2019/09

DP

José Miguel Quesada Pérez, Jean-Sébastien Tancrez and Jean-Charles Lange

Multi-hub express shipment service network design with complex routes

CORE
CENTER FOR OPERATIONS RESEARCH AND ECONOMETRICS

CORE
Voie du Roman Pays 34, L1.03.01
B-1348 Louvain-la-Neuve
Tel (32 10) 47 43 04
Email: immaq-library@uclouvain.be
https://uclouvain.be/en/research-institutes/lidam/core/discussion-papers.html

Multi-Hub Express Shipment Service Network Design with Complex Routes

José Miguel Quesada Pérez[a,*], Jean-Sébastien Tancrez[a], Jean-Charles Lange

[a]*CORE - Center for Operations Research and Econometrics, Université catholique de Louvain, Chaussée de Binche, 151, 7000 Mons, Belgium.*

Abstract

The *Express Shipment Service Network Design* (ESSND) problem consists in defining a network of flights that enables the overnight flow of express packages from their origins to their destinations at minimum cost. This problem is normally solved considering only one-leg, multi-leg and ferry routes. Assessing the value of more complex route types is an open question of academic and practical importance. In this article, we present a mixed integer programming model that includes five types of complex routes: two-hub, transload, direct, inter-hub and early routes. We assess their economic impact by performing many experiments built from an instance provided by FedEx Express Europe. Inter-hub and early routes have the best performance, with significant average savings (from 0.5% to 3.5%).

Keywords: Service network design, express integrator, multiple hubs, flexible hub assignment, mixed integer programming, complex routes.

1. Introduction

Express integrators provide fast and reliable door-to-door delivery services worldwide. Their service proposition has appointed them as key supporters of the global market. According to Boeing Commercial Airplanes (2018), between 2010 and 2017, the international express air cargo market grew at an average rate of 7.7% per year and, in 2017, the air express operations generated 43% of the world air cargo revenue. To increase their competitiveness, the express carriers are highly dependent on improving the efficiency of their operations. The main service of the express carriers is the overnight delivery of packages within regions as large as the US

*Corresponding author

Email addresses: jose.quesada@uclouvain.be (José Miguel Quesada Pérez), js.tancrez@uclouvain.be (Jean-Sébastien Tancrez)

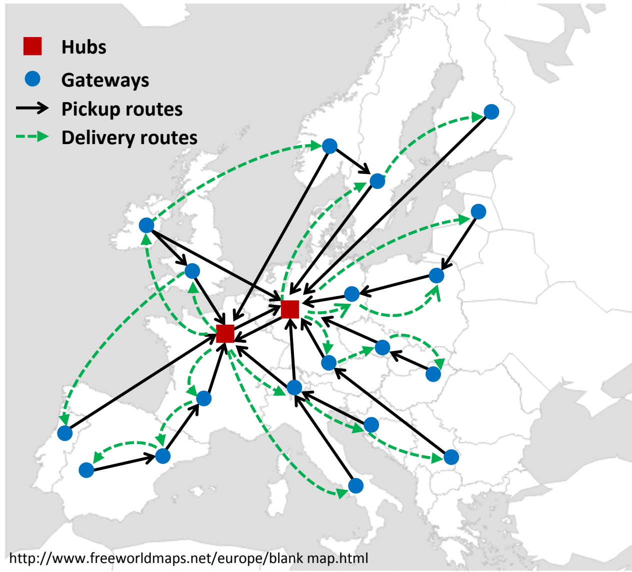

Figure 1: Simplified representation of the air operations for the pickup and delivery processes of express integrators.

or Europe. For these overnight deliveries, the packages collected from the origin customers are gradually consolidated, first at city *stations*, then at airport facilities called *gateways*, and finally at airport facilities called *hubs*. At the hubs, the packages are sorted by destination, and then are gradually deconsolidated, by hauling them from hubs to gateways, then to city stations, and finally to the destination customers. Moving the packages from the origin customers to hubs is known as the *pickup process*, and from hubs to the destination customers as the *delivery process*.

The aircraft operations, from origin gateways to hubs to destination gateways, constitute a large component of the total operational cost of the overnight delivery process. In this article, we focus on studying the network design of the air operations of express integrators (see Figure 1), a task known as the express shipment service network design (ESSND) problem, which was introduced by Kuby and Gray (1993) and fully characterized by Kim et al. (1999). To solve the ESSND problem, potential routes, that respect the operative constraints, are first predetermined, and an optimization model then selects the routes that most efficiently deliver the packages. A route is a sequence of coordinated flights performed with an aircraft (two aircrafts for transload routes), with the purpose of either moving packages between nodes or to reposition aircrafts. Each route can fly one leg (i.e. the movement of an aircraft between two airports) or more. Due to the complexity on solving the ESSND problem on real size instances, researchers and practitioners have paid attention to consider enough routes linking hubs and

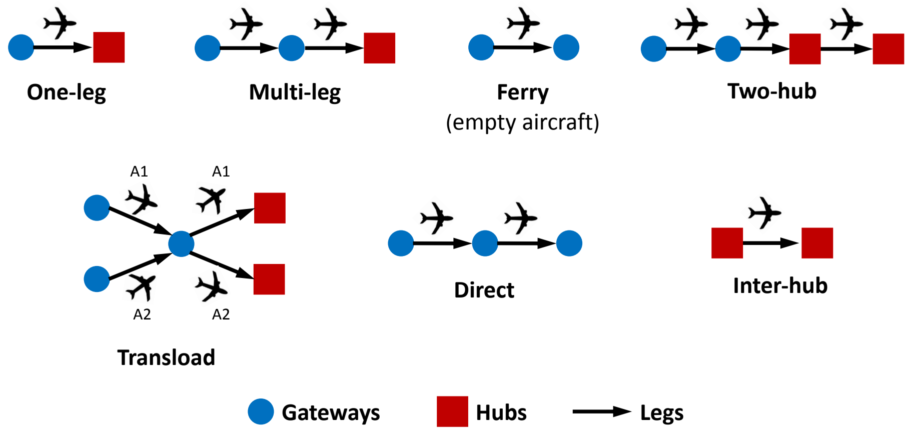

Figure 2: Examples of the route types (pickup routes for one-leg, multi-leg, two-hub and transload routes).

gateways to obtain useful solutions, but not too many so as to keep models tractable. Most of the ESSND models in the literature include three route types that we call *standard*: one-leg, multi-leg and ferry routes. In this article, our goal is to assess the contribution of five route types that we call *complex*, and that have not been or are rarely explored in the literature (see Section 2): two-hub, transload, direct, inter-hub and early routes. The complex routes are particularly relevant when designing multi-hub networks. Figure 2 shows a representation of these standard and complex routes. They are described in details in Section 3.2. For both academics and practitioners, it is still an open question to determine what is the potential benefit of complex routes and when they are effective despite the additional computational cost they imply. The five types of complex routes we address are the first in line when thinking about promising routes to add when solving the ESSND problem. The two-hub, transload and inter-hub routes could bring benefits by reducing the number of aircrafts required to connect groups of gateways with groups of hubs, and the direct and early routes could bring savings by decongesting the sorting systems at hubs. In this article, we show how these complex routes can be added to the strategic ESSND problem and we show that the resulting model can be solved efficiently, thereby making the assessment of their relevance and economic impact possible.

For our analysis, we develop an exact mixed integer programming model for the strategic ESSND problem with flexible hub assignment (i.e. the assignment of packages to hubs is part of the decisions of the model), in which complex routes are not enforced in the solution nor their loads predefined. The economic contribution of the complex routes is then analyzed with many numerical experiments, varying several input parameters to understand their influence (demand volumes, hub sorting capacities and distance between hubs). The experimental instances were built based on a fictitious instance provided by FedEx Express Europe, built to be fully representative of real industrial instances (2 hubs, 77 gateways and 7 aircraft types). The objective of the collaboration with FedEx Express Europe is the development of optimization models highly grounded in the industrial reality.

The remainder of the article is structured as follows. The related literature is reviewed in Section 2. Section 3 presents in details the express shipment service network design problem, its restrictions and the eight route types considered in this article. In Section 4, we introduce our Route and Hub model with Cuts and Complex Routes for the multi-hub ESSND problem. Finally, in Section 5, we discuss the results of our numerical experiments.

2. Literature review

The multi-commodity network design problem consists in determining an efficient infrastructure for transporting goods or information between multiple origins and destinations (Gendron et al., 1999). Therefore, it has applications in many contexts, e.g., telecommunications, air and ground transportation of freight or passengers. Surveys and reviews of the multi-commodity network design problem can be found, among others, in Gendron et al. (1999), Agarwal (2002) and Tong et al. (2015). The conventional formulation of the network design problem provides poor linear programming relaxations, and is difficult to solve for large instances (Armacost et al., 2002). This difficulty is increased when additional characteristics are included to the standard problem, e.g., by having unsplittable commodities (Benhamiche et al., 2016), by considering crew and maintenance decisions (Crainic, 2000) or by considering the network survivability (Gouveia et al., 2018). Reviews on the modeling variants and solution approaches developed to overcome such difficulties are presented by Crainic (2000), Wieberneit (2008), Wang et al. (2013) and Meuffels (2015).

