.27 |
| Subset 30 div. all | 1.4499 | 1.4618 | 0.82 | 50779 | 157.05 |
| Subset 30 div. any | 1.4025 | 1.4095 | 0.50 | 16994 | 130.38 |
| Subset 30 2 it cap 400 | 1.4208 | 1.4273 | 0.46 | 41245 | 56.31 |
| Subset 30 10 it cap 400 | 1.4666 | 1.4735 | 0.47 | 85721 | 325.12 |
| Subset 15 (base, 5 it) | 1.4525 | 1.4597 | 0.50 | 47871 | 93.14 |
| Subset 15 cap 50 | 1.4407 | 1.4468 | 0.42 | 21032 | 48.93 |
| Subset 15 cap 200 | 1.4540 | 1.4610 | 0.49 | 48400 | 75.81 |
| Subset 15 cap 400 | 1.4537 | 1.4608 | 0.49 | 59610 | 94.40 |
| Subset 15 div. all | 1.4425 | 1.4492 | 0.46 | 36110 | 94.04 |
| Subset 15 div. any | 1.3949 | 1.4019 | 0.50 | 12946 | 86.04 |
| Subset 15 2 it cap 400 | 1.4281 | 1.4341 | 0.42 | 37615 | 33.63 |
| Subset 15 10 it cap 400 | 1.4561 | 1.4634 | 0.50 | 67690 | 201.06 |

Table 6 presents the results of various CG-specific acceleration strategies, all implemented with the subset pre-processing method. For each subset limit, the first instance serves as the baseline, featuring only pre-processing without any additional CG acceleration techniques.

The **cap** strategy, which limits the number of paths generated per request, shows a trade-off between solution quality and computational time. Note that enforcing this limit does not necessarily lead to a reduction in the total number of paths generated. As we remove requests from \mathcal{R}_u the pricing problem is forced to find solutions that include other requests, leading to fewer duplicate paths among the 20 best solutions. Consequently, more new paths being accepted into the set \mathcal{P} . We observe slightly improved profits as the cap value increases at the cost of longer computation times. The cap 200 instance achieves a good balance, maintaining solution quality while significantly reducing computational time compared to the baseline case.

The **diversification** strategies (all and any) aim to generate a more diverse set of paths. The diversification all approach maintains reasonable solution quality while reducing computation time. However, the diversification any strategy, while fastest, results in a significant drop in the profit value, suggesting that it may be too restrictive.

Varying the number of iterations, such as 2 iterations or 10 iterations demonstrates the impact of iteration count on solution quality and computational time. While fewer iterations drastically reduce computation time, it comes at the cost of solution quality. Conversely, more iterations slightly improve the objective value but at the expense of longer computation times.

In all instances, **pricing problem aggregation** was utilized. For a randomly selected instance within the set of algorithmic tests, the average execution time for solving the pricing problem was 2.792×10^{-2} seconds, and this problem was executed 14,670 times. In contrast, simply copying the existing paths from an identical ULD occurred 5,225 times and took only 2.527×10^{-4} seconds on average. This demonstrates a significant improvement in computational speed, specifically a 26% improvement over the computational time of the total CG algorithm. While this approach is beneficial from a computational standpoint, copying paths from identical ULDs introduces a potential vulnerability to symmetrical solutions. Since these copied paths are essentially equivalent in terms of their contribution to the objective function, they can lead to multiple equivalent solutions that differ only in which of the identical ULDs is being used.

In conclusion, the subset pre-processing method, limiting 15 ULDs per request combined with a moderate cap value (e.g., cap 200), which is highlighted in bold in Table 6, appears to offer the best balance between solution quality and computational efficiency for our CG algorithm. This combination significantly reduces computation time while maintaining high-quality solutions, making it a promising approach for larger, more complex instances.

7.2 Numerical results

In Table 7, we report the final results for 12 test instances, comparing the performance of the integrated, sequential, and fully arc-based approaches. Each instance had a combined runtime of one hour, which includes the time for CG iterations and reinstating integrality constraints for the integrated approach, and the combined time for the first and second stages in the sequential approach. The results for the sequential approach reported in Table 7 correspond to the second stage of the method. The test instances encompassed various compatibility levels (denoted by \mathbb{H} for high, \mathbb{M} for medium and \mathbb{L} for low) and take different assumed payloads per outgoing flight (indicated by the numbers 35, 50, 65, and 80). All instances utilized the full-sized schedule within the restricted scenario, leading us to apply the simplified linearization method for the pricing problem, as outlined in (63). Additionally, we applied a pre-processing step by selecting a subset of 15 ULDs per request. For the CG method, we imposed a cap of 200 paths per request, execute 5 full iterations, make use of pricing problem aggregation, but do not apply any additional diversification techniques.

Table 7: Comparison between the integrated, sequential, and fully arc-based solution methods for the 12 restricted instances.

| Inst. | Integrated | | | Sequential | | | Arc-based | | |
|-------|--------------|---------|---------|------------|---------|---------|-----------|---------|---------|
| | P [M€] | BB [M€] | Gap [%] | P [M€] | BB [M€] | Gap [%] | P [M€] | BB [M€] | Gap [%] |
| H35 | 1.793 | 1.828 | 1.92 | 1.513 | 1.583 | 4.67 | 1.226 | 2.169 | 76.9 |
| H50 | 3.486 | 3.543 | 1.65 | 3.201 | 3.402 | 6.28 | 2.726 | 4.251 | 55.9 |
| H65 | 4.801 | 4.869 | 1.42 | 4.546 | 4.803 | 5.65 | 2.814 | 6.269 | 123 |
| H80 | 5.864 | 5.919 | 0.93 | 5.534 | 5.844 | 5.62 | 2.224 | 8.120 | 265 |
| M35 | 1.679 | 1.694 | 0.88 | 1.350 | 1.482 | 9.76 | 1.403 | 2.018 | 43.8 |
| M50 | 3.349 | 3.375 | 0.78 | 2.988 | 3.228 | 8.06 | 2.903 | 4.077 | 40.4 |
| M65 | 4.662 | 4.702 | 0.85 | 4.247 | 4.635 | 9.13 | 3.825 | 6.694 | 75.0 |
| M80 | 5.707 | 5.738 | 0.54 | 5.295 | 5.702 | 7.69 | -3.694 | 9.049 | 345 |
| L35 | 1.623 | 1.638 | 0.90 | 1.315 | 1.406 | 6.93 | 1.153 | 2.005 | 73.9 |
| L50 | 3.292 | 3.327 | 1.05 | 2.951 | 3.181 | 7.80 | 2.528 | 4.208 | 66.5 |
| L65 | 4.601 | 4.640 | 0.85 | 4.271 | 4.590 | 7.46 | -3.674 | 6.624 | 280 |
| L80 | 5.632 | 5.673 | 0.71 | 5.168 | 5.625 | 8.83 | 4.476 | 8.394 | 87.6 |

The integrated approach outperforms the other approaches in all instances, as evidenced by the boldface profit values. This outperformance is more pronounced in scenarios with lower cargo compatibility and/or lower cargo demand. Specifically, the performance improvement ranges from 5.62% in the H65 instance to 24.42% in the M35 instance. The integrated approach is solved to good optimality gaps within the hour runtime, with many gaps below 1%. This suggests that for the paths generated, we can safely assume these are very close to the optimal solutions possible. To estimate the performance of the integrated model relative to the exact optimal solution, a subset of instances in the arc-based formulation was allowed to run for 12 hours to determine its best bounds

as is shown in Table 8. For the instances $\mathbb{H}50$, $\mathbb{M}50$, and $\mathbb{L}50$, these bounds were 3.646, 3.482, and 3.427 million euros, respectively. This means the relative gaps from the profits achieved by the integrated method after one hour are approximately 4%, compared to the exact best bounds after 12 hours. These results underline that the integrated method provides near-optimal solutions in excellent computational time.

