



LOUVAIN

School of Management
RESEARCH INSTITUTE

WORKING PAPER

2016/07

A multi-hub express shipment
service network design model
with flexible hub assignment

José Miguel Quesada,
Jean-Sébastien Tancrez,
Louvain School of Management

Jean-Charles Lange

Louvain School of Management

Working Paper Series

Editor : Prof. Frank Janssen
(president-ilsm@uclouvain.be)

A multi-hub express shipment service network design model with flexible hub assignment

José Miguel Quesada, Louvain School of Management
Jean-Sébastien Tancrez, Louvain School of Management
Jean-Charles Lange

Summary

The Express Shipment Service Network Design problem consists in designing a network of flights that, at minimum cost, enables the overnight transportation of packages from their origins to their destinations. To ease the tractability of the multi-hub version of this problem, most of the modeling approaches in the literature assume that the allocation of packages to hubs is an input of the problem. In this paper, we develop a formulation in which the allocation of packages to hubs is a decision of the model. Our formulation is strengthened with three families of valid inequalities and the forcing constraints are reformulated for reducing the number of variables and constraints. The efficiency and effectiveness of our model is tested with an extensive set of numerical experiments. First, on limited size instances (20 gateways and 6 equipment types), we show that our model is efficient (a few percent optimality gap in two hours) and compares favorably to approaches from the literature. Then, the performance of the model is assessed on testing instances built for us by FedEx Express Europe so as to be representative of real size instances (76 gateways and 7 equipment types), and also shows a significant improvement compared to models from the literature.

Keywords: service network design; express integrator; air network; flexible hub assignment; valid inequalities.

JEL Classification: C61

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

Corresponding author:

PhD student José Miguel Quesada.
Center for Excellence: CESC
Louvain School of Management / Campus Mons
Université catholique de Louvain
151, Chaussée de Binche
7000 Mons, BELGIUM
Email : jose.quesada@uclouvain.be

The papers in the WP series have undergone only limited review and may be updated, corrected or withdrawn without changing numbering. Please contact the corresponding author directly for any comments or questions regarding the paper.
President-ilsm@uclouvain.be, ILSM, UCL, 1 Place des Doyens, B-1348 Louvain-la-Neuve, BELGIUM
www.uclouvain.be/ilsm and www.uclouvain.be/lsm_WP

A Multi-Hub Express Shipment Service Network Design Model with Flexible Hub Assignment

José Miguel Quesada

Université catholique de Louvain, Chaussée de Binche, 151, 7000 Mons, Belgium. jose.quesada@uclouvain.be

Jean-Charles Lange

Jean-Sébastien Tancrez

Université catholique de Louvain, Chaussée de Binche, 151, 7000 Mons, Belgium. js.tancrez@uclouvain.be

The Express Shipment Service Network Design problem consists in designing a network of flights that, at minimum cost, enables the overnight transportation of packages from their origins to their destinations. To ease the tractability of the multi-hub version of this problem, most of the modeling approaches in the literature assume that the allocation of packages to hubs is an input of the problem. In this paper, we develop a formulation in which the allocation of packages to hubs is a decision of the model. Our formulation is strengthened with three families of valid inequalities and the forcing constraints are reformulated for reducing the number of variables and constraints. The efficiency and effectiveness of our model is tested with an extensive set of numerical experiments. First, on limited size instances (20 gateways and 6 equipment types), we show that our model is efficient (a few percent optimality gap in two hours) and compares favorably to approaches from the literature. Then, the performance of the model is assessed on testing instances built for us by FedEx Express Europe so as to be representative of real size instances (76 gateways and 7 equipment types), and also shows a significant improvement compared to models from the literature.

Key words: service network design; express integrator; air network; flexible hub assignment; valid inequalities.

1. Introduction

In the transport industry, express carriers promise the fastest and most flexible logistic solutions, which makes them an attractive option for the transportation of high profit goods. According to a study by Boeing Commercial Airplanes (2014), in recent years the international express traffic has continued to grow faster than the traffic of the world air cargo, expanding by 8.9% in 2012 and 5.8% in 2013. In particular, express carriers offer different types of services for the door-to-door transportation of packages, each guaranteeing delivery before a specific date and time. One of the main service options offered by these carriers is the overnight deliveries within regions as large as the US, Europe or the Middle East. Carrying large numbers of packages within tight time windows

represents a coordination challenge for these carriers, which includes the operations of vans, trucks and aircrafts. Within these coordination efforts, one of the most important tasks is the tactical network design of their inter-city operations.