The *Express Shipment Service Network Design* (ESSND) problem is a subclass of the multi-commodity network design problem, faced by carriers that offer door-to-door delivery of goods with tight due dates within regions as large as the US or Europe. This problem has received attention from a limited number of researchers. Most of the contributions come from collaborations with express companies and have reported significant savings on real cases (e.g., Barnhart and Schneur, 1996; Armacost et al., 2004; Fleuren et al., 2013; Louwerse et al., 2014). Table

Table 1: Model characteristics and route types from works in the literature of the ESSND problem.

	Route types							**Hub modeling**		
	Standard routes		**Complex routes**					**Hub assignment**		
	One-leg and multi-leg	Ferry	Two-hub	Trans-load	Direct	Inter-hub	Early	Fixed	Flexible	Multi-hub
Kuby and Gray (1993)	X							X		
Barnhart and Schneur (1996)	X	X			X$_1$			X		
Kim et al. (1999)	X	X		X	X$_2$				X	X
Armacost et al. (2002)	X	X		X$_3$				X		X
Shen (2004)	X	X							X	X
Fleuren et al. (2013)	X	X			X$_2$			X		
Louwerse et al. (2014)	X	X			X$_2$	X$_4$		X		X
Meuffels (2015)	X	X				X$_4$		X		X
Quesada Pérez et al. (2018)	X	X							X	X

X$_1$: Routes operated by a third party with a weight-based cost.
X$_2$: Direct routes are enforced to fly with predefined flows as input.
X$_3$: The transloaded flows assigned to each aircraft of the route are predefined.
X$_4$: Each gateway is served by a single preassigned hub, and thus, the inter-hub connections are implied.

1 reports on the main contributions in solving the ESSND problem. As our article aims at studying the impact of complex routes when they are considered in the ESSND problem, Table 1 details when these routes have been considered in the literature. As can be seen, they are scarcely studied. This table also shows if the articles assume a fixed hub assignment (the hub assignment is an input of the model) or a flexible hub assignment (the hub assignment is a decision of the model), and if they address single hub or multi-hub instances.

The study of the ESSND problem starts with Kuby and Gray (1993), who develop a single-hub model without sorting, counting, circulation nor slot constraints. They compare the cost-effectiveness of networks designed with only one-leg routes against networks that also include multi-leg routes. Based on the FedEx case in the west coast of the US, they show that the multi-leg routes allow much more efficient network designs (73.6% less expensive). Barnhart and Schneur (1996) develop a single-hub model in which they incorporate the sorting, counting and slot constraints, but in which each gateway is served by a single aircraft type. They include direct routes in their formulation as commercial flights operated by a third party, with a weight-based cost. Among the researchers that include direct routes, Barnhart and Schneur (1996) are the only ones that do not enforce them in the solution. However, they do not make an analysis of their benefit. They solved a large instance from an express carrier in the US with a column generation technique, uncovering potential savings in the order of tens of millions of dollars per year. Kim et al. (1999) present the route and flow model, a classical multi-commodity network design approach applied to the multi-hub ESSND problem with flexible hub assignment. In their implementation, Kim et al. (1999) include direct routes that are enforced in the solution. They also add transload routes assuming that the packages can be transferred between aircrafts at any gateway, but at most once. In their approach, the transload routes are not enforced nor their loads predefined, but their economic impact is not evaluated. For solving the problem, the authors propose a heuristic method with generation of columns and rows. They apply it to three data sets provided by an express integrator and compare their results to the solutions provided by the company, obtaining average savings of 10%. Armacost et al. (2002) introduce the aircraft routing model for solving the multi-hub ESSND problem with fixed hub assignment. They include transload routes in their formulation, but they need to determine their loads during the route generation process. Armacost et al. (2002) apply their formulation to the UPS case in the US, which led to savings of tens of millions of dollars per year. They also test the benefit of transload routes, finding a cost improvement of 1.67%.

The second model addressing the multi-hub ESSND problem with flexible hub assignment is the route and hub model, introduced by Shen (2004). This model is a reformulation of the route and flow model by Kim et al. (1999) with a reduced number of variables and constraints. Shen (2004) does not study any complex route. Applying his approach to the UPS case in the US and by fixing 60% of the hub assignments, he reports savings of 7.5% compared to the solution obtained by UPS. Fleuren et al. (2013) apply the aircraft routing model by Armacost et al. (2002) for a single hub case, allowing the user to enforce some direct routes for coordinating the international and domestic networks of TNT, which are designed independently. Using this approach, Fleuren et al. (2013) report savings of 132 million dollars between 2008 and 2009. Louwerse et al. (2014) use the aircraft routing model by Armacost et al. (2002) in the development of a heuristic algorithm for solving air or ground ESSND problems with multiple service types. They enforce some direct routes in the solution and also some inter-hub routes by assuming that each gateway is served by a single hub. The authors test modified instances from an express carrier, getting up to 18.6% cost savings compared to the solutions provided by the carrier. Meuffels (2015) develops a heuristic algorithm for the hub decongestion of a dual-hub case for an express integrator in Europe. She includes inter-hub routes with the same approach as Louwerse et al. (2014). Applying her method, Meuffels (2015) reports bypassing 16% of the network demand from the sorting process. Finally, Quesada Pérez et al. (2018) present the route and hub model with cuts and covers, an evolution of the route and hub model by

Shen (2004), enhanced with three families of valid inequalities (commodity connectivity, strong linking and strong cover inequalities) and with covers to reduce the number of flow variables and capacity constraints. On realistic instances from an express carrier in Europe, their model outperforms the route and hub model by more than 20% in the optimality gap and cost. In the present article, we extend this model by adding five types of complex routes and the hub sorting staggering constraints.

As it can be seen, the complex routes are scarcely addressed in the literature and their benefit is usually not analyzed. Moreover, for tractability reasons, most of the contributions enforce the complex routes in the solution and/or predefine their loads. The main contributions of our article are the following. First, we introduce two-hub routes and early routes, which are new to the literature. Second, we present the Route and Hub model with Cuts and Complex Routes (RHCCR), a strategic multi-hub ESSND model with flexible hub assignment that incorporates five types of complex routes: transload, direct, inter-hub, two-hub and early routes. In this model, we do not enforce the complex routes in the solution nor predefine their loads. Third, the benefits of these complex routes are analyzed with many numerical experiments performed with realistic instances.

3. The multi-hub ESSND problem with complex routes

The express shipment service network design problem aims at selecting a set of routes that enables the transportation of all packages from their origins to their destinations, overnight and at minimal cost, while respecting the operative conditions of the gateways, hubs and aircrafts. To reach its destination, a package is loaded into a single pickup route at its origin, which hauls the package to a hub. There, the package is sorted and then loaded into a single delivery route, which hauls the package to its destination. In our setting, we consider a network with multiple hubs. A group of packages sharing the same origin and destination is called a *commodity*, and it can be split and hauled through different routes and hubs to achieve its

To solve the ESSND problem, most academics as well as practitioners follow a two-stage process. First, a set of feasible routes is computed by enumeration, a process called route generation. Second, the routes generated become the input of a mathematical model which is solved so as to design a network that ensures the overnight delivery of packages at minimal cost. Generating the routes prior to solving the model reduces the complexity for modeling the

ESSND problem, as during this process, the schedules and costs of the routes are calculated and fixed, and thus do not need to be included explicitly in the model of the second stage. The route generation process enumerates all of the feasible routes that the express integrator may perform for each possible combination of aircraft types, hubs, and gateways, considering their operative constraints (see Section 3.2). We refer to Kim et al. (1999) for a detailed mathematical description of the route generation process. The cost of a route includes the variable, cycle and ownership costs. The variable costs are incurred for each hour flown. The cycle costs are related to each take-off and landing cycle. The ownership costs are the daily depreciation or leasing costs of the aircraft. Note that the package handling costs are mainly determined by the installation and operation costs of the hubs and gateways, which for this problem are both assumed to be fixed and are thus omitted.

3.1. Problem restrictions

In this section, we detail the problem restrictions, particularly coming from the operative constraints of the facilities and aircrafts. The restrictions related to the routes are presented in the next section, 3.2. The operative constraints of the facilities are the following. The gateways/hubs have a *release time*, i.e. the moment at which they release packages to the pickup/delivery process. They also have a *due time*, i.e. the latest moment at which they receive packages from the delivery/pickup process and ensures the on time delivery. Additionally, hubs also perform the sorting process with automated conveyor systems that have a sorting capacity per unit of time. This capacity cannot be exceeded due to the arrival of packages from the pickup process, given that each package must be sorted before the hub's release time. The hubs and gateways may also establish curfews banning the landings and departures of aircrafts. A final limitation of hubs is the *slots constraint*, i.e. the maximum number of take-offs and landings that a company can perform at each hub during the night.

The aircrafts have to respect the following operative constraints. First, aircrafts have a limited capacity in terms of the weight they can carry. Second, they have a maximum flying range (without fuel refill). Third, they have traveling speeds that determine if they can respect the release and due times of the visited airports. Additionally, each time an aircraft lands at an airport, it must remain there for a minimum *stop time* before departing, which allows the loading/unloading of packages and refueling. Finally, there is a limit to the total number of aircrafts of each type available in the network.

Beside the constraints related to the facilities and aircrafts, two other restrictions are important. First, the routes schedule is designed for one day of operations, and is repeated daily for months. To ensure this repeatability, referred to as the *circulation problem*, each hub and gateway must start and end daily operations with the same number of aircrafts of each type. Second, a constraint imposed by some express integrators is referred to as the *main hub connectivity condition* (Barnhart et al., 2002; Kim et al., 1999; Shen, 2004), which forces each gateway to be connected by at least one route to the main hub (the one with the largest sorting capacity) for both the pickup and the delivery process. Being a restriction imposed for business reasons, the main hub connectivity condition is not necessary and may be omitted in the ESSND models (Quesada Pérez et al. (2018) compare results with or without this condition).