The gap for the sequential approach is significantly larger after an hour of runtime, ranging from about 5% to 10%, indicating there could still be some room for improvement in the second stage of the solution method. However, the best bounds for the sequential approach are still lower than the objective value of the counterpart for the integrated approach in all but one instance ($\mathbb{H}65$). This means that even if the best bound turns out to be the final optimum for the sequential approach, it would still underperform compared to the integrated approach, further underlining the superior performance of the integrated method.

The fully arc-based formulation generally underperforms in Table 7 with respect to the other two methods. In some cases with lower compatibility and higher demand, it is not even able to find a positive profit solution, as seen in the negative profit values for $\mathbb{M}80$ and $\mathbb{L}65$. The optimality gaps for the arc-based method are extremely large, occasionally even exceeding 100%, indicating that this method struggles to find high-quality solutions within the given time limit.

In summary, the results clearly demonstrate the superiority of the integrated approach across all tested scenarios, with the sequential approach performing reasonably well but still falling short, and the arc-based formulation proving to be inadequate for large complex instances.

Table 8: Comparison of the integrated solution method (1 hour) and the fully arc-based method (12 hours) across three instances.

| Inst. | Integrated (1 hour) | | | Arc-based (12 hours) | | | Integrated P vs. Arc-based BB | | |
|----------------|---------------------|---------|---------|----------------------|---------|---------|-------------------------------|---------|---------|
| | P [M€] | BB [M€] | Gap [%] | P [M€] | BB [M€] | Gap [%] | P [M€] | BB [M€] | Gap [%] |
| $\mathbb{H}50$ | 3.486 | 3.543 | 1.65 | 2.921 | 3.646 | 24.8 | 3.486 | 3.646 | 4.59 |
| $\mathbb{M}50$ | 3.349 | 3.375 | 0.78 | 2.912 | 3.482 | 19.6 | 3.349 | 3.482 | 3.97 |
| $\mathbb{L}50$ | 3.292 | 3.327 | 1.05 | 2.683 | 3.427 | 27.7 | 3.292 | 3.427 | 4.10 |

7.3 Managerial insights

In Table 9, a detailed breakdown of the profit for the different instances is presented, along with additional data offering insights into the revenue and costs associated with the operations. The profit consists of the revenue generated by transporting cargo, denoted as $R^{\mathcal{R}}$, minus the operational costs for both aircraft and ULD transportation, represented by $C^{\mathcal{FK}}$ and $C^{\mathcal{FU}}$ respectively. The percentage of cargo requests that are transported is denoted by $\%_{tr}^{\mathcal{R}}$, while the total number of transported cargo requests is represented by $|\mathcal{R}_{tr}|$. The total number of ULDs utilized is indicated by $|\mathcal{U}_{tr}|$, and $v_{avg}^{\mathcal{U}}$ reflects the average ULD volume factor as this was often a limiting constraint during the bin packing problems. Finally, Φ is the percentage of non-special cargo requests transported in an unnecessarily advanced ULD. This refers to cases where cargo without special handling needs is placed in a more expensive ULD designed for such special cargo (e.g., LD3+ or LD7+). It is important to note that minimizing Φ is not an absolute priority. In many cases, it is still more beneficial to transport a non-special request in an already assigned LD3+ or LD7+ ULD rather than using an additional standard ULD or, worse, leaving the cargo unassigned. Nonetheless, when the number of transported requests is similar across instances, a lower Φ is an indicator of more efficient ULD type assignment.

As shown in Table 9, the integrated model consistently outperforms the sequential approach across all instances. This is primarily driven by two factors: increased revenue and reduced ULD operational costs.

The revenue improvement is consistent across all instances, with the integrated approach yielding an increase of approximately 2-4 percentage points in $\%_{tr}^{\mathcal{R}}$ compared to the sequential approach for each instance pair. The improvement is a result of the integrated decision-making process, which allows the bin-packing problems for all ULDs to be solved more efficiently. This is evident by the higher volume factors, indicating that the ULDs are packed more effectively to their capacity. Consequently, more cargo can be transported, leading to higher overall revenue.

Table 9: Profit breakdown of the integrated and sequential solution methods for the 12 restricted instances.

| | Inst. | P [€] | R^R [€] | C^{FK} [€] | C^{FU} [€] | $\%R_{tr}^R$ [%] | $ R_{tr} $ | $ U_{tr} $ | v_{avg}^{fU} [%] | Φ [%] |
|------|-------|-----------|------------|--------------|--------------|------------------|------------|------------|--------------------|------------|
| Int. | H35 | 1,793,198 | 6,346,160 | 3,586,592 | 966,369 | 95.51 | 2683 | 872 | 92.94 | 26.22 |
| | H50 | 3,485,699 | 8,169,040 | 3,586,592 | 1,096,747 | 87.81 | 3507 | 1111 | 95.12 | 26.68 |
| | H65 | 4,800,808 | 9,557,160 | 3,586,592 | 1,169,758 | 78.42 | 4034 | 1319 | 95.31 | 24.50 |
| | H80 | 5,864,252 | 10,712,400 | 3,606,212 | 1,241,934 | 71.90 | 4550 | 1489 | 95.84 | 24.52 |
| Seq. | H35 | 1,512,717 | 6,197,520 | 3,586,593 | 1,098,210 | 92.42 | 2596 | 936 | 86.82 | 26.54 |
| | H50 | 3,201,182 | 8,023,040 | 3,601,477 | 1,220,382 | 85.90 | 3431 | 1165 | 89.84 | 28.59 |
| | H65 | 4,545,542 | 9,497,800 | 3,606,213 | 1,346,045 | 77.24 | 3973 | 1379 | 91.13 | 24.76 |
| | H80 | 5,533,634 | 10,571,040 | 3,622,016 | 1,415,390 | 69.72 | 4412 | 1576 | 91.25 | 26.33 |
| Int. | M35 | 1,679,378 | 6,284,600 | 3,586,592 | 1,018,628 | 94.52 | 2655 | 887 | 89.76 | 38.54 |
| | M50 | 3,348,962 | 8,083,160 | 3,589,140 | 1,145,057 | 86.55 | 3457 | 1125 | 91.86 | 41.19 |
| | M65 | 4,662,153 | 9,488,440 | 3,586,592 | 1,239,693 | 77.70 | 3997 | 1319 | 92.93 | 39.13 |
| | M80 | 5,706,930 | 10,619,600 | 3,606,212 | 1,306,456 | 71.10 | 4499 | 1490 | 93.73 | 38.13 |
| Seq. | M35 | 1,349,774 | 6,103,640 | 3,586,593 | 1,167,273 | 90.49 | 2542 | 976 | 83.78 | 43.95 |
| | M50 | 2,987,626 | 7,925,520 | 3,613,083 | 1,324,812 | 84.65 | 3381 | 1220 | 85.17 | 44.53 |
| | M65 | 4,247,176 | 9,268,440 | 3,606,213 | 1,415,051 | 75.16 | 3866 | 1407 | 86.79 | 43.26 |
| | M80 | 5,294,782 | 10,417,480 | 3,633,622 | 1,489,077 | 68.44 | 4331 | 1599 | 87.92 | 44.22 |
| Int. | L35 | 1,623,330 | 6,282,560 | 3,586,592 | 1,072,636 | 94.45 | 2653 | 877 | 90.24 | 62.57 |
| | L50 | 3,292,370 | 8,083,560 | 3,589,140 | 1,202,049 | 86.45 | 3453 | 1122 | 92.23 | 63.31 |
| | L65 | 4,600,712 | 9,495,360 | 3,586,592 | 1,308,055 | 77.78 | 4001 | 1332 | 93.17 | 56.82 |
| | L80 | 5,632,494 | 10,624,480 | 3,606,212 | 1,385,772 | 71.21 | 4506 | 1502 | 93.62 | 56.48 |
| Seq. | L35 | 1,314,749 | 6,116,480 | 3,586,593 | 1,215,138 | 90.89 | 2553 | 966 | 83.92 | 65.89 |
| | L50 | 2,950,631 | 7,901,800 | 3,601,477 | 1,349,692 | 84.30 | 3367 | 1203 | 85.72 | 64.04 |
| | L65 | 4,271,112 | 9,349,320 | 3,606,213 | 1,471,996 | 75.74 | 3896 | 1394 | 87.81 | 60.79 |
| | L80 | 5,168,281 | 10,342,320 | 3,622,016 | 1,552,023 | 67.84 | 4293 | 1589 | 87.62 | 65.41 |