Network design problems arise when it is necessary either to install a set of facilities (e.g., pipelines or telecommunication antennas) or to determine a set of routes, so that they enable the transportation of a given amount of goods or information between specific points in a region. The demand between an origin and a destination is referred to as a *commodity*. The objective of the network design problem is to minimize both the cost of installing/operating the facilities/routes, and the handling and flow costs of the commodities through the network. The overnight service operated by the express integrators includes both ground and air operations. However, the cost of air operations represents, by far, the largest component of their operative costs. Thus, for the express integrators, defining the network design for the overnight deliveries is a critical task, known as the Express Shipment Service Network Design (ESSND) problem. The ESSND is a sub-class of the network design problem, because the express carriers must also consider additional characteristics.

The express integrators perform three processes to connect all the origins with all the destinations, (see Figure 1). The *pickup process* consolidates the packages, in stages, by transporting them from the origin customers to ground stations, then to airport facilities called gateways, and then to hubs. The hubs are large airport facilities that perform the *sorting process*, i.e. they receive the packages ordered by origin and sort them by destination. The *delivery process* deconsolidates the packages by transporting them from hubs to gateways, then to stations, and then to the final customer. Typically, the movements of packages are performed with vans between customers and stations, with trucks between stations and gateways, and with aircrafts between gateways and hubs. Normally, each vehicle is operated in both the pickup and the delivery processes and, eventually, they may also be repositioned to ensure the repeatability of the daily schedules of the routes. All the movements of a vehicle within either the pickup or the delivery processes, and each repositioning movement are called a *route*.

The routes have four main characteristics. First, they have to respect the time windows for loading and unloading the packages at hubs and gateways. Second, they have to observe the curfew and slot constraints imposed by the airports. Third, for each vehicle type, the number of units used in the network cannot surpass the number of vehicles available. Finally, to ensure the repeatability of the daily schedules of the routes, they must be selected such that, for each aircraft type, each gateway and hub must start and end with the same number of aircrafts. For the hubs, there are two elements to consider. First, their sorting capacity has to be respected. Second, if there is more than one hub within a region, the allocation of packages to hubs becomes a critical decision, known as the *hub assignment*. When the hub assignment is provided as an input to the ESSND, it is

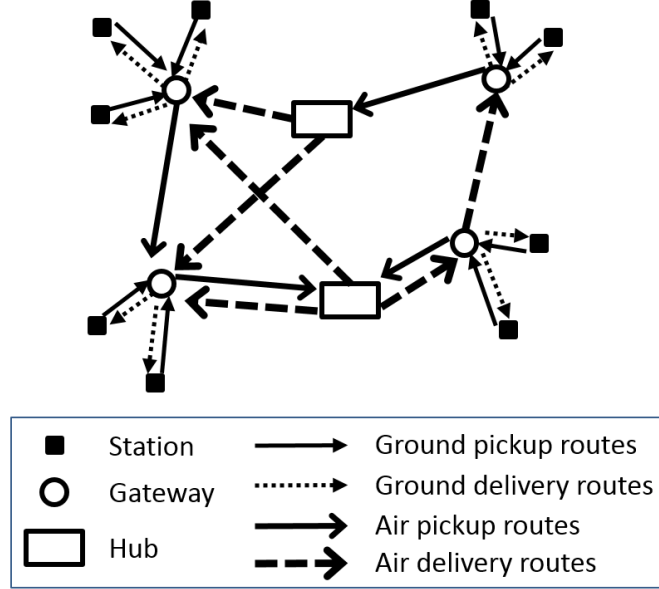


Figure 1 Representation of a network for the overnight express packages delivery processes.

referred to as a *fixed hub assignment*. If the hub assignment is an explicit decision of the ESSND, then it is called a *flexible hub assignment*. In this paper, one of our originalities is to consider a flexible hub assignment, while most of the latest solution approaches are based on the fixed hub assignment (e.g., Armacost, Barnhart, and Ware 2002, Fleuren et al. 2013, Louwerse et al. 2014, Meuffels 2015) to ease the tractability of large instances of the ESSND.

In this research, we develop an optimization model for addressing the Express Shipment Service Network Design problem, for the next day deliveries within a region, with multiple hubs and flexible hub assignment. Starting from an existing model, we first strengthen it with three families of valid inequalities: commodity connectivity inequalities, strong linking inequalities and strong cutset inequalities. Then, we use the route covers to reduce the number of variables and constraints. Our model is tested using extensive numerical experiments to show the value added in terms of efficiency and effectiveness. First, a comparison of our approach to the different formulations from the literature is presented on reduced size instances (20 gateways and 6 equipment types). The same instances allow us to analyze the contribution of each family of valid inequalities and of the route covers. Then, the performance of our model is analyzed for variations on the input data, such as the demand, the sorting capacity, the number of equipments, the distance between hubs and the number of gateways. Finally, our model is tested on realistic instances (76 gateways and 7 equipment types) within the European region. We make the following specific contributions. We develop an efficient method for solving the multi-hub ESSND with flexible hub assignment. We present two families of valid inequalities, the commodity connectivity inequalities, and the strong cutset inequalities, which strengthen the LP relaxation for the multi-hub ESSND. We introduce

the route covers which reduces the number of variables and constraints. Also, the efficiency of our method is shown with extensive numerical experiments.