3.2. Detailed routes descriptions

In this section, we describe in detail the eight route types that are considered in this article. They are illustrated in Figure 2, and their goal and operations are detailed in the following list. We call the last five in this list complex, as they are scarcely explored in the literature (see Section 2). For the route types with pickup or delivery variants (one-leg, multi-leg, two-hub, transload and early routes), we only explain the pickup variants, since their delivery counterparts follow the same logic but in inverse direction.

- *One-leg routes* are flown by an aircraft hauling packages directly from a gateway to a hub.
- *Multi-leg routes* are flown by an aircraft picking packages sequentially from two or three gateways and delivering them to a hub.
- *Ferry routes* are flown by an empty aircraft between two airports, after the end of the delivery process. Their goal is to reposition aircrafts so that they are able to perform the pickup routes they have to fly the following day (i.e. to comply with the circulation problem, see Section 4.1). An aircraft performing a ferry route previously performed either a pickup and a delivery route, or a direct route.
- *Two-hub routes* are extensions of one-leg or multi-leg routes that connect to two hubs. They move packages to a hub where they release a fraction of their loads and then, without picking

additional packages, they fly to a second hub for delivering the remaining packages. Unlike inter-hub routes, these routes do not load packages at the first hub. Otherwise, their stop time at that hub would be increased, thus significantly reducing the number of two-hub routes that can be operated. Two-hub routes require the pickup gateways to perform a pre-sort of packages per destination hub, which is needed for all gateways in a multi-hub environment. The use of a two-hub route reduces the number of aircrafts required to connect a group of gateways to two hubs.

- *Transload routes* involve two aircrafts departing from two different gateways and meeting at a third one. There, they exchange packages among them and load packages from the third gateway. Finally, each aircraft flies to a different hub. Besides the pre-sorting ability of the pickup gateways, the transload operation implied at the third gateway requires more time than the regular stop time. The use of a transload route reduces the number of aircrafts required to connect three gateways to two hubs. The transload routes are the only routes that involve two aircrafts and that allow transfers of packages among aircrafts at gateways.

- *Direct routes* connect two or more gateways without touching any hub, picking and/or delivering packages at each stop. These routes reduce the quantity of packages to be sorted at hubs. An aircraft performing a direct route can only perform additionally a ferry route within the same day (while other aircrafts perform at least a pickup route and a delivery route, and may perform additionally an inter-hub and/or a ferry route). Moreover, a gateway served by a direct route must be able to pre-sort packages that are loaded for the gateways served by the route afterwards. Figure 3 shows the five variants of direct routes produced during our route generation process.

- *Inter-hub routes* are flown by an aircraft between two hubs. They consolidate packages from the pickup process at one hub and move them to another hub where the packages are sorted and released to the delivery process. They allow to reduce the number of direct connections between gateways and hubs. In the following, we mention several features and conditions of the inter-hub routes. First, as for previous routes, packages need to be pre-sorted per hub at the origin gateways, so that the transfer into inter-hub routes can be performed directly. Second, an aircraft performing an inter-hub route also performs one pickup route and one delivery route, and the pickup route thus has to complete its operations

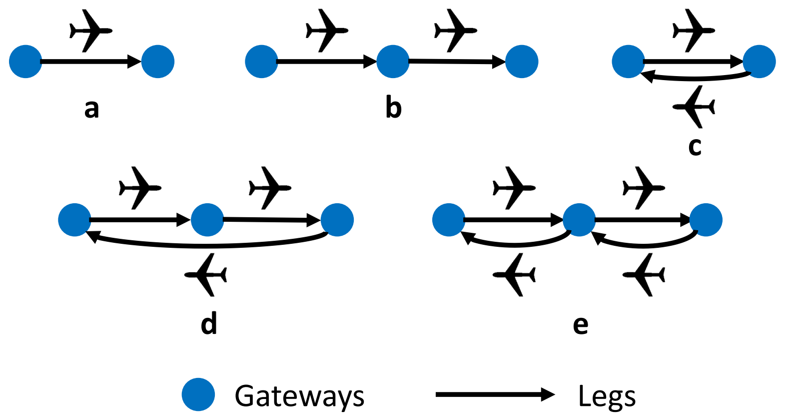

Figure 3: Direct route variants.

before the departure of the inter-hub route. Third, pickup routes are allowed to transfer packages into the inter-hub routes if they arrive at the origin hub a given time before the departure of the inter-hub route. Fourth, an inter-hub route may be allowed to arrive after the due time of its destination hub, as long as its load can be sorted before the destination hub release time. This allows the inter-hub routes to pick packages from more gateways, and thus increases their ability to reduce the number of direct gateway-hub connections. Fifth, transload and two-hub routes cannot transfer packages to inter-hub routes. Although this assumption reduces the number of potential gateway-hub connections due to inter-hub routes (for networks with three or more hubs, as two-hub and transload routes already connect to two hubs), it simplifies the formulation and computational tractability of these routes. Although both the inter-hub and two-hub routes visit two hubs, they are different as inter-hub routes consolidate packages from several gateways without visiting them, while two-hub routes only haul packages from the gateways they visit.

- *Early routes* form a special type of one-leg and multi-leg routes that are allowed to depart at an *early release time*, i.e. a given time earlier than the regular gateway/hub release time (e.g. 30 minutes earlier), carrying a fraction of a gateway's demand (called *early demand)*. In practice, this is often a reasonable assumption as the release times in the ESSND problem represent the end of granular processes. For pickup gateways, the arrival of packages from stations is staggered (as explained in detail by Schenk and Klabjan (2010), who study the transportation of packages between stations and gateways). For hubs, the sorting operations release packages gradually to the delivery process. Considering early routes means cutting these granular processes, leading to two release times instead of a single one, assuming that a fraction of the total demand is available a certain amount of time before the regular release time. Early routes may enable gateway-hub connections with slower aircraft types that cannot be achieved otherwise. Also, if they advance the arrivals of some packages to hubs, the sorting load of hubs at peak hours may be reduced. When using early routes, the early release times are a key choice. Choosing an earlier time will allow to connect earlier to the hub and to decongest the sorting operations (or use a slower aircraft), but will reduce the amount of packages that already reached the gateway (the early demand) and that can thus be hauled by the early route. The calculation of the correspondent early demands is based on historic data, and is different for each gateway. Importantly, to better evaluate the contribution of early routes, the early demands are allowed but not forced to depart early. They may be carried by other route types at the regular release times.

4. The ESSND model with multiple hubs and complex routes

Now that the problem has been presented in details, we propose the *Route and Hub model with Cuts and Complex Routes* (RHCCR), a mixed integer program for solving the multi-hub ESSND problem when complex routes are considered.

4.1. Route and Hub model with Cuts

In this section, we give the Route and Hub model with Cuts (RHC), introduced by Quesada Pérez et al. (2018). It serves as a basis for our RHCCR model, which adds complex routes and time-dependent hub sorting capacity constraints. Before proceeding, we introduce the notation of the RHC model.

Sets and indices

F	is the set of aircraft types, indexed by f.
R	is the set of routes, indexed by r or r'.
R_P^f	is the set of pickup routes performed with aircraft type f.
H	is the set of hubs, indexed by h or h'.
N_P	is the set of pickup gateways.
N_D	is the set of delivery gateways, where $N_P \cap N_D = \emptyset$.
N	is the set of gateways, indexed by n, where $N = N_P \cup N_D$.
N_r	is the set of gateways served by route r.
G	is the set of airports, whether they are hubs, pickup gateways or delivery gateways, indexed by g, i.e. $G = H \cup N_P \cup N_D$.
K	is the set of commodities (i.e. each origin-destination pair with positive demand), indexed by k.
K^n	is the set of commodities whose origin is gateway n (if $n \in N_P$), or whose destination is gateway n (if $n \in N_D$).
R^{nh}	is the set of pickup (delivery) routes that connect gateway n with hub h.

Parameters

b^k	is the demand for commodity k, in pounds.
b^n	is the pickup demand of gateway n (if $n \in N_P$) or its delivery demand (if $n \in N_D$), in pounds.
c_r	is the cost of flying route r, in dollars.
e^h	is the sorting capacity of hub h during the full sorting period, in pounds.
m^f	is the number of units of aircraft type f available.
s^h	is the number of take-off and landing slots available at hub h.
u_r	is the capacity of route r, in pounds.
α_r^{nh}	equals 1 if route r connects gateway n with hub h, 0 otherwise.
γ_r^h	is the number of times route r lands and/or departs from hub h.
η_r^{fg}	equals 1 if route r starts its itinerary at airport g with aircraft type f; -1 if route r ends its itinerary at airport g with aircraft type f; 0 otherwise.
φ_r^n	equals 1 if route r serves gateway n, 0 otherwise.