In terms of ULD costs, the integrated model offers two key advantages. First, it uses fewer ULDs, with a reduction of 60-100 units, translating to a 4-9% decrease. This reduction occurs while transporting more cargo requests, highlighting the superior efficiency of the integrated approach. The approach optimizes cargo routing while simultaneously considering bin-packing and compatibility constraints, allowing it to allocate more cargo to fewer ULDs. In contrast, the sequential approach assigns cargo to a flight in the first stage based on total aircraft (payload) weight and volume constraints without considering any bin-packing constraints. Only in the second stage does it become apparent that much of the cargo may have incompatible commodity types, differing OD pairs, wide-ranging release times and delivery deadlines, or weights and volumes that awkwardly fit into a single ULD. These issues result in the need for more ULDs to transport less cargo.

The second key advantage is that the integrated model selects the ULD types more efficiently, as shown by the lower values for Φ . This is particularly evident in instances with high cargo demand and low compatibility. For example, in the L80 instance, the percentage of non-special requests transported in an unnecessarily advanced ULD decreases from 65.41% in the sequential approach to 56.48% in the integrated approach.

Finally, it is worth noting that the integrated model also outperforms the sequential model in terms of aircraft operating costs. However, this gain is less significant than the improvements in revenue and ULD costs. This is largely because every flight arc is mandatory, and there is only a slight difference in operating costs between the two available aircraft types.

7.3.1 Managerial insights: Myopic first stage solution of the sequential approach

In the first stage of the sequential approach, the simplification of packing constraints creates a myopic view of the problem, leading to an overestimation of potential revenue. Once real packing constraints are reinstated in the second stage a significant portion of the assigned requests may become infeasible.

In Table 10 we compare the revenue generated in both stages, highlighting the reduction once packing constraints are applied. In the first stage, where these constraints are absent, revenue is the same across different compatibility levels. This is because the aircraft's overall capacity, consisting of

the payload weight and combined weight-volume capacities of available ULD slots, is the only limiting factor on how many cargo requests can be accommodated. Once the actual packing constraints are applied in the second stage, the revenue drops. For the \mathbb{L} and \mathbb{M} compatibility levels, the percentage point decrease in $\%_{tr}^{\mathcal{R}}$ between the two stages ranges from 4 to 7 points, while for \mathbb{H} the decrease is slightly smaller, between 2.5 and 5 points. The optimistic revenue projections from the first stage result in a reduction of approximately €300,000 to €600,000, translating to a 3-6% revenue loss.

Table 10: Comparison between the myopic revenue and the actual revenue for the sequential approach.

| Inst. | 1st Stage | | | 2nd Stage | | | Revenue Loss | |
|-------|-----------------------|-----------------------------|----------------------|-----------------------|-----------------------------|----------------------|--------------------------------------|------------------------------|
| | $R^{\mathcal{R}}$ [€] | $\%_{tr}^{\mathcal{R}}$ [%] | $ \mathcal{R}_{tr} $ | $R^{\mathcal{R}}$ [€] | $\%_{tr}^{\mathcal{R}}$ [%] | $ \mathcal{R}_{tr} $ | $\Delta\%_{tr}^{\mathcal{R}}$ [%pt.] | $\Delta R^{\mathcal{R}}$ [%] |
| H35 | 6,495,680 | 97.44 | 2737 | 6,197,520 | 92.42 | 2596 | -5.02 | -4.59 |
| H50 | 8,332,760 | 89.46 | 3573 | 8,023,040 | 85.90 | 3431 | -3.56 | -3.72 |
| H65 | 9,815,160 | 80.19 | 4125 | 9,497,800 | 77.24 | 3973 | -2.95 | -3.23 |
| H80 | 10,951,800 | 72.44 | 4584 | 10,571,040 | 69.72 | 4412 | -2.72 | -3.48 |
| M35 | 6,495,680 | 97.44 | 2737 | 6,103,640 | 90.49 | 2542 | -6.94 | -6.04 |
| M50 | 8,332,760 | 89.46 | 3573 | 7,925,520 | 84.65 | 3381 | -4.81 | -4.89 |
| M65 | 9,815,160 | 80.19 | 4125 | 9,268,440 | 75.16 | 3866 | -5.03 | -5.57 |
| M80 | 10,951,800 | 72.44 | 4584 | 10,417,480 | 68.44 | 4331 | -4.00 | -4.88 |
| L35 | 6,495,680 | 97.44 | 2737 | 6,116,480 | 90.89 | 2553 | -6.55 | -5.84 |
| L50 | 8,332,760 | 89.46 | 3573 | 7,901,800 | 84.30 | 3367 | -5.16 | -5.17 |
| L65 | 9,815,160 | 80.19 | 4125 | 9,349,320 | 75.74 | 3896 | -4.45 | -4.75 |
| L80 | 10,951,800 | 72.44 | 4584 | 10,342,320 | 67.84 | 4293 | -4.60 | -5.57 |

The higher aircraft operating costs of the sequential approach can also be attributed to the myopic nature of the first stage. Due to the overestimation of the cargo a flight can accommodate, the model tends to select the more expensive aircraft type, which has a higher capacity. However, after introducing the real packing constraints in the second stage, it becomes clear that the larger aircraft was not actually necessary, leading to unnecessary operating costs.

7.4 Sensitivity analysis

To assess the robustness of the solution methods under different problem settings, a sensitivity analysis was conducted by running all 12 instances again, this time under the free scenario. In this scenario, each flight arc becomes optional, introducing a flight selection decision to the problem and thereby increasing its complexity. We maintain consistent conditions regarding CG algorithm specifics, accelerating strategies, and computational time.

Table 11: Comparison between the integrated, sequential, and fully arc-based solution methods for the 12 free instances, with 1 hour computational time.

| Inst. | Integrated | | | Sequential | | | Arc based | | |
|-------|--------------|---------|---------|--------------|---------|---------|-----------|---------|---------|
| | P [M€] | BB [M€] | Gap [%] | P [M€] | BB [M€] | Gap [%] | P [M€] | BB [M€] | Gap [%] |
| H35 | 1.922 | 2.453 | 27.67 | 1.993 | 2.095 | 5.09 | 1.193 | 3.935 | 230 |
| H50 | 3.429 | 3.925 | 14.45 | 3.501 | 3.697 | 5.59 | 0.251 | 5.440 | 2070 |
| H65 | 4.792 | 5.157 | 7.62 | 4.667 | 4.936 | 5.77 | 0 | 7.328 | - |
| H80 | 5.787 | 6.017 | 3.97 | 5.545 | 5.870 | 5.85 | 0 | 9.433 | - |
| M35 | 1.837 | 2.324 | 26.49 | 1.851 | 2.027 | 9.50 | 1.349 | 3.814 | 183 |
| M50 | 3.332 | 3.742 | 12.30 | 3.301 | 3.557 | 7.75 | 0.728 | 5.555 | 664 |
| M65 | 4.634 | 4.975 | 7.37 | 4.495 | 4.802 | 6.83 | 0 | 8.427 | - |
| M80 | 5.650 | 5.864 | 3.80 | 5.316 | 5.695 | 7.13 | 1.611 | 9.302 | 477 |
| L35 | 1.810 | 2.273 | 25.62 | 1.799 | 1.978 | 9.94 | 1.504 | 3.753 | 149 |
| L50 | 3.244 | 3.720 | 14.70 | 3.239 | 3.524 | 8.81 | 0.733 | 5.450 | 643 |
| L65 | 4.568 | 4.891 | 7.08 | 4.465 | 4.739 | 6.14 | 0 | 8.359 | - |
| L80 | 5.562 | 5.787 | 4.05 | 5.252 | 5.608 | 6.78 | 0 | 10.011 | - |

The results of this analysis, presented in Table 11, indicate that while the performance gap between the integrated and sequential approaches narrows in the free scenario, the integrated approach

still generally outperforms the sequential one. However, it is important to note that the clear advantage of the integrated approach seen in Section 7.2 diminishes here significantly, with the sequential approach even yielding superior results in some instances.