All the developments presented are the result of a close research collaboration with FedEx Express Europe. The objective of this collaboration is the development of optimization models highly grounded in the industrial reality. FedEx Express Europe built all the testing instances we used for our numerical experiments so that they are representative of real life instances yet fictitious. Additionally, our industrial partner ran its own phase of testing that confirmed the results we achieved.

The remainder of the paper is structured as follows. A review of the related literature is exposed in Section 2. A detailed problem description and a modeling discussion is introduced in Section 3. In Section 4, we propose our Route and Hub formulation with Cuts and Covers. We discuss the results of our numerical experiments in Section 5. The conclusion is presented in Section 6.

2. Literature Review

Network design problems have been studied in many contexts, e.g., in telecommunications (Belotti, Malucelli, and Brunetta 2007), passenger airlines (Sherali, Bae, and Haouari 2013), and ground transportation (Crainic 2000). Surveys and reviews of the network design problem can be found, among others, in Agarwal (2002), Gendron, Crainic, and Frangioni (1999), Magnanti and Wong (1984), and Minoux (1989). The conventional formulation of the network design problem provides poor LP relaxations, and it is difficult to solve for large instances (Armacost, Barnhart, and Ware 2002, Gendron, Crainic, and Frangioni 1999). Many solution approaches have been proposed for improving the dual and primal bounds of this problem, such as cutting plane methods, Benders decomposition, Lagrangian relaxations, heuristic algorithms, approximation algorithms and hybrid methods. Reviews on the solution approaches for solving the network design problem are presented, by Shen (2004), and Wieberneit (2008).

The service network design problem, a sub-class of the network design problem, is faced by carriers that offer a high quality service in terms of speed, flexibility and reliability (Crainic 2000). Reviews of models and applications focused on the service network design problem are presented by Crainic (2000) and Wieberneit (2008).

The service network design problem faced by the express integrators has specific characteristics that make it particularly challenging (see Section 1) and is known as the Express Shipment Service Network Design problem. It has received attention from a limited number of researchers. Several of these contributions came 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). These researchers have focused on three main sub-classes of the ESSND:

single-hub models, multi-hub models with fixed hub assignment, and multi-hub models with flexible hub assignment.

In one of the first papers tackling the ESSND, Kuby and Gray (1993) developed an MIP formulation for a simplified version of the single-hub case which only models the flows of the commodities using capacitated aircrafts. Based on the FedEx case in the west coast of the US (34 gateways and 3 equipment types), they compare the cost-effectiveness between networks with one-leg routes against those with multi-leg routes and feeders, showing that the latter configuration allows more efficient network designs (73.6% less expensive). Barnhart and Schneur (1996) presented an uncapacitated single-hub ESSND model that incorporates the circulation, sorting, slots and counting constraints, but limits each gateway to be served by one aircraft type. Using a column generation technique, they solved a large instance of an express carrier in the US, uncovering potential savings in the order of tens of millions of dollars per year.

Regarding the multi-hub ESSND with fixed hub-assignment, Armacost, Barnhart, and Ware (2002) proposed the aircraft routing model. By using composite variables, they reduce the ESSND to a set covering formulation that has strong LP relaxations. The composite variables represent groups of routes carrying the total volume of a subset of gateways. The deployment of this approach to the UPS case in the US (101 gateways and 7 aircraft types) led to savings of tens of million dollars per year (Armacost et al. 2004). Different applications of the composite variables formulation have been presented for the generic (air or ground) multi-hub ESSND with flexible hub assignment. With minor adjustments on the sorting and counting constraints, Fleuren et al. (2013) report using the approach with savings of 132 million dollars at TNT between 2008 and 2009. Louwerse et al. (2014) use the composite variables in the development of a heuristic algorithm for solving generic ESSND problems with multiple service types. They combine iteratively a column generation technique and a search algorithm to find primal solutions. They test modified instances from two regions of an express carrier with realistic size (44 ramps and 2 equipment types; 60 ramps and 2 equipment types), getting up to 18.6% cost savings compared to the solutions provided by the carrier. Using the composite variables, Meuffels (2015) develops a heuristic algorithm for the hub decongestion of a dual-hub case. With this approach, Meuffels (2015) reports bypassing 16% of the demand from the sorting process, for an instance built in collaboration with an express integrator in Europe (40 ramps and 2 equipment types).