Decision variables

x_r^{nh}	is the demand of gateway n shipped on pickup (delivery) route r directed to (from) hub h, in pounds. They are referred to as *aggregated flow variables.*
y_r	is the number of times that route r is operated. They are referred to as *design variables.*
z^{kh}	is the demand from commodity k assigned to hub h, in pounds. They are referred to as *hub assignment variables.*

The RHC model is formulated as follows.

$$\min \quad \sum_{r \in R} c_r y_r \tag{1}$$

s.t.

$$\sum_{h \in H} z^{kh} \geq b^k, \qquad \forall\, k \in K, \tag{2}$$

$$\sum_{r \in R^{nh}} x_r^{nh} - \sum_{k \in K^n} z^{kh} \geq 0, \qquad \forall\, n \in N,\ h \in H, \tag{3}$$

$$\sum_{h \in H} \sum_{n \in N_r} \alpha_r^{nh} x_r^{nh} \leq u_r y_r, \qquad \forall\, r \in R, \tag{4}$$

$$\sum_{r \in R} \eta_r^{fg} y_r = 0, \qquad \forall\, g \in G, \;\; f \in F, \tag{5}$$

$$\sum_{r \in R_P^f} y_r \leq m^f, \qquad \forall\, f \in F, \tag{6}$$

$$\sum_{r \in R} \gamma_r^h y_r \leq s^h, \qquad \forall\, h \in H, \tag{7}$$

$$\sum_{k \in K} z^{kh} \leq e^h, \qquad \forall\, h \in H, \tag{8}$$

$$\sum_{r \in R^{nh'}} y_r \geq 1, \qquad \forall\, n \in N,\ h' \in H : e^{h'} = \max_{h \in H}(e^h), \tag{9}$$

$$\sum_{r \in R^{nh}} y_r \geq \frac{z^{kh}}{b^k}, \qquad \forall\, n \in N,\ k \in K^n,\ h \in H, \tag{10}$$

$$\min(b^n, u_r) y_r \geq \sum_{h \in H} \alpha_r^{nh} x_r^{nh}, \qquad \forall\, r \in R,\ n \in N_r, \tag{11}$$

$$\sum_{r \in R} \min\Big(\sum_{n \in S} \varphi_r^n b^n, u_r\Big) y_r \geq \sum_{n \in S} b^n, \qquad \forall\, S \subseteq N_P \text{ and } S \subseteq N_D, \tag{12}$$

$$y_r \in \mathbb{Z}_+, \qquad \forall\, r \in R, \tag{13}$$

$$x_r^{nh}, z^{kh} \in \mathbb{R}_+, \qquad \forall\, r \in R, \;\; n \in N, \;\; h \in H,\ k \in K. \tag{14}$$

The objective function (1) minimizes the total cost of the routes operated. Equations (2)--(3), called the *forcing* constraints, enforce the flow of the commodities from their origins to their destinations. The aircraft capacity constraints (4) limit the weight hauled by each aircraft. Constraints (5) model the circulation problem for the gateways and hubs. The counting constraints (6) ensure that no more aircrafts are employed than those available per type. The slots constraints (7) limit the number of landing and departures that can be operated at hubs during the night. The sorting constraints (8) ensure that the total sorting capacity of the hubs during the night is respected. The main hub connectivity condition (9) forces each gateway to be connected to the main hub (i.e. the one with the largest sorting capacity). Equations (10)-(12) are three families of valid inequalities employed in the RHC model to improve the speed for solving the ESSND problem[1]. The commodity connectivity inequalities (10) ensure that there is an origin-hub-destination connection for each commodity. When possible, the strong linking inequalities (11) reduce the capacity of a route to the maximum demand of a served gateway, thus reducing the fractionality of the linear relaxation. The strong cutset inequalities (12) ensure that the routes serving a subset of gateways have enough capacity to haul the gateways' demands. Finally, equations (13)-(14) define the domain of the variables.

4.2. Complex routes modeling

In this section, we study how to modify the RHC model so as to incorporate the complex routes. Most of the constraints have to be modified and new variables and constraints are needed. First, we explain the changes required for each complex route. Second, the concept of critical legs is introduced, as it is required to model the aircraft capacities of all the route types. Finally, the changes required in the valid inequalities are discussed. We introduce the notations required to model the complex routes progressively, when they are needed for the new or modified constraints.

Two-hub routes

The two-hub pickup (delivery) routes do not allow loading (unloading) packages at the first (second) hub they visit, and are thus similar to multi-leg routes, limiting the required modifications to the model. Their capability to haul packages destined to two hubs can be modeled with the aggregated flow variables x_r^{nh}. Additionally, since a two-hub pickup (delivery) route lands and departs from the first (second) hub, the parameter γ_r^h can take a value of 2.

[1] In Quesada Pérez et al. (2018), covers are also included, but they are removed here as they are not compatible with the time-dependent hub capacity constraints that we will add to the model in Section 4.3.

Transload routes

Transload routes carry packages destined for multiple hubs, a capability that can be modeled with the aggregated flow variables x_r^{nh}.

Direct routes

As direct routes bypass the hubs, their flows need to be considered in the flow forcing constraints (2). Additionally, the aircrafts performing direct routes have to be included in the counting constraints (6), because they cannot perform pickup or delivery routes (see Section 3.2). The new flow forcing constraints and counting constraints are written as follows.

$$\sum_{h \in H} z^{kh} + \sum_{r \in B^k} x_r^k \geq b^k, \qquad \forall\, k \in K, \tag{15}$$

$$\sum_{r \in R_P^f} y_r + \sum_{r \in B^f} y_r \leq m^f, \qquad \forall\, f \in F, \tag{16}$$

with *sets*

- B is the set of direct routes;
- B^k is the set of direct routes that can pickup and deliver commodity k;
- B^f is the set of direct routes flown with aircraft type f;

with *variables*

- x_r^k with $r \in B$, is the demand from commodity k shipped on direct route r, in pounds; they are referred to as *commodity flow variables*.

Inter-hub routes

The first consequence of introducing the inter-hub routes is that they allow one-leg and multi-leg routes to haul packages directed to multiple hubs, which can be modeled with the aggregated flow variables x_r^{nh}. For pickup gateways, the set R^{nh} thus has to include the routes that connect gateway n to hub h directly and indirectly (via an inter-hub route). Additionally, to enable the sorting of their loads at their destination hub, the inter-hub routes may depart before the due time of their origin hub. As only some of the pickup routes can connect packages to an inter-hub route on time, the flow conservation of packages has to be modeled explicitly. The new constraints (17) ensure the on-time connectivity of packages between pickup and inter-hub routes. The flow conservation of aircrafts has to be explicitly modeled too, because an inter-hub route is not operated with a dedicated aircraft and thus a pickup route must provide it the aircraft. The new equations (18) ensure that, to fly an inter-hub route, a pickup route performed with the same aircraft must have arrived to the origin of the inter-hub route, early enough to allow the unloading of its pickup packages and the loading of the inter-hub packages.

$$\sum_{r' \in I^{hh'}} \lambda_r^{r'h} x_{r'} - \sum_{r' \in R_P} \sum_{n \in N_{r'}} \Lambda_r^{r'h} x_{r'}^{nh'} \geq 0, \qquad \forall\, h \in H,\ h' \in H \setminus \{h\},\ \ r \in I^{hh'}, \tag{17}$$

$$\sum_{r' \in R_P} \Delta_r^{r'h} y_{r'} - \sum_{r' \in I : \eta_{r'}^{fh} = 1} \delta_r^{r'h} y_{r'} \geq 0, \qquad \forall\, h \in H,\ f \in F,\ r \in I : \eta_r^{fh} = 1, \tag{18}$$

with *sets*

- I is the set of inter-hub routes;
- $I^{hh'}$ is the set of inter-hub routes flying from hub h to hub h';
- R_P is the set of pickup routes;

with parameters

- $\lambda_r^{r'h}$ with r and $r' \in I$, equals 1 if inter-hub routes r and r' depart both from hub h and route r' departs at the same time or after route r, 0 otherwise;
- $\Lambda_r^{r'h}$ with $r \in I$ and $r' \in R_P$, equals 1 if the pickup route r' can transfer packages to the inter-hub routes departing from h at the departure time of inter-hub route r or later, but not to those departing before; 0 otherwise;
- $\Delta_r^{r'h}$ with $r \in I$ and $r' \in R_P$, equals 1 if the aircraft of pickup route r' arrives to hub h early enough to perform inter-hub route r, 0 otherwise; notice that this value can only be 1 if routes r and r' are performed with the same aircraft type;
- $\delta_r^{r'h}$ with r and $r' \in I$, equals 1 if the inter-hub routes r and r' depart both from hub h and route r' departs at the same time or before route r, 0 otherwise.

with *variables*

- x_r with $r \in I$, is the demand hauled by inter-hub route r, in pounds.

To give more intuition about these new constraints, we give them for the example given in Figure 4. The constraints (17) are $x_1 + x_2 - x_3^{a1} - x_4^{a1} - x_4^{b1} \geq 0$, for inter-hub route 1; and $x_2 - x_4^{a1} - x_4^{b1} \geq 0$, for inter-hub route 2. They force the pickup demand of route 4 directed to hub 1 to be transferred to the inter-hub route 2, and the pickup demand from route 3 to be transferred to inter-hub route 1 or 2. Assuming that all the routes are performed with the same aircraft type, the equations (18) are $y_3 - y_1 \geq 0$, for inter-hub route 1; and $y_3 + y_4 - y_1 - y_2 \geq 0$, for inter-hub route 2. The equation for inter-hub route 2 ensures that this route can be flown only if either pickup routes 3 or 4 are flown, and that inter-hub routes 1 and 2 are both flown only if pickup routes 3 and 4 are flown.