The increased complexity of the free scenario significantly impacts the integrated approach as it has to deal with an expanded decision space. This results in a larger optimality gap than in the restricted scenario. The sequential approach, however, remains largely unaffected in terms of optimality gap, as the added flight selection complexity only impacts the first stage, which is computationally straightforward. The computational intensity of the sequential approach is determined primarily by its second stage, and thus remains stable.

To further assess performance under extended computational time, a subset of instances was resolved with a 6 hour runtime for both the integrated and sequential approaches (the original arc-based method was excluded due to its poor performance in the initial tests). These results are presented in Table 12 along with a profit breakdown. The integrated approach continues to deliver superior performance (except for H50), with slightly better profits compared to the sequential method. The optimality gap has improved across all instances due to the extended computational time. A new metric introduced in the table is \mathcal{F}_0 , representing the number of flight arcs opted not to operate.

Table 12: Comparison of results between the integrated and sequential solution methods for 9 free instances, with 6 hour computational time, including a profit breakdown.

| | Inst. | P [M€] | BB [M€] | Gap [%] | R^R [€] | C^{FK} [€] | C^{FU} [€] | $\%_{tr}^R$ [%] | $ \mathcal{U}_{tr} $ | $ \mathcal{F}_0 $ |
|------|-------|--------------|---------|---------|------------|--------------|--------------|-----------------|----------------------|-------------------|
| Int. | H50 | 3.520 | 3.884 | 10.30 | 7,559,360 | 3,020,427 | 1,018,899 | 81.12 | 1058 | 5 |
| | H65 | 4.858 | 4.988 | 2.67 | 9,214,560 | 3,195,316 | 1,161,057 | 75.66 | 1296 | 3 |
| | H80 | 5.831 | 5.924 | 1.60 | 10,353,280 | 3,330,037 | 1,191,991 | 69.48 | 1457 | 2 |
| | M50 | 3.379 | 3.505 | 3.72 | 7,472,080 | 3,020,427 | 1,072,624 | 80.15 | 1092 | 5 |
| | M65 | 4.675 | 4.783 | 2.31 | 8,895,160 | 3,065,673 | 1,153,998 | 72.76 | 1278 | 4 |
| | M80 | 5.686 | 5.744 | 1.02 | 10,279,560 | 3,335,033 | 1,258,327 | 68.87 | 1468 | 2 |
| | L50 | 3.321 | 3.520 | 5.99 | 7,449,920 | 3,011,772 | 1,116,849 | 79.57 | 1093 | 5 |
| | L65 | 4.609 | 4.745 | 2.94 | 8,701,560 | 2,931,478 | 1,160,961 | 70.78 | 1267 | 5 |
| | L80 | 5.593 | 5.673 | 1.43 | 10,095,080 | 3,210,200 | 1,292,045 | 67.34 | 1458 | 3 |
| Seq. | H50 | 3.538 | 3.620 | 2.32 | 7,741,680 | 3,055,753 | 1,148,183 | 81.90 | 1120 | 5 |
| | H65 | 4.733 | 4.859 | 2.67 | 9,207,880 | 3,227,631 | 1,247,485 | 74.42 | 1328 | 3 |
| | H80 | 5.621 | 5.768 | 2.62 | 10,645,480 | 3,615,629 | 1,408,992 | 70.32 | 1584 | 0 |
| | M50 | 3.356 | 3.457 | 3.00 | 7,628,960 | 3,055,754 | 1,216,832 | 80.42 | 1152 | 5 |
| | M65 | 4.571 | 4.713 | 3.12 | 9,120,120 | 3,238,657 | 1,310,536 | 73.44 | 1343 | 3 |
| | M80 | 5.397 | 5.576 | 3.32 | 10,502,120 | 3,626,656 | 1,478,726 | 69.15 | 1578 | 0 |
| | L50 | 3.302 | 3.413 | 3.35 | 7,610,680 | 3,055,754 | 1,252,839 | 80.37 | 1143 | 5 |
| | L65 | 4.519 | 4.653 | 2.97 | 9,110,960 | 3,227,631 | 1,364,668 | 73.46 | 1355 | 3 |
| | L80 | 5.340 | 5.518 | 3.33 | 10,516,880 | 3,626,656 | 1,550,095 | 69.12 | 1579 | 0 |

In high-demand scenarios, the integrated approach consistently operates fewer flight arcs than the sequential approach. This difference stems from the tendency of the sequential approach to overestimate revenue in its first stage, where flight selection decisions are made without full packing constraints. This initial overestimation makes more flights appear profitable, resulting in more arcs being selected for operation. However, when packing constraints are reintroduced in the second stage, a significant percentage of the projected revenue disappears.

In contrast, the integrated approach makes flight selection and bin packing decisions simultaneously, providing a more accurate view of how much cargo each flight can realistically carry. This realism allows the integrated model to avoid operating certain flight arcs, saving on expensive flight operating costs, while still transporting a sufficient number of requests. Furthermore, the integrated decision-making is inherently more efficient in minimizing ULD costs compared to the sequential approach. Interestingly, in cases where both models operate the same number of flights, the sequential approach may yield slightly higher revenue. Nonetheless, the integrated model remains more advantageous overall, as its reduced ULD costs more than offset the revenue difference.

In conclusion, the sensitivity analysis reveals that adding the complexity of flight selection decisions significantly impacts the integrated approach, as its decision space expands. While it still

outperforms the sequential method, the clear advantage observed in the restricted scenario diminishes. The sequential approach is less affected, benefiting from a simpler first stage problem. Nonetheless, the integrated model remains adept at opting not to operate certain flight arcs, leading to worthwhile cost reductions.

8 Conclusion

This work developed an integrated model for cargo operations with the primary goal of addressing fleet assignment within a full freighter network, ULD routing, and cargo assignment to ULDs simultaneously. Two mathematical formulations were presented: an arc-based formulation and a hybrid arc- and path-based formulation. The arc-based formulation models the routing of all three commodities as sequences of arcs. In contrast, the hybrid formulation employs an arc-based approach for routing aircraft while using path-based decision variables to route ULDs containing requests.

We explored three distinct solution approaches: the integrated, the sequential, and the original fully arc-based methods. The integrated approach simultaneously addresses fleet assignment, ULD routing, and cargo packing. It is based on the hybrid formulation and employs a column generation algorithm to identify beneficial paths. In contrast, the sequential approach consists of two stages. The first stage focuses on fleet assignment and the allocation of requests to flights, intentionally overlooking the assignment of cargo requests to ULDs. In the second stage, cargo requests are assigned to ULDs and ULDs are allocated to flights, thereby reinstating the bin-packing and compatibility constraints.

To evaluate the models, numerous instances were generated using synthetic datasets that aim to reflect actual air cargo operations. The instance generation is subject to various cargo demand levels, compatibility levels (high, medium, low), and scenarios (full vs. reduced schedules, free vs. restricted aircraft placement).