The multi-hub ESSND problem with flexible hub assignments integrates the hub assignment, the flows of the commodities and the selection of routes in a single problem. Therefore, it is more challenging to solve, and exact approaches have only been studied in two papers, to the extent of our knowledge. Nevertheless, good solution methods for solving this problem are expected to generate more benefits compared to those obtained with fixed hub assignment methods. There are

two main models for addressing the ESSND with flexible hub assignment. First, the route and hub model, proposed by Kim et al. (1999), is based on the classic multi-commodity network design formulation, as it tracks the flow of the commodities through each arc. To solve large instances, the authors propose a heuristic algorithm with three main iterative steps. First, it solves a reduced model, then it generates lifted cutset inequalities, and finally it generates columns. With three data sets provided by an express integrator (31 gateways and 5 aircraft types; 91 gateways and 7 aircraft types; 141 ramps and 7 aircraft types), Kim et al. (1999) compared the performance of their algorithm against the solutions provided by the company’s planners, obtaining average savings of 10%. The second model addressing the ESSND with flexible hub assignment is the route and hub model, introduced by Shen (2004). It tracks the commodities only at hubs and it considers aggregated flows everywhere else. Therefore, it reduces significantly the number of variables and constraints compared to the route and flow model. Shen (2004) also proposed a model extension called gateways cover flow formulation that includes the composite variables as cover cuts for improving the LP relaxation, and developed a column generation technique for solving it. Applying the gateway cover flow formulation to the case of UPS in the US (85 gateways and 4 aircraft types), the numerical experiments for this algorithm showed a slow convergence for finding primal solutions. Nevertheless, by fixing the hub assignment of 60% of the demand and letting a flexible flow assignment for the remaining demand, Shen (2004) reports theoretical savings of 7.5% compared to the solution obtained by UPS using the aircraft routing model (Armacost, Barnhart, and Ware 2002). This percentage accounts for savings of tens of millions of dollars per year.

In this paper, to strengthen the LP relaxation of the model and ease its tractability, we add three families of valid inequalities: commodity connectivity, strong linking and strong cutset inequalities. Additionally, by introducing the route covers, we reduce the number of variables and constraints required for formulating the multi-hub ESSND with flexible hub assignment. Our approach allows us to efficiently solve realistic instances with 76 gateways and 7 vehicle types.

3. Problem description and modeling discussion

In this section, a detailed description of the multi-hub Express Shipment Service Network Design problem with flexible hub assignment is presented (Section 3.1). Then, the typical two-stage modeling approach for addressing it is discussed. In the first stage, an algorithm for generating the set of feasible routes is executed (Section 3.2). In the second stage, using the routes of the previous stage as an input, a mathematical formulation is employed to solve the problem (Section 3.3). The second stage is illustrated by presenting the Route and Hub model (RH), introduced by Shen (2004), which is the starting point of our developments.

3.1. Problem description

The companies facing the multi-hub ESSND problem need to decide a set of routes to be operated every day that enables the flow of their demands from their origins to their destinations, as efficiently as possible, overnight and via intermediary hubs. Due to the costs of flight operations, it is key to use the capacity of the aircrafts efficiently, notably by using multi-stop flights when possible. However, guaranteeing overnight deliveries is a key component of the service proposition of express companies, thus, delivering the packages on time might imply low utilization rates for certain aircrafts.

In general, a *route* is the sequence of movements performed by a vehicle, and each movement is called a *leg*. However, a route cannot contain legs from both the pickup and the delivery processes. Therefore, the routes are classified as *pickup routes* and *delivery routes* according to the process in which they are operated. The route characteristics and route generation process are explained in Section 3.2.

Due to the limited time for performing the pickup and delivery processes, the packages are routed from their origin to their destination as follows. Each package needs to be loaded into a single pickup route at its origin, which hauls the package to a hub. At the hub, the package is sorted and then loaded into a single delivery route, which hauls the package to its destination. The packages from a commodity can be separated and loaded into different routes, and they can be sorted at different hubs, i.e., they may have a different hub assignment.

The features of the facilities and of the vehicles characterize the ESSND as a sub-class of the service network design problem. The operative constraints of the facilities are the following. The gateways have a *pickup release time*, minimum time at which the commodities can be loaded into a pickup route. They also have a *delivery due time*, maximum time at which they receive delivery routes and that ensures the on time delivery of the commodities. The hubs require a time window for performing the sorting process, defined by the window between the *pickup due time*, which is the maximum time at which they receive pickup routes, and the *delivery release time*, which is the time at which the sorting process has finished and the delivery routes can begin their operations. Since hubs have a sorting speed, this time window determines their *sorting capacity*, the maximum number of packages they can receive from the pickup process. The airports hosting the hubs and gateways also establish curfews banning the landings and departures of aircrafts. We call *timing conditions* of the facilities to all the curfew, due times and release times they impose. A final limitation of the airports hosting hubs is the *slots constraint*, the maximum number of take-offs and landings that a company can perform at that airport.