Early routes

The early routes require a new family of equations to ensure that they haul not more than the available early demands of each gateway. Since the early demands are allowed but not forced

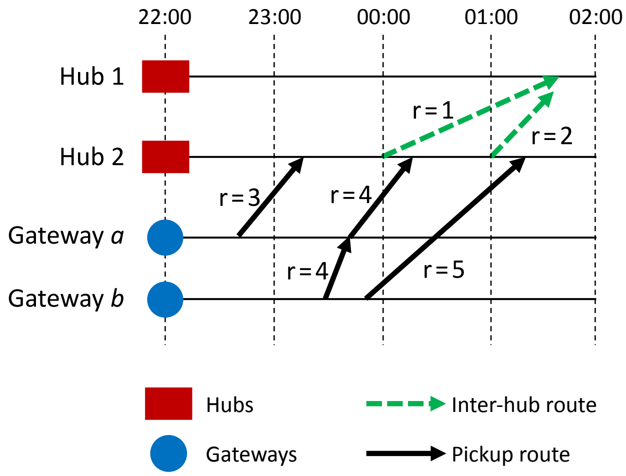

Figure 4: Time space network illustrating the connectivity of pickup routes to inter-hub routes.

to depart early (see Section 3.2), these equations are inequalities. The aggregated flow forcing constraints (3) ensure the early demands to be carried either by early routes or by other route types. The early flows can be modeled as follows.

$$\sum_{r\in R}\sum_{h\in H} \varepsilon_r^n x_r^{nh} \leq b^{n\varepsilon}, \qquad \forall\, n \in N, \tag{19}$$

with *parameters*

$b^{n\varepsilon}$ is the early demand of gateway n, in pounds;

ε_r^n equals 1 if route r is an early route serving the gateway n, 0 otherwise.

Critical legs

For modeling the ESSND problem, a capacity constraint is required for each route leg. However, in the RHC model, it is assumed that the pickup (delivery) routes can only release (seize) packages at hubs. For a multi-leg route, this means that only the leg connecting a gateway and a hub requires a capacity constraint, because the other legs have the same aircraft capacity and are loaded with fewer packages. This helps reducing the complexity of the model as the aircraft capacity constraints typically lead to highly fractional linear relaxations (Armacost et al., 2002; Gendron et al., 1999). However, this assumption of the RHC model is not valid for direct and transload routes as they can load and unload packages at each intermediary stop. To include only the required capacity constraints when the complex routes are considered, we introduce the concept of *critical legs*, which are the legs of a route for which a capacity constraint needs to be included. For one-leg, multi-leg, two-hub, transload and inter-hub routes, each leg connecting a gateway and a hub (the two hubs for inter-hub routes) is critical. The additional legs of transload routes are critical if the demands of the gateways they serve are larger than their capacity. For direct routes, each leg is critical since they can load and unload packages at each stop. In sum, by concentrating on the critical legs, half or more of the aircraft capacity constraints of multi-leg and two-hub routes can be omitted, as well as up to half of the capacity constraints of transload routes. The capacity constraints for pickup/delivery routes, for direct routes and for inter-hub routes can be modeled respectively with the following equations.

$$\sum_{h\in H}\sum_{n\in N_r}\alpha_a^{nh}x_r^{nh} \le u_a y_r, \qquad \forall\, r\in R\setminus(B\cup I),\ a\in\mathcal{A}_r, \tag{20}$$

$$\sum_{k\in K_r}\alpha_a^{k}x_r^{k} \le u_a y_r, \qquad \forall\, r\in B,\ a\in\mathcal{A}_r, \tag{21}$$

$$x_r \le u_r y_r, \qquad \forall\, r\in I, \tag{22}$$

with *sets*

$\mathcal{A}_r$	is the set of critical legs of route r, indexed by a;
K_r	is the set of commodities that can be loaded into route r;

with *parameters*

u_a	is the capacity of critical leg a, in pounds;
α_a^{nh}	equals 1 if critical leg a connects gateway n with hub h, even if it is done indirectly via an inter-hub route, 0 otherwise;
α_a^{k}	equals 1 if commodity k can be served by critical leg a, 0 otherwise.

Valid inequalities

The commodity connectivity constraints (10) ensure that there is an origin-hub-destination connection for each commodity assigned to a hub. With inter-hub routes, a commodity package can visit more than one hub to achieve its destination, thus increasing the number of such connections. To account for this, the adjusted commodity connectivity constraints can be formulated as follows.

$$\sum_{r\in R^{nh}}\beta_r^{nh}y_r + \sum_{r\in I^{nh}}\frac{1}{2}y_r \ge \frac{z^{kh}}{b^k}, \qquad \forall\, n\in N,\ k\in K^n,\ h\in H, \tag{23}$$

with *sets*

I^{nh} is the set if inter-hub routes that can connect indirectly pickup gateway n to hub h;

with *parameters*

β_r^{nh} equals 1 if route r connects directly gateway n with hub h; $\dfrac{1}{2}$ if it connects them indirectly via an inter-hub route; 0 otherwise.

The strong linking inequalities (11) need to be modified in two ways. First, two families of inequalities have to be created: one for the aggregated flows of pickup and delivery routes, and one for the commodity flows of direct routes. Second, as early routes can transport less packages than other routes, their strong linking inequalities can be strengthened by modifying the demand parameter b^n. This leads to the following equations.

$$\min(b_r^n, u_r)y_r \geq \sum_{h\in H} x_r^{nh}, \qquad \forall\, r \in R \setminus B,\ n \in N_r, \tag{24}$$

$$\min(b^k, u_r)y_r \geq x_r^k, \qquad \forall\, r \in B,\ k \in K_r \tag{25}$$

with *parameters*

b_r^n is the maximum demand from gateway n available to be loaded into route r, in pounds. For early routes $b_r^n = b^{n\varepsilon}$, for other routes $b_r^n = b^n$;

The strong cutset inequalities (12) state that the capacity of the routes serving a cutset must be sufficient to carry the total cutset demand. Therefore, the demand parameter b_r^n is used to account for early routes. The capacity of direct routes cannot be used directly, because these routes can load and unload packages at any intermediate stop, and the total commodity demand they haul may thus be higher than their total capacity. To obtain the new strong cutset inequalities, the capacity of direct routes can be substituted by their flows, leading to the following equations.

$$\sum_{r\in R\setminus B} \min(\sum_{n\in S} \varphi_r^n b_r^n, u_r)y_r + \sum_{r\in B^S} \sum_{k\in K_r^S} x_r^k \geq \sum_{n\in S} b^n, \qquad \forall\, S \subseteq N_P \text{ or } S \subseteq N_P, \tag{26}$$

with *sets*

B^S is the set of direct routes that can pickup and deliver packages from $S \subseteq N$;

K_r^S is the set of commodities that can be loaded into route r with origin in S (if $S \subseteq N_P$), or with destination in S (if $S \subseteq N_D$).

4.3. Hub sorting capacity

As hub operations represent an important cost for express integrators, the sorting capacity of the hubs is an important constraint of the ESSND problem. In the literature, some researchers neglect this constraint (e.g., Armacost et al., 2002; Louwerse et al., 2014), while others include one time-independent sorting capacity constraint per hub (e.g., Shen, 2004; see Section 4.1). These two modeling approaches do not prevent a concentrated arrival of aircrafts to a hub during a specific time slot. In particular, if this occurs close to the hubs due times, the sorting process may be delayed, rendering a solution infeasible in practice. To avoid this issue, a more complex but more accurate approach is needed, in the form of time-dependent sorting capacity constraints. Moreover, direct and early pickup routes help relieving the sorting load of hubs during the peak hours, as direct routes bypass the hubs and early pickup routes advance the arrival of packages to hubs. Therefore, the contribution of these routes can be better evaluated with the use of time-dependent sorting capacity constraints. that Time-dependent sorting capacity can be modeled with the so-called *staggering constraints* (Barnhart and Schneur, 1996; Kim et al., 1999). They are obtained by dividing the total hub sorting period into non-overlapping intervals $t \in T$. Then, one sorting capacity constraint is formulated for each interval, limiting the number of packages arriving between the start time of t and the hub release time to the maximum hub sorting capacity during that time. The arrival of packages to hubs during a given time span can be computed from the arrival times of the routes that carry them, which are computed during the route generation process. This is true even for the inter-hub flows because all the inter-hub routes sharing the same destination hub are scheduled to land within the same sorting interval. The staggering constraints can be stated as follows.

$$\sum_{r \in R_P} \sum_{n \in N_r} \theta_r^{th} x_r^{nh} \leq e^{th}, \qquad \forall\, h \in H,\ t \in T, \tag{27}$$

with *sets*

- T — is the set of sorting intervals, indexed by t;

with *parameters*

- e^{th} — is the sorting capacity of hub h from the sorting interval t to the end of the sorting process (i.e. the hub release time), in pounds;
- θ_r^{th} — equals 1 if packages from route r arrive to hub h during the sorting interval t or afterwards, including indirect connections via an inter-hub route; 0 otherwise.

4.4. Route and Hub model with Cuts and Complex Routes

In this section, we can now give the Route and Hub model with Cuts and Complex Routes (RHCCR), in which five complex route types (two-hub, transload, direct, inter-hub and early routes) are considered, and in which the hub sorting capacity constraints are time-dependent. The RHCCR model is given by the following equations.

$$\min \quad \sum_{r\in R} c_r y_r \tag{28}$$

s.t.