Twelve instances were tested across the three solution methods: the integrated, sequential, and exact arc-based approaches. The results clearly indicated the superiority of the integrated approach, consistently achieving the highest profit values and solving close to or even below a 1% optimality gap in under one hour. Profit improvements ranged from 5.62% to 24.42% over the sequential approach. While the sequential method demonstrated reasonable performance, its best bound values remained below the profit achieved by the integrated method, indicating that even with additional runtime, it would not surpass the integrated approach. The arc based method significantly underperformed, sometimes even yielding non-profitable solutions with optimality gaps exceeding 100%.

The integrated approach also excelled in handling compatibility and bin-packing constraints, transporting 2-4% more of the demanded cargo and using 4-9% fewer ULDs. This efficiency was also reflected in better ULD selection, with fewer non-special requests transported in expensive advanced ULDs. In contrast, the sequential approach, was limited by the myopic results of the first stage after using oversimplified packing constraints, leading to overestimated revenue projections. There was a significant revenue drop of 3-6% when real packing constraints are applied.

A sensitivity analysis incorporating flight selection added complexity to the problem. While the integrated approach typically outperformed the sequential method, the advantage was greatly reduced compared to the previous restricted scenario. In certain instances, the sequential approach even yielded better results. The integrated model faced larger optimality gaps due to the added complexity, while the sequential model was able to maintain a similar gap, as flight selection occurred in the simpler first stage of the solution process. Although extended computational time enhanced performance for both methods, the improvements were not groundbreaking.

Based on our findings, we recommend implementing the integrated model across all freighter operations when working with a fixed schedule. The integrated approach consistently outperformed the sequential model, generating higher revenue and lowering ULD costs. It proved to be the superior choice in every scenario, including cases with minimal compatibility constraints. While the sequential approach is simpler from a modeling perspective, its lack of integrated decision-making across stages resulted in less efficient outcomes. In scenarios where flight selection is part of the decision-making process, the advantages of the integrated model diminished. Although the integrated model remains

the optimal choice in these cases, airlines might find the sequential approach more appealing due to its simplicity and only a modest impact on profitability.

To enhance the effectiveness and applicability of this model in real-world scenarios, several key recommendations are proposed. First, addressing symmetrical solutions is critical. Due to the identical nature of certain ULDs and the aggregation in the pricing problem, symmetrical solutions can arise, leading to redundant exploration of the solution space and slower convergence. This can be mitigated through strategies such as path prioritization, where specific paths are prioritized for identical ULDs, ensuring that only one representative solution is considered from each set of symmetrical solutions. Another approach is to impose lexicographic constraints, which enforce an arbitrary order on path assignments for identical ULDs. While this work incorporated steps to limit the solution space and the number of new columns, further exploration of advanced column management techniques, such as dynamic column generation, pooling, pruning, and adaptive selection, could help keep the RMP manageable while ensuring high-quality solutions.

From a practical standpoint, improving access to detailed, real data on cargo, including specific commodity types, will make the model more realistic and suitable for implementation. Additionally, a deeper understanding of the operating costs and availability of ULD types, especially for temperature-controlled units, would further improve the model's accuracy. The synthetic data currently used assumes an abundance of ULDs, including special units, which may not reflect actual availability. Ideally, real data from an airline case study would improve both these points. Additionally, incorporating ULD and cargo handling times, the ability to partially repack requests at intermediate airports, and accounting for ULD repositioning constraints (particularly for special ULDs) is critical for practical application. The current model would also benefit from exploring more accurate, ideally three-dimensional, bin-packing approaches to better capture real-world packing challenges. And finally, it is essential to consider the inherent uncertainty in cargo demand. Airlines often receive orders up until the last hour, leading to unpredictable fluctuations in actual cargo demand. Adapting the model to account for this stochastic demand will enhance its robustness and make it more applicable to real-world operations.

Appendix A Sets, parameters, and decision variables for the fully arc-based formulation.

Table A.13: Notation for the arc-based formulation.

| Sets | |
|---|---|
| \mathcal{N} | Set of nodes $n \in \mathcal{N}$ defining the TSN |
| \mathcal{G} | Set of ground arcs $g \in \mathcal{G}$ |
| \mathcal{F} | Set of flight arcs $f \in \mathcal{F}$ |
| \mathcal{E} | Set of arcs $a \in \mathcal{E}$ connecting source node s with the TSN |
| \mathcal{O} | Set of arcs $a \in \mathcal{O}$ connecting the TSN with sink node t |
| $\{a_{bp}\}$ | By-pass arc |
| \mathcal{A} | Set of all arcs $a \in \mathcal{A}$ defining the network |
| \mathcal{K} | Set of aircraft types $k \in \mathcal{K}$ |
| \mathcal{V} | Set of ULD types $v \in \mathcal{V}$ |
| \mathcal{U} | Set of ULDs $u \in \mathcal{U}$ |
| \mathcal{R} | Set of cargo requests $r \in \mathcal{R}$ |
| $\mathcal{A}^{\mathcal{K}} \subset \mathcal{A}$ | Subset of arcs a that aircraft can use |
| $\mathcal{A}^{\mathcal{U}} \subset \mathcal{A}$ | Subset of arcs a that ULDs can use |
| $\mathcal{A}_k^{\mathcal{K}} \subseteq \mathcal{A}^{\mathcal{K}}$ | Subset of arcs a that aircraft of type k can use |
| $\mathcal{A}_u^{\mathcal{U}} \subseteq \mathcal{A}^{\mathcal{U}}$ | Subset of arcs a that ULD u can use |
| $\mathcal{A}_{k,n^+}^{\mathcal{K}} \subset \mathcal{A}_k^{\mathcal{K}}$ | Subset of arcs a usable by aircraft of type k incident to node $n \in \mathcal{N}$ |
| $\mathcal{A}_{k,n^-}^{\mathcal{K}} \subset \mathcal{A}_k^{\mathcal{K}}$ | Subset of arcs a usable by aircraft of type k emanating from node $n \in \mathcal{N}$ |
| $\mathcal{A}_{u,n^+}^{\mathcal{U}} \subset \mathcal{A}_u^{\mathcal{U}}$ | Subset of arcs a usable by ULD u incident to node $n \in \mathcal{N}$ |
| $\mathcal{A}_{u,n^-}^{\mathcal{U}} \subset \mathcal{A}_u^{\mathcal{U}}$ | Subset of arcs a usable by ULD u emanating from node $n \in \mathcal{N}$ |
| $\mathcal{F}_{ru^+} \subset \mathcal{F}$ | Subset of flight arcs f such that $l_f^{o,\mathcal{F}} = l_r^{o,\mathcal{R}} = l_u^{o,\mathcal{U}}$, $DT_f \geq RT_r^{\mathcal{R}}$, and $DT_f \geq RT_u^{\mathcal{U}}$ |
| $\mathcal{F}_{ru^-} \subset \mathcal{F}$ | Subset of flight arcs f such that $l_f^{d,\mathcal{F}} = l_r^{d,\mathcal{R}} = l_u^{d,\mathcal{U}}$, $AT_f \leq DD_r^{\mathcal{R}}$, and $AT_f \leq DD_u^{\mathcal{U}}$ |
| $\mathcal{F}_k \subseteq \mathcal{F}$ | Subset of flight arcs f that aircraft type k can use |
| $\mathcal{F}_u \subseteq \mathcal{F}$ | Subset of flight arcs f that ULD u can use |
| $\mathcal{U}_r \subseteq \mathcal{U}$ | Subset of ULDs u where cargo request r can be packed |
| $\mathcal{U}_f \subseteq \mathcal{U}$ | Subset of ULDs u that can make use of flight arc f |
| $\mathcal{U}_v \subseteq \mathcal{U}$ | Subset of ULDs u that are of ULD type v |
| $\mathcal{R}_u \subseteq \mathcal{R}$ | Subset of requests r that can be transported with ULD u |
| $\mathcal{R}_{inc} \subseteq \mathcal{R} \times \mathcal{R}$ | Subset of incompatible item pairs (r_1, r_2) |
| $\mathcal{K}_f \subseteq \mathcal{K}$ | Subset of aircraft types k that are capable of operating flight arc f |
| Parameters | |
| C_{fk} | Operational fixed cost of operating flight f with aircraft type k |
| C_{fu} | Operational fixed cost of transporting ULD u on flight f |
| C_e | Fictitious cost used to encourage the model to use the by-pass arc a_{bp} |
| R_r | Revenue of cargo request r |
| W_{fk} | Maximum cargo payload transportable with aircraft type k on flight arc f |
| N_k | Number of aircraft of type k |
| N_{kv} | Maximum number of ULDs of type v for aircraft type k |
| $W_v^{\mathcal{V}}$ | Maximum transportable weight of ULDs of type v |
| $V_v^{\mathcal{V}}$ | Maximum transportable volume of ULDs of type v |
| $W_r^{\mathcal{R}}$ | Weight of cargo request r |
| $V_r^{\mathcal{R}}$ | Volume of cargo request r |
| Variables | |
| $x_{ak} \in \mathbb{N}_0$ | number of aircraft of type k using arc a |
| $y_{au} \in \{0, 1\}$ | unitary if ULD u uses arc a |
| $z_{ru} \in \{0, 1\}$ | unitary if cargo request r is assigned to ULD u |
| $w_{ruf} \in \{0, 1\}$ | unitary if cargo request r is assigned to ULD u and uses flight arc f |