The vehicles also have some operative constraints. First, they have a capacity, the maximum weight they can carry between two nodes. Second, they have a travel time that determines if they

can respect the timing conditions of two facilities when they travel between them. Additionally, each time the vehicles visits a node, they must remain there for a minimum *stop time* before departing, which allows among other operations, the loading/unloading of packages and refueling. Finally, if the vehicles are aircrafts, they have a maximum flying range without fuel refill.

3.2. Route characteristics and route generation process

The route generation process is executed as the first stage for modeling the ESSND problem. It computes, by enumeration, all the routes that can be operated feasibly for all the combinations of equipment types, hubs, and subsets of gateways with up to three elements. A mathematical description of the route generation process can be found in Kim et al. (1999). The route generation reduces significantly the formulation complexity of the ESSND problem, because for each route, it determines the nodes it visits, it fixes its departure and arrival times at each node, and it calculates its total operating cost. Therefore, the route generation process allows the routes to be modeled with discrete variables in the second stage of the solution approach, i.e. decide how many times each route is operated.

There are three validity conditions of the routes for the ESSND problem. First, considering the routes individually, each one must satisfy the timing conditions of the facilities it visits and must respect the operative constraints of the vehicle with which it is operated. Second, considering the routes globally, they must allow the daily repeatability of the schedule, i.e. they must be selected such that, for each aircraft type, each gateway and hub must start and end with the same number of aircrafts. This characteristic, known as the *circulation problem*, makes the ESSND particularly challenging compared to other multi-commodity network design problems (Chouman and Crainic 2014, Pedersen, Crainic, and Madsen 2009). Third, for each aircraft type, the number of units employed on the network design cannot exceed their availability. This validity condition may be added to the formulation or not, depending on the modeling objectives, e.g., if the objective is to study the optimal fleet composition, this constraint can be omitted.

The route generation process only produces feasible pickup/delivery routes that end/start at a hub. Since only feasible routes are generated, a package loaded into a route will also comply with the timing conditions of the facilities. As a result, the flow decisions in the mathematical formulation can be modeled by assigning each package to a single pickup route and to a single delivery route, and by ensuring the flow conservation at the hubs.

The cost of the air routes is composed by the variable, cycle and ownership costs. The variable costs are incurred for each hour flown. The cycle costs are related to each take off-landing cycle, and are, for example, due to airport fees and tires wear and tear. The ownership costs are the daily depreciation or leasing costs of the vehicle. The cost of the ground routes is so low compared to the

cost of any air route, that it is assumed to be free. Regarding the package handling costs, they are mainly determined by the installation and fixed operation costs of the hubs and gateways, which for this problem are both assumed to be fixed, and can thus be omitted.

The types of routes that are produced with our route generation process are the following.

One-leg routes are flown by an aircraft directly from a gateway/hub to a hub/gateway.

Multi-leg routes are flown by an aircraft from a gateway/hub to a hub/gateway, with intermediate stops at one or more gateways. For each in-between landing, there exists a minimum stop time before the aircraft departs again. Due to the timing conditions, only two-leg and three-leg routes are generated.

Ferry routes are flown by an empty aircraft between two nodes after the end of the delivery process.

Their objective is to reposition the aircrafts so as to comply with the circulation problem. They are not subject to timing conditions since they fly during the day and have enough time to complete their movements. They belong neither to the pickup nor the delivery processes.

Ground routes are performed by a truck traveling from a gateway/hub to a hub/gateway. They are assumed to be free and uncapacitated because their operative cost is much cheaper than those of any air route.

3.3. Route and Hub formulation

Two main equivalent formulations exist for the multi-hub ESSND with flexible hub assignment. First, the route and flow formulation, proposed by Kim et al. (1999). It is based on the classical multi-commodity network design formulation (Magnanti and Wong 1984, Crainic and Rousseau 1986, Minoux 1989) which tracks the commodities throughout the network. Second, the route and hub (RH) formulation, introduced by Shen (2004), which considers aggregated flows over the routes and only tracks the commodities at the hubs. The RH formulation reduces significantly the number of variables and constraints compared to the route and flow formulation.

In this paper we use the RH formulation as a starting point. Before presenting this formulation, we introduce the notation used throughout the paper.