Commodity flow constraints:

$$\sum_{h\in H} z^{kh} + \sum_{r\in B^k} x_r^k \geq b^k, \qquad \forall\, k\in K, \tag{29}$$

$$\sum_{r\in R^{nh}} x_r^{nh} - \sum_{k\in K^n} z^{kh} \geq 0, \qquad \forall\, n\in N,\ h\in H, \tag{30}$$

$$\sum_{r'\in I^{hh'}} \lambda_r^{r'h} x_{r'} - \sum_{r'\in R_P}\sum_{n\in N_{r'}} \Lambda_r^{r'h} x_{r'}^{nh'} \geq 0, \qquad \forall\, h\in H,\ h'\in H\setminus\{h\},\ \ r\in I^{hh'}, \tag{31}$$

$$\sum_{r\in R}\sum_{h\in H} \varepsilon_r^n x_r^{nh} \leq b^{n\varepsilon}, \qquad \forall\, n\in N, \tag{32}$$

Aircraft capacity and flow constraints:

$$\sum_{h\in H}\sum_{n\in N_r} \alpha_a^{nh} x_r^{nh} \leq u_a y_r, \qquad \forall\, r\in R\setminus(B\cup I),\ a\in\mathcal{A}_r, \tag{33}$$

$$\sum_{k\in K_r} \alpha_a^k x_r^k \leq u_a y_r, \qquad \forall\, r\in B,\ a\in\mathcal{A}_r, \tag{34}$$

$$x_r \leq u_r y_r, \qquad \forall\, r\in I, \tag{35}$$

$$\sum_{r\in R} \eta_r^{fg} y_r = 0, \qquad \forall\, g\in G,\ \ f\in F, \tag{36}$$

$$\sum_{r\in R_P^f} y_r + \sum_{r\in B^f} y_r \leq m^f, \qquad \forall\, f\in F, \tag{37}$$

$$\sum_{r'\in R_P} \Delta_r^{r'h} y_{r'} \ - \sum_{r'\in I:\eta_{r'}^{fh}=1} \delta_r^{r'h} y_{r'} \geq 0, \qquad \forall\, h\in H,\ f\in F,\ r\in I:\eta_r^{fh}=1, \tag{38}$$

Hub slot and sorting capacity constraints:

$$\sum_{r\in R} \gamma_r^h y_r \leq s^h, \qquad \forall\, h\in H, \tag{39}$$

$$\sum_{r\in R_P}\sum_{n\in N_r}\theta_r^{th}x_r^{nh} \leq e^{th}, \qquad \forall\, h\in H,\ t\in T, \tag{40}$$

Main hub connectivity condition:

$$\sum_{r\in R^{nh'}} y_r \geq 1, \qquad \forall\, n\in N,\ h'\in H : e^{h'} = \max_{h\in H}(e^h), \tag{41}$$

Valid inequalities:

$$\sum_{r\in R^{nh}}\beta_r^{nh}y_r + \sum_{r\in I^{nh}}\frac{1}{2}y_r \geq \frac{z^{kh}}{b^k}, \qquad \forall\, n\in N,\ k\in K^n,\ h\in H, \tag{42}$$

$$\min(b_r^n, u_r)y_r \geq \sum_{h\in H} x_r^{nh}, \qquad \forall\, r\in R\setminus B,\ n\in N_r, \tag{43}$$

$$\min(b^k, u_r)y_r \geq x_r^k, \qquad \forall\, r\in B,\ k\in K_r, \tag{44}$$

$$\sum_{r\in R\setminus B}\min(\sum_{n\in S}\varphi_r^n b_r^n, u_r)y_r + \sum_{r\in B^S}\sum_{k\in K_r^S} x_r^k \geq \sum_{n\in S} b^n, \qquad \forall\, S\subseteq N_P \text{ or } S\subseteq N_P, \tag{45}$$

Logical constraints:

$$y_r \in \mathbb{Z}_+, \qquad \forall\, r\in R, \tag{46}$$

$$x_r^{nh}, x_r, x_r^k, z^{kh} \in \mathbb{R}_+, \qquad \forall\, r\in R,\ \ n\in N,\ \ h\in H,\ k\in K. \tag{47}$$

The objective function (28) minimizes the total cost of the routes operated. The commodity assignment constraints (29) assign the commodities either to a hub or to a direct route. The aggregated flow forcing constraints (30) ensure that the gateway demands carried by the pickup and delivery routes match the hub assignments. The inter-hub flow conservation constraints (31) ensure that the packages hauled by an inter-hub route are connected on time by a pickup route. Equations (32) limit the loads of early routes to the available early demands of the gateways. The aircraft capacity constraints (33)-(35) limit the weight carried by the pickup and delivery, by the direct and by the inter-hub routes respectively. Constraints (36) model the circulation problem for the gateways and hubs. The counting constraints (37) ensure that no more aircrafts are employed than those available of each type. Equations (38) ensure the flow conservation of aircrafts performing an inter-hub route. The slots constraints (39) limit the number of landing and departures that can be performed at a hub during the night. The staggering constraints (40) ensure that the hub sorting capacities are respected for each time interval. The main hub connectivity condition (41) enforces each gateway to be connected to the main hub. The commodity connectivity inequalities (42) ensure that there exists an origin-hub-destination connection for each commodity assigned to a hub. The strong linking inequalities (43)–(44) reduce the capacity of the routes when it is larger than the demands they can serve. The strong cutset inequalities (45) ensure that the routes serving a subset of gateways have enough capacity to carry the gateways' demands.[2] Finally, equations (46)-(47) are the logical constraints, which define the domain of the variables. Note that since our RHCCR model focuses on strategic decisions, it does not include some constraints required for the tactical decision making process of express integrators. Specifically, the RHCCR formulation does not include the runway staggering constraints (Barnhart and Schneur, 1996; Meuffels, 2015) nor accounts for the exchange of unit load devices between aircrafts (Meuffels, 2015).

5. Computational experiments

In this section, we present a series of computational experiments aiming to assess the impact of the five types of complex routes considered in our ESSND model: two-hub, transload, direct, inter-hub and early routes. We start from an instance with 2 hubs, 77 gateways, 7 aircraft types and 4,345 commodities, built for the purpose of this research by FedEx Express Europe as to be representative of real life instances. Two basic instances are then obtained from the initial instance by changing the hub sorting capacities. In the first instance, within the interval between its due and release times, the first (second) hub can sort an amount of packages equivalent to 70% (30%) of the total network demand. In the second instance, these percentages are swapped between the hubs. Twelve other instances are built by modifying the value of one important parameter while keeping the other parameters to their default values. Three parameters are modified. The *network demand* is increased by 50% or 100% (with the hub sorting capacities increased proportionally). The *hub sorting capacities* are reduced by 10% or 20%.[3] The *distance between the two hubs* is multiplied by a factor of 2 or 3.

A total of 140 experiments are performed on these 14 instances to study the contribution of the complex routes. First, the 14 instances are solved considering only the standard routes.

[2] In our experiments, the strong cutset inequalities are included for $|S| \leq 2$ and for $|S| = 2$ they are only generated for pairs of gateways that share at least one route.

[3] Reducing the capacity by 30% would render the instances infeasible if only standard routes were considered. Reducing the capacity by 20% thus leads to congested hubs.

Table 2: Performance of the RHCCR model when adding the complex routes to the standard routes.

	Standard routes only	Standard routes plus								**All routes**[b]
		Two-hub	Trans-load	Direct	Inter-hub	Early[a] 10p30m	Early[a] 10p60m	Early[a] 20p30m	Early[a] 20p60m	
All instances										
Avg. no. of non-ferry routes[c]	7,382	7,492	7,865	15,777	7,394	9,225	10,131	10,214	12,821	**21,821**
Avg. optimality gaps (%)	2.89	2.71	2.89	5.01	2.64	3.63	4.57	3.74	5.05	**5.49**
Avg. cost change[d] (%)	0.00	-0.16	-0.02	0.07	-3.44	-0.42	-0.96	-1.57	-1.47	**-4.86**
Instances with improved costs when the complex routes are used in the network										
No. of improved instances		2	2	6	14	8	9	11	7	**11**
With proven improvement[e]		0	0	1	5	0	2	4	3	**5**
Avg. cost change[d] (%)		-0.30	-0.07	-1.04	-3.44	-0.95	-1.73	-2.04	-3.55	**-6.42**
Best cost change (%)		-0.50	-0.07	-3.17	-10.93	-2.38	-5.19	-4.26	-9.59	**-13.75**

[a] Early ApBm = Early routes can serve A% of the total gateway demands B minutes early.
[b] With Early 20p60m.
[c] Even though the average number of ferry routes is large (13,878), it does not significantly affect the computational cost of the RHCCR model, because these routes do not require aircraft capacity constraints (4).
[d] Compared to the corresponding instance that include only standard routes.
[e] Larger than the optimality gap of the corresponding instance with only standard routes.