Appendix B (Additional) sets, parameters and decision variables for the hybrid arc- and path-based formulation

Table B.14: Notation for the hybrid arc- and path-based formulation.

| Sets | |
|---|--|
| \mathcal{P} | Set of paths $p \in \mathcal{P}$ |
| $\mathcal{P}_r \subseteq \mathcal{P}$ | Subset of paths containing request r |
| $\mathcal{P}_u \subseteq \mathcal{P}$ | Subset of paths for ULD u |
| $\mathcal{P}_u^f \subseteq \mathcal{P}$ | Subset of paths for ULD u operating on flight arc f |
| $\mathcal{P}_u^r \subseteq \mathcal{P}$ | Subset of paths for ULD u containing request r |
| $\mathcal{U}_v^f \subseteq \mathcal{U}$ | Subset of ULDs u of ULD type v able to operate on flight arc f |
| Parameters | |
| R_{up} | Revenue of the requests carried by ULD u along path p |
| C_{up} | Operational costs of the transportation of ULD u along path p |
| W_{up} | Cumulative weight of the requests carried by ULD u along path p |
| Variables | |
| $x_{ak} \in \mathbb{N}_0$ | number of aircraft of type k using arc a |
| $z_{up} \in \{0, 1\}$ | unitary if ULD u is assigned to path p |

Appendix C Visualisation of the schedules in a Time-Space Network.

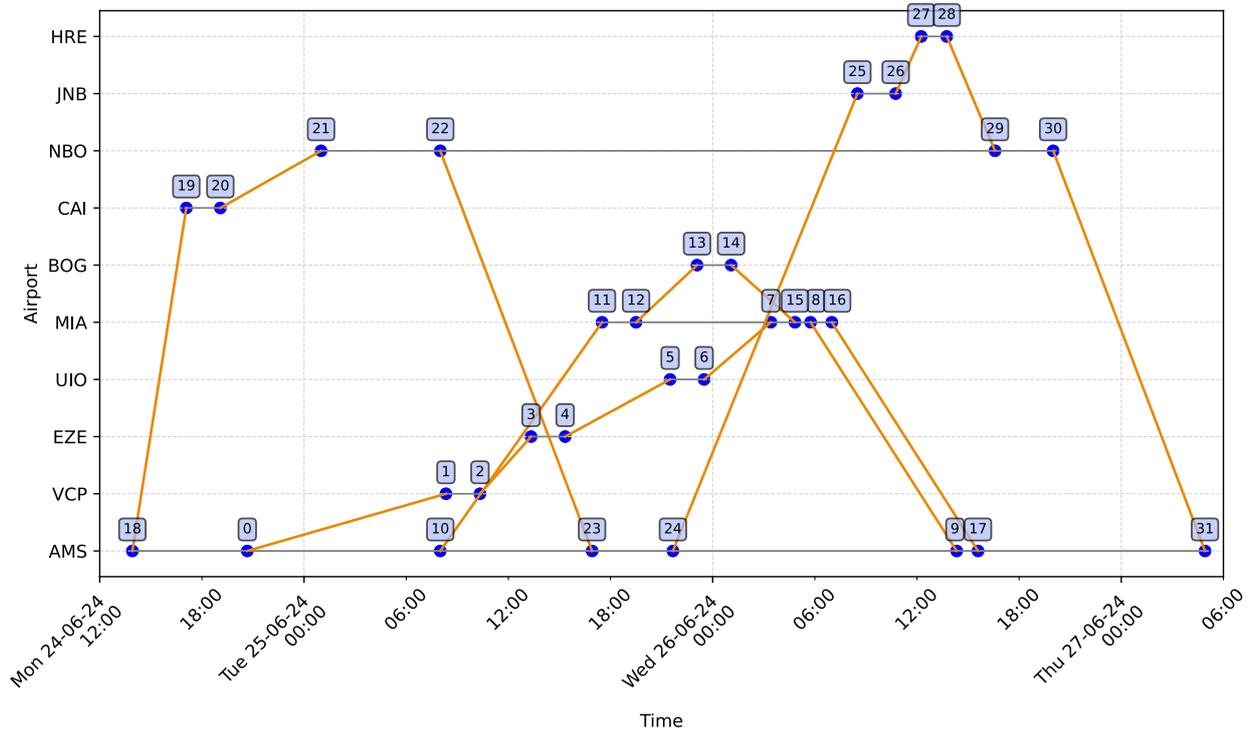


Figure C.4: Reduced schedule.