Notation

F	is the set of fleet types, indexed by f , including aircrafts and trucks.
H	is the set of hubs, indexed by h .
K	is the set of commodities, i.e. each origin-destination pair with positive demand, indexed by k .
R	is the set of routes, indexed by r .
N	is the set of gateways, indexed by n .
$N_P \subset N$	is the set of gateways in the pickup process.
$N_D \subset N$	is the set of gateways in the delivery process.
$N^r \subset N$	is the set of gateways served by route r .
$N_P^r \subseteq N_P$	is the set of gateways served by route r in the pickup process.
$N_D^r \subseteq N_D$	is the set of gateways served by route r in the delivery process.
$K_O^n \subset K$	is the set of commodities whose origin is the gateway n .
$K_D^n \subset K$	is the set of commodities whose destination is the gateway n .
$R^f \subset R$	is the set of routes operated with fleet type f .
$R_h^f \subset R$	is the set of routes operated with fleet type f that land or take-off from h .
$R_P \subset R$	is the set of pickup routes.
$R_D \subset R$	is the set of delivery routes.
$R_P^f \subset R$	is the set of pickup routes operated with fleet type f .
$R_P^S \subset R$	is the set of pickup routes that serve the subset of gateways $S \subseteq N_P$.
$R_D^S \subset R$	is the set of delivery routes that serve the subset of gateways $S \subseteq N_D$.
$R_P^{n,h} \subset R$	is the set of pickup routes that serve gateway n and end at hub h .
$R_D^{h,n} \subset R$	is the set of delivery routes that start at hub h and serve gateway n .
$R_c \subseteq R$	is the set of routes that are a cover.
$R_{P,c}^{n,h} \subseteq R_c$	is the set of pickup routes that are a cover for ramp n and end at hub h .
$R_{D,c}^{h,n} \subseteq R_c$	is the set of delivery routes that are a cover for ramp n and start at hub h .

Parameters

a_h	is the number of take-off and landing slots available at hub h .
b^k	is the demand for commodity k in pounds.
b_P^n	is the pickup demand for gateway n in pounds, i.e. the sum of the demands of all the commodities originated in n .
b_D^n	is the delivery demand for gateway n in pounds, i.e. the sum of the demands of all the commodities with destination to n .
b_P^S	is the pickup demand of gateways $S \subseteq N_P$ with destination N_D in pounds.
b_D^S	is the delivery demand of gateways $S \subseteq N_D$ originated in N_P in pounds.
c_r	is the cost of route r in dollars.
e_h	is the sorting capacity of hub h in pounds per night.
m^f	is the available number of units of fleet type f .
u_r	is the capacity of route r in pounds.
$\mathcal{O}(r)$	is the origin gateway of route r .
$\mathcal{D}(r)$	is the destination gateway of route r .
$\mathcal{O}(k)$	is the origin gateway of commodity k .
$\mathcal{D}(k)$	is the destination gateway of commodity k .
φ_r^n	1 if route r serves gateway n , 0 otherwise.

Decision variables

x_r^n is the weight from gateway n shipped on route r in pounds, referred to as flow variables.

y_r is the number of times that route r is operated, referred to as design variables.

z_h^k is the weight from commodity k assigned to hub h in pounds, referred to as hub assignment variables.

The RH formulation, introduced by Shen (2004), is given by the following equations.

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

$$\text{s.t.} \quad \sum_{h \in H} z_h^k = b^k, \quad \forall k \in K, \quad (2)$$

$$\sum_{r \in R_P^{n,h}} x_r^n - \sum_{k \in K_{\mathcal{O}}^n} z_h^k = 0, \quad \forall n \in N_P, h \in H, \quad (3)$$

$$\sum_{r \in R_D^{h,n}} x_r^n - \sum_{k \in K_{\mathcal{D}}^n} z_h^k = 0, \quad \forall n \in N_D, h \in H, \quad (4)$$

$$\sum_{n \in N^r} x_r^n \leq u_r y_r, \quad \forall r \in R, \quad (5)$$

$$\sum_{r \in R^f: \mathcal{O}(r)=n} y_r - \sum_{r \in R^f: \mathcal{D}(r)=n} y_r = 0, \quad \forall n \in N, f \in F, \quad (6)$$

$$\sum_{r \in R^f: \mathcal{O}(r)=h} y_r - \sum_{r \in R^f: \mathcal{D}(r)=h} y_r = 0, \quad \forall h \in H, f \in F, \quad (7)$$

$$\sum_{r \in R_P^f} y_r \leq m^f, \quad \forall f \in F, \quad (8)$$

$$\sum_{f \in F} \sum_{r \in R_h^f} y_r \leq a_h, \quad \forall h \in H, \quad (9)$$

$$\sum_{k \in K} z_h^k \leq e_h, \quad \forall h \in H, \quad (10)$$

$$y_r \in \mathbb{Z}_+, \quad \forall r \in R, \quad (11)$$

$$x_r^n, z_h^k \in \mathbb{R}_+, \quad \forall r \in R, n \in N, h \in H, k \in K. \quad (12)$$