Second, they are solved with each complex route type added separately to the standard routes.[4] Regarding the early routes, we represent the early demands and early release times with the notation ApBm, meaning that A% of the total demand of a gateway can be released B minutes early. Four different early demand availabilities are considered: 10p30m, 10p60m, 20p30m and 20p60m. Regarding the inter-hub routes, a pickup route can transfer packages to an inter-hub route or its aircraft can perform the inter-hub route only if the pickup route arrives 75 minutes before the departure of the inter-hub route. Moreover, the due time of inter-hub routes is set to 80 minutes before the hub release time (overriding the regular hub due time by 94 minutes). Third, the instances are solved with all the route types considered (early routes with 20p60m). All the experiments are solved with our RHCCR formulation, using Java/Gurobi 7.0, running on a 2.7GHz Intel Core i7 CPU, with 16 GB in RAM. The solver is let run for 8 hours (all instances reached this limit) using 3 threads.

Tables 2 and 3 display the results of the 140 experiments. Table 2 gives the aggregated results on all instances and on instances with an improvement, while Table 3 details the impact of the network parameters (whether the complex routes are used in the network or not). As shown in Table 2, the optimality gaps remain low when complex routes are added to the standard routes

[4]Routes are generated if the demand of the gateways they serve can fill at least 30% of their aircraft capacity. This value is increased to 70% for early routes. For direct routes, it is increased to 60% when the network demands have their initial values, or to 80% or 90% when the network demands are increased by 50% or 100% respectively.

Table 3: Average cost change, in percentage, when the complex routes are added to the standard routes, for variations of some input parameters.

Dem.[a]	**Cap.**[b]	**Dist.**[c]	Standard routes plus								All route types[e]
			Two-hub	Trans-load	Direct	Inter-hub	Early[d] 10p30m	Early[d] 10p60m	Early[d] 20p30m	Early[d] 20p60m	
0	**0**	**1**	-0.21	-0.06	-0.49	-5.89	-0.72	-1.30	-2.25	-2.95	-7.68
+50%	**0**	**1**	-0.07	0.07	1.18	-0.76	0.46	0.41	-0.17	0.22	-0.68
+100%	**0**	**1**	-0.41	-0.06	0.78	-0.42	-0.25	0.26	-0.20	0.72	1.25
0	**−10%**	**1**	-0.10	-0.05	-0.75	-5.84	-0.48	-1.36	-2.18	-1.05	-7.30
0	**−20%**	**1**	-0.30	0.15	-0.91	-4.87	-0.89	-1.80	-2.78	-1.25	-7.72
0	**0**	**×2**	-0.25	-0.14	0.50	-4.18	-0.25	-0.85	-0.79	-1.29	-4.49
0	**0**	**×3**	0.21	-0.07	0.18	-2.09	-0.81	-2.11	-2.62	-4.73	-7.39
All instances			-0.16	-0.02	0.07	-3.44	-0.42	-0.96	-1.57	-1.47	**-4.86**

[a] Network demand increase.
[b] Hub sorting capacity reduction.
[c] Distance between hubs multiplier.
[d] Early ApBm = Early routes can serve A% of the total gateway demands B minutes early.
[e] With E 20p60m.

(maximum gap increase of 2.6% when all route types are included), with improved or similar costs. We thus showed how to model our five types of complex routes efficiently, without enforcing them in the solution nor fixing their loads, which allows to analyze their practical value. In the following, we discuss the contribution of each type of complex route individually, and then that of all routes together.

Inter-hub routes

As seen in Table 2, the inter-hub routes have a low impact on the computational cost of the ESSND model, as they reduce the optimality gap (to 2.64%) and their number is very small (12). They improve the results for all 14 instances, 5 of them with proven improvement (i.e. larger than the optimality gap of the corresponding instances with only standard routes), with average savings of 3.44%.

The benefits of inter-hub routes firstly come from reducing the number of aircrafts in the network, mainly by sparing small aircrafts directly connecting gateways with the secondary hub. In our experiments, by transferring 3.8% of the network demand, inter-hub routes lead, on average, to a reduction of 11% of the number of aircrafts and an increase of 76% of the number of gateways connected to the secondary hub. Consistently, Table 3 shows that savings are larger when the network demand is smaller. When the network demand is large, a lower number of small aircrafts is employed, which reduces the possibilities for inter-hub routes to consolidate demands from different origins and thus to spare aircrafts. Table 3 also shows that the savings

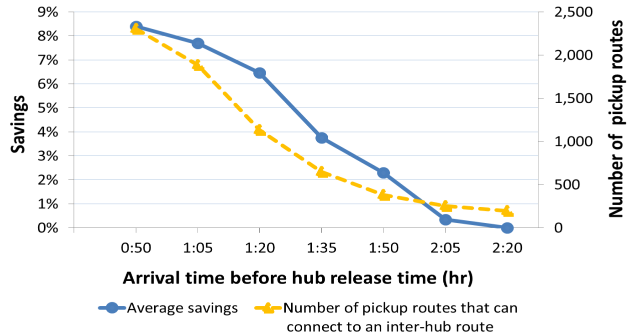

Figure 5: Average savings with different due times for the inter-hub routes, on the two basic instances.

coming from inter-hub routes tend to increase when hubs are located closer to the gateways' geographical center (Dist. 1). When hubs are more distant (Dist. ×3), the hubs are farther from this center, and thus the average distance between gateways and hubs increases (by 15% on average in our experiments). This reduces almost by half the number of pickup routes that can connect to the inter-hub routes on time. Inter-hub routes thus become less effective for consolidating packages and for reducing the number of direct gateway-hub connections.

To analyze the impact of the due times of inter-hub routes, Figure 5 shows how the savings and the number of pickup routes that can connect evolve with the due times on our two basic instances. It clearly appears that more savings are brought when more pickup routes can transfer packages to inter-hub routes. Naturally, the assumption we make that inter-hub routes are allowed to arrive after the hub due times impacts directly the underlying savings they can uncover. However, although inter-hub routes are given the possibility to arrive after the hub due time, their loads must respect the time-dependent hub sorting constraints (40), which ensures that these loads can be sorted on time at the hubs.

In sum, our results show that inter-hub routes yield savings in a broad range of scenarios, without representing a computational burden. Considering these routes is particularly beneficial when the inter-hub due times allow them to consolidate packages from a large number of pickup routes and when there is a significant number of small aircrafts in the network.

Direct routes

As shown in Table 2, when direct routes are included, the model can still be solved efficiently (optimality gap around 5%). In fact, the optimality gap increase (around 2 percentage points) is lower than may be expected, given that the inclusion of direct routes practically double the

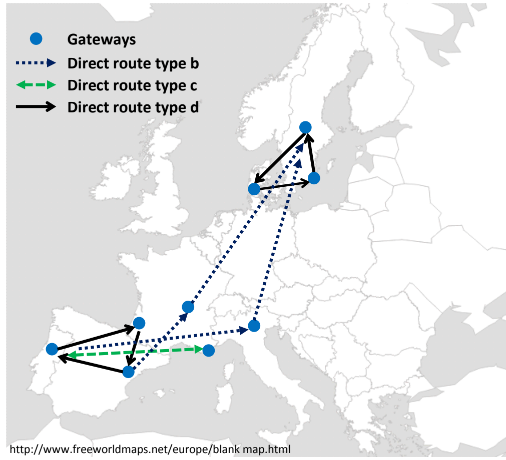

Figure 6: Direct routes from an instance with improved costs.

number of non-ferry routes in the network, thus considerably increasing the number of variables (+21%) and constraints (+92%). Table 2 also shows that, in almost half of the instances, direct routes yield average savings of around 1%. These savings are achieved by only slightly modifying the network: direct routes constitute 2.4% of the network routes and their loads represent 1.6% of the network demand. Table 3 shows that direct routes yield savings mainly when the hub sorting capacities decrease. Tight sorting capacities force to serve some gateways with faster, larger and more expensive aircrafts to advance the arrival of some packages to the hub sorting processes. As they bypass the hubs, direct routes relieve the sorting process and thus allow to avoid some of these inefficient flights.

Looking at the types of direct routes found in the solutions, we see that direct routes are generally performed with small aircrafts. In the instances in which they yield savings, 73% of the direct routes are performed with the smallest aircraft type, and all with the 3 smallest aircraft types. Furthermore, 54% of the used direct routes are simple two-leg non-cyclic routes (type b in Figure 3) and, together with the two-leg and three-leg cyclic routes (types c and d), they account for 95% of the direct routes. One-leg direct routes (type a) were never used. To illustrate, Figure 6 shows all the selected direct routes for one instance with improved costs (they are all of types b, c or d). It also shows that direct routes can connect both close or far gateways and that some gateways may be served by more than one direct route.

In conclusion, direct routes are worthwhile considering, as they are beneficial in certain network configurations, mainly when the hub sorting capacities are congested. Given their computational cost, it is recommendable to test whether they are beneficial for the specific network, and then decide if they should be included. Our results hint that the most promising direct routes are multi-leg and performed with small aircrafts.

Two-hub routes

As shown in Table 2, two-hub routes do not worsen the optimality gaps, and they produce low savings sporadically, in 2 out of 14 instances.[5] In these two instances, the used two-hub routes connect a far away gateway to the two hubs when the latter are close (with Dist. 1), showing that they are sometimes useful to avoid doubling the number of aircrafts needed on long distances.

Transload routes

The impact of transload routes shown in Tables 2 and 3 is unnoticeable both economically and in computational costs. The two instances in which the transload routes yield savings are those in which the hubs are more distant (Dist. 3), with a single transload route in each instance, transloading packages at a gateway located between the hubs.