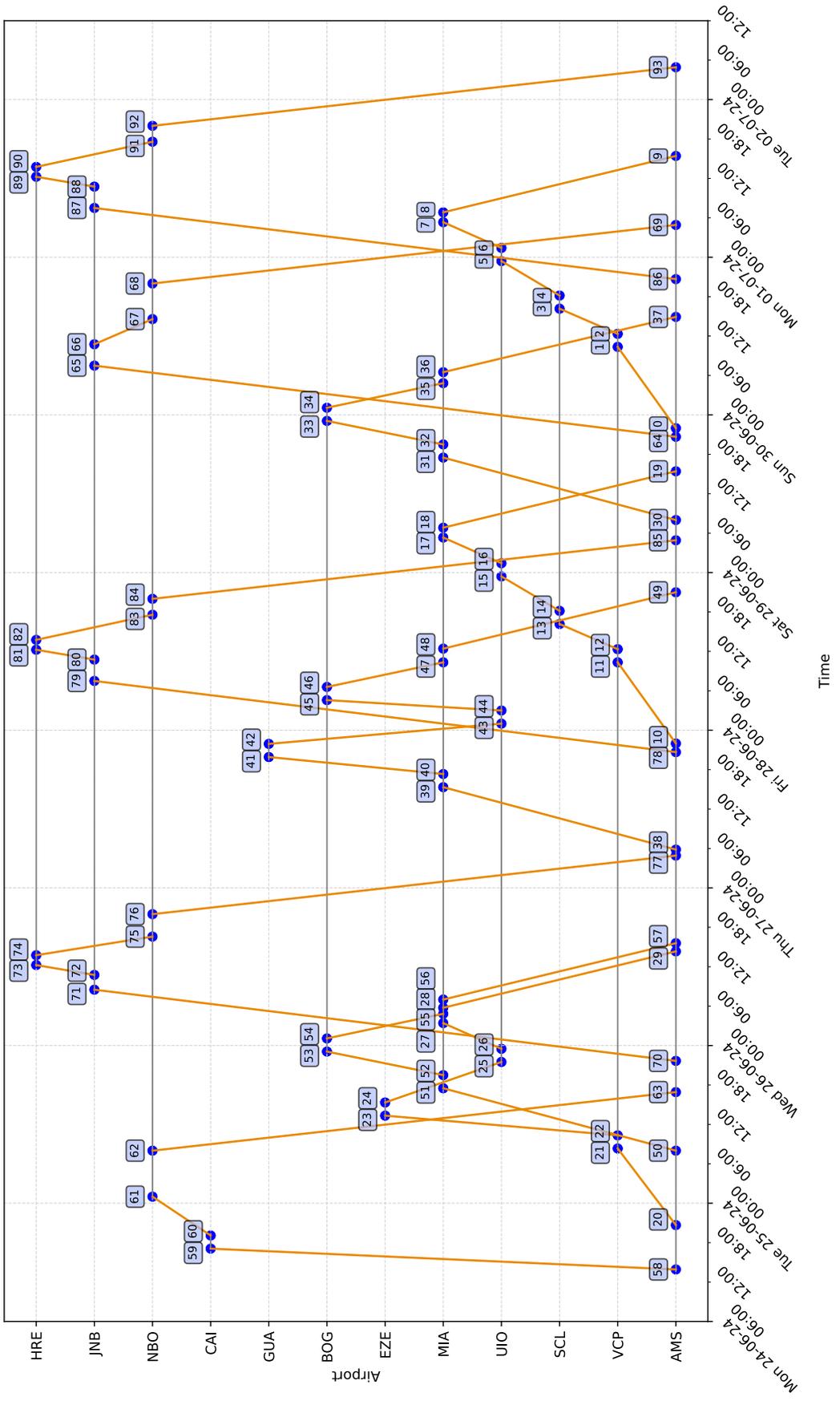


Figure C.5: Full schedule.

Appendix D Origin-Destination demand matrix formulation

Table D.15: Number of distinct paths matrix.

| To\From | AMS | MIA | BOG | UIO | SCL | VCP | EZE | GUA | CAI | JNB | NBO | HRE |
|---------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| AMS | 0 | 6 | 3 | 4 | 2 | 3 | 1 | 1 | 1 | 4 | 4 | 3 |
| MIA | 6 | 0 | 3 | 4 | 2 | 3 | 1 | 1 | 1 | 2 | 3 | 2 |
| BOG | 3 | 3 | 0 | 2 | 1 | 2 | 1 | 1 | 1 | 2 | 2 | 2 |
| UIO | 4 | 3 | 2 | 0 | 2 | 3 | 1 | 1 | 1 | 2 | 3 | 2 |
| SCL | 2 | 2 | 2 | 1 | 0 | 2 | 0 | 0 | 1 | 2 | 2 | 2 |
| VCP | 3 | 2 | 2 | 2 | 1 | 0 | 1 | 0 | 1 | 2 | 2 | 2 |
| EZE | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| GUA | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 1 | 1 |
| CAI | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| JNB | 4 | 3 | 3 | 3 | 1 | 2 | 1 | 1 | 1 | 0 | 3 | 2 |
| NBO | 5 | 3 | 3 | 3 | 1 | 2 | 1 | 1 | 1 | 4 | 0 | 3 |
| HRE | 3 | 2 | 2 | 2 | 1 | 2 | 1 | 1 | 1 | 3 | 2 | 0 |

Table D.16: OD pair rating matrix.

| To\From | AMS | MIA | BOG | UIO | SCL | VCP | EZE | GUA | CAI | JNB | NBO | HRE |
|---------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| AMS | 0 | A | C | C | C | A | C | C | A | A | A | A |
| MIA | A | 0 | A | B | A | A | B | A | E | E | E | E |
| BOG | C | B | 0 | C | C | C | C | C | E | E | E | E |
| UIO | D | E | D | 0 | C | C | C | C | E | E | E | E |
| SCL | D | B | D | D | 0 | C | 0 | 0 | E | E | E | E |
| VCP | A | D | D | D | D | 0 | D | 0 | E | E | E | E |
| EZE | B | 0 | 0 | 0 | 0 | A | 0 | 0 | 0 | 0 | 0 | 0 |
| GUA | E | E | C | D | 0 | C | C | 0 | E | E | E | E |
| CAI | AAA | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| JNB | B | E | E | E | E | E | E | E | D | 0 | C | C |
| NBO | B | E | E | E | E | E | E | E | B | B | 0 | B |
| HRE | C | E | E | E | E | E | E | E | D | C | D | 0 |

Table D.17: OD pair demand percentage.

| \From | AMS | MIA | BOG | UIO | SCL | VCP | EZE | GUA | CAI | JNB | NBO | HRE |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| To\ | [%] | [%] | [%] | [%] | [%] | [%] | [%] | [%] | [%] | [%] | [%] | [%] |
| AMS | 0.00 | 45.61 | 19.14 | 24.04 | 20.00 | 23.12 | 14.31 | 17.67 | 26.42 | 35.86 | 43.63 | 34.24 |
| MIA | 24.43 | 0.00 | 31.97 | 31.97 | 33.40 | 23.12 | 19.03 | 29.51 | 5.22 | 3.54 | 6.47 | 4.51 |
| BOG | 7.32 | 18.16 | 0.00 | 12.02 | 10.00 | 9.23 | 14.31 | 17.67 | 5.22 | 3.54 | 4.31 | 4.51 |
| UIO | 6.53 | 4.51 | 8.55 | 0.00 | 20.00 | 13.84 | 14.31 | 17.67 | 5.22 | 3.54 | 6.47 | 4.51 |
| SCL | 3.27 | 12.11 | 8.55 | 4.03 | 0.00 | 9.23 | 0.00 | 0.00 | 5.22 | 3.54 | 4.31 | 4.51 |
| VCP | 12.22 | 6.10 | 8.55 | 8.05 | 6.70 | 0.00 | 9.59 | 0.00 | 5.22 | 3.54 | 4.31 | 4.51 |
| EZE | 3.24 | 0.00 | 0.00 | 0.00 | 0.00 | 7.71 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| GUA | 0.80 | 1.50 | 6.38 | 4.03 | 0.00 | 4.61 | 14.31 | 0.00 | 5.22 | 1.77 | 2.16 | 2.26 |
| CAI | 5.68 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| JNB | 12.97 | 4.51 | 6.32 | 5.95 | 3.30 | 3.05 | 4.72 | 5.83 | 10.60 | 0.00 | 19.60 | 13.67 |
| NBO | 16.22 | 4.51 | 6.32 | 5.95 | 3.30 | 3.05 | 4.72 | 5.83 | 21.04 | 28.56 | 0.00 | 27.27 |
| HRE | 7.32 | 3.00 | 4.21 | 3.97 | 3.30 | 3.05 | 4.72 | 5.83 | 10.60 | 16.10 | 8.75 | 0.00 |

References

- Baxter, G., Kourousis, K., 2015. Temperature Controlled Aircraft Unit Load Devices: The Technological Response to Growing Global Air Cargo Cool Chain Requirements. *Journal of Technology Management & Innovation* 10, 157–172. doi:<http://dx.doi.org/10.4067/S0718-27242015000100012>.
- van Bockstaele, V., Buyle, S., Dewulf, W., 2023. Solving the mystery of discrepancies and double counting in air cargo through demand and supply big data analysis. *Journal of the Air Transport Reserach Society* 1, 81–100. doi:<https://dx.doi.org/10.59521/6A961EF46EB809C5>.