The objective function is represented in (1), minimizing the total cost of the operated routes. Equations (2)-(4) are the *forcing* constraints, they enforce the flow of the commodities from their origins to their destinations and preserve the flow integrity throughout the network. Specifically, equation (2) ensures the *hub assignment* of the commodities, and equations (3)-(4) assign aggregated commodity demands to routes so that they match the hub assignment for the delivery and pickup processes respectively. Equation (5) is the *capacity* constraint, it limits the weight each vehicle can carry. It is also referred to as *linking* or *bundle* constraint, as it associates the design

and flow variables. Equations (6)-(7) model the *circulation problem* for the ramps and hubs respectively. For each type of vehicle, they force each node to start and end the operation process with the same number of units. Equation (8) is the *counting* constraint for each vehicle type, it limits the number of vehicles operated to the amount available. Equation (9) is the *slots* constraint, it limits the number of take-offs and landings at each hub. Equation (10) is the *sorting* capacity constraint for each hub. In this paper, we assume this constraint is sufficient to ensure the sorting feasibility. If required, more advanced sorting capacity constraints can be built using the *staggering* constraints, presented by Kim et al. (1999). Finally, equations (11)-(12) define the domain of the variables.

4. Strengthening of the Route and Hub model

In this section, we propose several improvements to the RH formulation presented in Section 3.3, in order to eventually be able to solve the multi-hub ESSND more efficiently. The RH formulation inherits some solving difficulties from the classical multi-commodity network design formulation, notably a poor LP relaxation (Gendron and Crainic 1994, Crainic, Frangioni, and Gendron 2001, Armacost, Barnhart, and Ware 2002) caused by the capacity constraints (5). To improve the resolution of the RH model, we add three families of valid inequalities: the commodity connectivity inequalities (Section 4.1), and the strong linking and strong cutset inequalities (Section 4.2), which help improving the LP relaxation of the problem and ease its tractability in the branching process. Then, we introduce the *route covers* (Section 4.3) to reformulate the forcing constraints (3)-(4), thus reducing the number of flow variables and capacity constraints. The section is closed with the introduction of our complete formulation, the Route and Hub formulation with Cuts and Covers (Section 4.4).

4.1. Commodity Connectivity Inequalities

The cover inequalities have been used to strengthen the LP relaxation of multi-commodity network design problems (Chouman, Crainic, and Gendron 2003, Barnhart et al. 2002). A basic cover inequality states that each pickup/delivery gateway must be served by at least one pickup/delivery route. It can be formulated as follows.

$$\sum_{h \in H} \sum_{r \in R_P^{n,h}} y_r \geq 1, \quad \forall n \in N_P, \quad (13)$$

$$\sum_{h \in H} \sum_{r \in R_D^{h,n}} y_r \geq 1, \quad \forall n \in N_D. \quad (14)$$

These cover inequalities consider all the routes serving a gateway regardless of the destination/origin hub of the routes. However, if the demand of a commodity is assigned to a specific

hub, there must exist a route connecting that hub to the origin/destination of the commodity. Therefore, we can disaggregate equations (13)-(14) by concentrating on the relationship between the hub assignment variables z_h^k with their commodity demands b^k as follows.

$$\sum_{r \in R_P^{n,h}} y_r \geq \frac{z_h^k}{b^k}, \quad \forall n \in N_P, k \in K : \mathcal{O}(k) = n, h \in H, \quad (15)$$

$$\sum_{r \in R_D^{h,n}} y_r \geq \frac{z_h^k}{b^k}, \quad \forall n \in N_D, k \in K : \mathcal{D}(k) = n, h \in H. \quad (16)$$

We call equations (15)-(16) the Commodity Connectivity Inequalities (CCI), since they express the connectivity requirements of commodities through hubs. In conjunction with the hub assignment constraint (2), the CCI imply that there must exist a valid origin-hub-destination path for each commodity. To the best of our knowledge, the commodity connectivity inequalities are new to the literature for the multi-hub ESSND problem.

4.2. Strong Linking and strong cutset inequalities

A known effective way for improving the LP relaxation of multi commodity network design models is to add the Strong Linking Inequalities (SLI) (Chouman and Crainic 2014). The SLI can be stated as follows.

$$\min(b_P^n, u_r) y_r \geq x_r^n, \quad \forall r \in R_P, n \in N_P^r, \quad (17)$$

$$\min(b_D^n, u_r) y_r \geq x_r^n, \quad \forall r \in R_D, n \in N_D^r. \quad (18)$$

For routes with idle capacity, the SLI produce tighter bounds than the capacity constraints (5), and thus the SLI yield better LP relaxations when such routes are part of the LP solution. The SLI are added to the model for each route and each gateway served by the route.