Early routes

The RHCCR model can still be solved efficiently when the early routes are included as, despite the high number of early routes, the optimality gaps remain low (between 3.6% and 5.1%, see Table 2). While early routes increase the required computational effort, they also uncover the potential and hidden benefits of being able to release a relatively low proportion of the network demand early. For example, we can see in Table 2 that, assuming 20p30m, early routes grants cost improvements of 1.57%. Overall, early routes bring benefits in 35 out of 56 experiments with the different early demand availabilities.

Table 3 shows that early routes seem to be particularly beneficial when the hubs are not centrally located (Dist. $\times 3$), as these routes allow more connections between gateways and hubs located far from each other. These routes also tend to yield higher savings when the hub sorting systems are congested (Cap. -20%), since early pickup routes can advance the arrivals

[5]Although Table 3 shows mild cost improvements for several input variations when two-hub routes are considered, most of the times the two-hub routes are not used in the solutions. Thus, such improvements are more related to changes in the branching process.

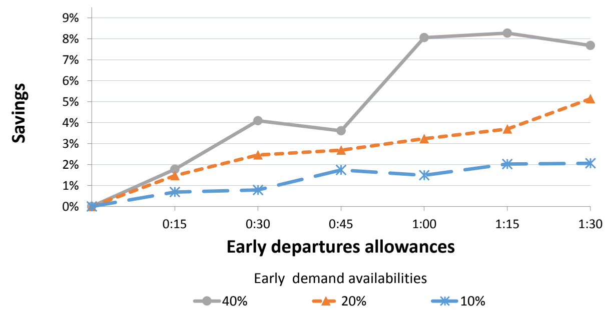

Figure 7: Average savings yielded by early routes with various early departure allowances, and various early demand availabilities, on the two basic instances.

of packages to hubs, and thus relieve the sorting load during critical hours.

Figure 7 shows how the early demand availabilities and the early departure allowances impact the capability of these routes to yield savings on the 2 basic instances. Unsurprisingly, the cost benefits increase with these two parameters. Larger economies of scales can be achieved when their early demand availabilities are increased, while the early departure allowances enable new gateway-hub connections with aircrafts that cannot achieve them otherwise. When 10% of the demand can be served early, savings are limited to 2%, while with 20% early availability, savings are already higher (around 2.5%) for an allowance of only 30 minutes. When a good proportion (40%) of the demand can be served early, savings rise up to 8%.

Overall, early routes can help to reduce the network costs if a carrier can release relatively small proportions of its demands early. These routes are particularly promising with congested hub sorting processes and with distant hubs.

All routes

Table 2 shows that, when all routes are included, their number is almost three times larger than the non-ferry standard routes. Yet, the increase of the optimality gap is limited (from 2.9% to 5.5%), which still allows us to assess the savings brought by all routes together. The complex routes lead to improve 11 out of 14 instances with average savings of 4.8%, and a maximum improvement of 13.7% in one instance. Complex routes give more flexibility when designing the network, ultimately leading to a more efficient network. In our experiments, on average, the complex routes reduce the total network capacity by 6% and the number of aircrafts by 9%. Naturally, the savings complex routes can bring depend on the fulfillment of their operating conditions, specially of the availability of early demands and the late arrival of inter-hub routes to their destination hubs. Table 3 shows that the savings are quite stable in most network configurations: the complex routes yield important savings regardless of the hub sorting capacities and hub locations (between 4.4% and 7.7%). However, when the network demand increases, results deteriorate. When it doubles, the number of complex routes is almost 10 times larger than the number of standard routes, vastly increasing the computational complexity of the problem (optimality gap of 7%).

6. Conclusions

The express shipment service network design (ESSND) problem consists in defining a network of flights that enables the overnight flow of packages from their origins to their destinations at minimum cost. A key characteristic of the models from the ESSND literature relates to the definition of the routes considered in the network design. The standard routes in these models are the one-leg, multi-leg and ferry routes. Few works also include some other route types (transload, direct and inter-hub routes) but usually by predefining their loads and/or by fixing them as part of the solution. Assessing the value of complex routes in the ESSND problem is an open question of academic and practical importance.

In this article, we study five types of complex routes: the two-hub, transload, direct, inter-hub and early routes. We develop the Route and Hub model with Cuts and Complex Routes (RHCCR), in which these five complex route types are modeled without enforcing them in the solution nor predefining their loads. Through numerical experiments, we show that our RHCCR model can be solved efficiently, as including all the complex routes causes an increase in the optimality gap (2.6 percentage points larger) that is low enough to allow the assessment of the economic contribution of the five types of complex routes. A close research collaboration with FedEx Express Europe helped us ground our development in the actual planning challenges faced by express air carriers. Our computational experiments are based on an instance built by our partner so as to be representative of a real industrial case (77 gateways, 2 hubs).

From the numerical experiments, the main lesson is that considering complex routes when solving the ESSND problem helps finding more efficient network designs. When looking at the route types separately, the inter-hub routes yield the best average savings (more than 3%). The early routes and the direct routes also lead to significant savings, in particular by relieving the traffic on sorting systems with limited capacity. The transload and two-hub routes yield mild savings on sporadic cases but do not represent a computational burden. When all route types are included in our experiments, they yield average savings of close to 5%. The benefits of complex routes are significant, as a cost improvement of 1% in the ESSND problem typically represent several millions of dollars per year for express integrators.

Some possible ways in which the research on the ESSND problem with complex routes can be extended are the following. First, column generation or heuristic methods could be developed to search for useful direct or early routes, instead of enumerating them all. Second, other types of complex routes could be designed and modeled, like multi-leg pickup/delivery routes that deliver/pickup packages directly at the intermediary gateways, routes that combine the operations of two or three aircrafts in a different way than transload routes, or ferry routes that are allowed to fly at any moment. Third, additional features could be included to our strategic RHCCR model to address some characteristics of the tactical ESSND problem. We could for example add time-dependent constraints on the number of take-offs and landings at hubs, or model explicitly the exchange of unit load devices between aircrafts. Finally, it would be of practical value to study if some of the complex routes can lead to more robust network designs when considering demand uncertainty in the ESSND problem.

Acknowledgment

This project was funded by the Brussels-Capital Region and Innoviris.

References

Agarwal, Y.K., 2002. Design of capacitated multicommodity networks with multiple facilities. Operations Research 50, 333–344.

Armacost, A.P., Barnhart, C., Ware, K.A., 2002. Composite variable formulations for express shipment service network design. Transportation Science 36, 1–20.

Armacost, A.P., Barnhart, C., Ware, K.A., Wilson, A.M., 2004. UPS optimizes its air network. Interfaces 34, 15–25.

Barnhart, C., Krishnan, N., Kim, D., Ware, K., 2002. Network design for express shipment delivery. Computational Optimization and Applications 21, 239–262.

Barnhart, C., Schneur, R.R., 1996. Air network design for express shipment service. Operations Research 44, 852–863.

Benhamiche, A., Mahjoub, A.R., Perrot, N., Uchoa, E., 2016. Unsplittable non-additive capacitated network design using set functions polyhedra. Computers & Operations Research 66, 105–115.

Boeing Commercial Airplanes, 2018. World air cargo forecast 2018-2037. Technical Report. Seattle, WA: The Boeing World Air Cargo Forecast Team, Boeing Commercial Airplanes.

Crainic, T.G., 2000. Service network design in freight transportation. European Journal of Operational Research 122, 272–288.

Fleuren, H., Goossens, C., Hendriks, M., Lombard, M.C., Meuffels, I., Poppelaars, J., 2013. Supply chain–wide optimization at TNT express. Interfaces 43, 5–20.

Gendron, B., Crainic, T.G., Frangioni, A., 1999. Multicommodity capacitated network design, in: Telecommunications network planning. Springer, pp. 1–19.

Gouveia, L., Joyce-Moniz, M., Leitner, M., 2018. Branch-and-cut methods for the network design problem with vulnerability constraints. Computers & Operations Research 91, 190–208.

Kim, D., Barnhart, C., Ware, K., Reinhardt, G., 1999. Multimodal express package delivery: A service network design application. Transportation Science 33, 391–407.

Kuby, M.J., Gray, R.G., 1993. The hub network design problem with stopovers and feeders: The case of federal express. Transportation Research Part A: Policy and Practice 27, 1–12.

Louwerse, I., Mijnarends, J., Meuffels, I., Huisman, D., Fleuren, H., 2014. Scheduling movements in the network of an express service provider. Flexible Services and Manufacturing Journal 26, 565–584.

Meuffels, W.J.M., 2015. The design of road and air networks for express service providers. Ph.D. thesis. Tilburg University, School of Economics and Management.

Quesada Pérez, J.M., Lange, J.C., Tancrez, J.S., 2018. A multi-hub express shipment service network design model with flexible hub assignment. Transportation Research Part E: Logistics and Transportation Review 120, 116–131.

Schenk, L., Klabjan, D., 2010. Intra market optimization for express package carriers with station to station travel and proportional sorting. Computers & Operations Research 37, 1749–1761.

Shen, S., 2004. Logistics service network design: Models, algorithms, and applications. Ph.D. thesis. Massachusetts Institute of Technology.

Tong, L., Zhou, X., Miller, H.J., 2015. Transportation network design for maximizing space–time accessibility. Transportation Research Part B: Methodological 81, 555–576.

Wang, S., Meng, Q., Yang, H., 2013. Global optimization methods for the discrete network design problem. Transportation Research Part B: Methodological 50, 42–60.

Wieberneit, N., 2008. Service network design for freight transportation: A review. OR spectrum 30, 77–112.