- Boeing, 2008. 747-400 boeing converted freighter. https://www.boeing.com/content/dam/boeing/boeingdotcom/company/about_bca/startup/pdf/freighters/747BCF.pdf. Accessed: August 2024.
- Boeing, 2010. 747-400/-400er freighters. https://www.boeing.com/content/dam/boeing/boeingdotcom/company/about_bca/startup/pdf/freighters/747-400f.pdf. Accessed: August 2024.
- Boeing, 2022. Boeing website: World Air Cargo Forecast 2022-2041. https://www.boeing.com/content/dam/boeing/boeingdotcom/market/assets/downloads/Boeing_World_Air_Cargo_Forecast_2022.pdf. Accessed: August 2024.
- Brandt, F., Nickel, S., 2019. The air cargo load planning problem - a consolidated problem definition and literature review on related problems. *European Journal of Operational Research* 275, 399–410. doi:<https://doi.org/10.1016/j.ejor.2018.07.013>.
- ColdChain, 2021. ColdChain website: Cold chain pallet shippers. <https://www.coldchaintech.com/product-details/traditional-pallet/>. Accessed: March 2023.
- Delgado, F., Mora, J., 2021. A matheuristic approach to the air-cargo recovery problem under demand disruption. *Journal of Air Transport Management* 90, 101939. doi:<https://doi.org/10.1016/j.jairtraman.2020.101939>.
- Delgado, F., Sirhan, C., Katscher, M., Larrain, H., 2020. Recovering from demand disruptions on an air cargo network. *Journal of Air Transport Management* 85, 101799. doi:<https://doi.org/10.1016/j.jairtraman.2020.101799>.
- Derigs, U., Friederichs, S., 2013. Air cargo scheduling: integrated models and solution procedures. *OR spectrum* 35, 325–362. doi:<https://doi.org/10.1007/s00291-012-0299-y>.
- Derigs, U., Friederichs, S., Schäfer, S., 2009. A new approach for air cargo network planning. *Transportation Science* 43, 370–380. doi:<https://doi.org/10.1287/trsc.1090.0282>.
- Desaulniers, G., Desrosiers, J., Solomon, M.M., 2002. Accelerating Strategies in Column Generation Methods for Vehicle Routing and Crew Scheduling Problems. Springer US, Boston, MA. pp. 309–324. URL: https://doi.org/10.1007/978-1-4615-1507-4_14, doi:10.1007/978-1-4615-1507-4_14.
- Feng, B., Li, Y., Shen, Z.J.M., 2015. Air cargo operations: Literature review and comparison with practices. *Transportation Research Part C: Emerging Technologies* 56, 263–280. doi:<https://doi.org/10.1016/j.trc.2015.03.028>.
- Gueret, C., Jussien, N., Lhomme, O., Pavageau, C., Prins, C., 2003. Loading aircraft for military operations. *Journal of the Operational Research Society* 54, 458–465. doi:<https://doi.org/10.1057/palgrave.jors.2601551>.
- Heidelberg, K.R., Parnell, G.S., Ames IV, J.E., 1998. Automated air load planning. *Naval Research Logistics* 45, 751–768. doi:[https://doi.org/10.1002/\(SICI\)1520-6750\(199812\)45:8<751::AID-NAV1>3.0.CO;2-R](https://doi.org/10.1002/(SICI)1520-6750(199812)45:8<751::AID-NAV1>3.0.CO;2-R).
- IATA, 2021. IATA website: How to Ship Perishable Goods by Air? <https://www.iata.org/en/publications/newsletters/iata-knowledge-hub/how-to-ship-perishable-goods-by-air/>. Accessed: August 2023.
- IATA, 2024. Air Cargo Market Analysis august 2024. <https://www.iata.org/en/iata-repository/publications/economic-reports/air-cargo-market-analysis-august-2024/>. Accessed: October 2024.
- Li, D., Huang, H.C., Chew, E.P., Morton, A., 2007. Simultaneous fleet assignment and cargo routing using benders decomposition. *Container Terminals and Cargo Systems: Design, Operations Management, and Logistics Control Issues* , 315–331.
- Lufthansa, 2024. Lufthansa cargo website: safe transport solutions for the chemical industry. <https://lufthansa-cargo.com/industries-chemicals>. Accessed: August 2024.
- Mesquita, A.C., Sanches, C.A., 2024. Air cargo load and route planning in pickup and delivery operations. *Expert Systems with Applications* 249. doi:10.1016/j.eswa.2024.123711.

- van der Meulen, S., Grijspaardt, T., Mars, W., van der Geest, W., Roest-Crollius, A., Kiel, J., 2023. Cost Figures for Freight Transport - final report (Kostenkengetallen voor Goederenvervoer - eindrapportage). Technical Report. Panteia. Zoetermeer. Commissioned by the Netherlands Institute for Transport Policy Analysis (KiM), part of the Dutch Ministry of Infrastructure and Water Management (IenW).
- Mongeau, M., Bes, C., 2003. Optimization of aircraft container loading. *IEEE Transactions on aerospace and electronic systems* 39, 140–150. doi:[10.1109/TAES.2003.1188899](https://doi.org/10.1109/TAES.2003.1188899).
- Nance, R., Roesener, A., Moore, J., 2011. An advanced tabu search for solving the mixed payload airlift loading problem. *Journal of the Operational Research Society* 62, 337–347. doi:<https://doi.org/10.1057/jors.2010.119>.
- Paquay, C., Limbourg, S., Schyns, M., 2018. A tailored two-phase constructive heuristic for the three-dimensional Multiple Bin Size Bin Packing Problem with transportation constraints. *European Journal of Operational Research* 267, 52–64. doi:<https://doi.org/10.1016/j.ejor.2017.11.010>.
- Paquay, C., Schyns, M., Limbourg, S., 2016. A mixed integer programming formulation for the three-dimensional bin packing problem deriving from an air cargo application. *International Transactions in Operational Research* 23, 187–213. doi:<https://doi.org/10.1111/itor.12111>.
- Vancroonenburg, W., Verstichel, J., Tavernier, K., Berghe, G.V., 2014. Automatic air cargo selection and weight balancing: a mixed integer programming approach. *Transportation Research Part E: Logistics and Transportation Review* 65, 70–83.
- Xiao, F., Guo, S., Huang, L., Huang, L., Liang, Z., 2022. Integrated aircraft tail assignment and cargo routing problem with through cargo consideration. *Transportation Research Part B: Methodological* 162, 328–351.
- Xu, Y., Wandelt, S., Sun, X., 2024. Airline scheduling optimization: literature review and a discussion of modelling methodologies. *Intelligent Transportation Infrastructure* 3, 1–24. URL: <https://doi.org/10.1093/itinfr/liad026>, doi:[10.1093/itinfr/liad026](https://doi.org/10.1093/itinfr/liad026).
- Yan, S., Lo, C.T., Shih, Y.L., 2006. Cargo container loading plan model and solution method for international air express carriers. *Transportation Planning and Technology* 29, 445–470. doi:<https://doi.org/10.1080/03081060601075674>.
- Zhao, X., Yuan, Y., Dong, Y., Zhao, R., 2021. Optimization approach to the aircraft weight and balance problem with the centre of gravity envelope constraints. *IET Intelligent Transport Systems* 15, 1269–1286. doi:[10.1049/itr2.12096](https://doi.org/10.1049/itr2.12096).
- Zheng, H., Sun, H., Zhu, S., Kang, L., Wu, J., 2023. Air cargo network planning and scheduling problem with minimum stay time: A matrix-based alns heuristic. *Transportation Research Part C: Emerging Technologies* 156. doi:[10.1016/j.trc.2023.104307](https://doi.org/10.1016/j.trc.2023.104307).
- Zhou, L., Liang, Z., Chou, C.A., Chaovalitwongse, W., 2020. Airline planning and scheduling: Models and solution methodologies. *Frontiers of Engineering Management* 7. doi:[10.1007/s42524-020-0093-5](https://doi.org/10.1007/s42524-020-0093-5).