Furthermore, both the routes and gateways can be aggregated to obtain the Strong Cutset Inequalities (SCI). Some graph theory concepts and notations are required to present these inequalities. In a digraph $G = (R, N)$, a cutset is a pair of non-empty subsets of nodes $(S, \bar{S})$, such that $S \subset N$ and $\bar{S} = N \setminus S$. The cutset is served by $R_{(S, \bar{S})}$, the set of routes that travel from S to $\bar{S}$. If $b_{(S, \bar{S})}$ is the positive demand originated in S with destination to $\bar{S}$, the basic cutset inequalities are defined as follows.

$$\sum_{r \in R_{(S, \bar{S})}} u_r y_r \geq b_{(S, \bar{S})}, \quad \forall S \subset N. \quad (19)$$

These inequalities enforce that, for any cutset $(S, \bar{S})$, its serving routes $R_{(S, \bar{S})}$ have to provide enough capacity for hauling the demand $b_{(S, \bar{S})}$ from S to $\bar{S}$. The basic cutset inequalities do not improve the LP relaxation of the ESSND (Chouman, Crainic, and Gendron 2003). However, for achieving this objective, they can be either lifted (Barnhart et al. 2002) or used for generating cover cuts (Chouman, Crainic, and Gendron 2003). For the latter case, if the basic cutset inequalities can be strengthened, the correspondent cover cuts will also be stronger.

For the ESSND, the cutset inequalities can be written independently for the pickup and delivery processes, since each of them is modeled as a sub-network. In the following explanation, we focus on the pickup process, but the same logic applies for the delivery process. In the pickup process, the hubs are the only destinations of the sub-network, and all the routes end at a hub. Therefore, any selection of nodes such that $S \subseteq N_P$ and $\bar{S} = H$ can be used for generating a cutset inequality ($N_P \setminus S$ cannot be destinations). Moreover, with the same logic used for generating the SLI, the idle capacity of the routes in a cutset can be eliminated to get strengthened cutset inequalities. Let φ_r^n be 1 if the route r serves gateway n , 0 otherwise; R_P^S the set of routes serving $S \subseteq N_P$, and R_D^S the set of routes serving $S \subseteq N_D$, then the strong cutset inequalities are as follows.

$$\sum_{r \in R_P^S} \min\left(\sum_{n \in S} \varphi_r^n b_P^n, u_r\right) y_r \geq b_P^S, \quad \forall S \subseteq N_P, \quad (20)$$

$$\sum_{r \in R_D^S} \min\left(\sum_{n \in S} \varphi_r^n b_D^n, u_r\right) y_r \geq b_D^S, \quad \forall S \subseteq N_D. \quad (21)$$

The strong cutset inequalities are lifted cutset inequalities, and thus they improve the LP relaxation when they are added to the RH formulation. However, the SLI (17)-(18) are stronger since they are a disaggregated version of the SCI (20)-(21). Nevertheless, even when the SLI (17)-(18) are present, the SCI (20)-(21) can be used to improve the LP relaxation since their knapsack structure allows to generate lifted cover cuts in a branch and cut algorithm (and modern commercial softwares have built-in functions doing this, also see Section 5.2). To the best of our knowledge, the strong cutset inequalities are new to the literature for the ESSND problem.

4.3. Route Covers

Based on the concept of minimal covers, Armacost, Barnhart, and Ware (2002) developed a set partitioning formulation for addressing the ESSND with fixed hub assignment. Their formulation eliminates the flow variables and the capacity constraint (5) from the RH, and yields stronger LP relaxations. However, their approach cannot be used for the case of flexible hub assignment. To improve the tractability of the latter case, we introduce the *route covers* (RC) which reduce the number of flow variables x_r^n and capacity constraints (5) from the RH.

A route cover for the pickup/delivery process is defined as a route whose capacity is larger than the total pickup/delivery demand of the gateways it serves. Thus, their capacity constraint (5) can be omitted. Yet, their design variables y_r still need to be linked to their correspondent flows, so that their costs can be accounted. This can be achieved by replacing the flow variables x_r^n in the forcing equations (3)-(4) by the design variables y_r , and multiplying them by the total pickup/delivery demand of the gateway served. Additionally, both equations need to be changed from equalities to inequalities so that, if a route cover is selected, the commodities of the gateways it serves can still be split to different hubs. As a result, the commodity flows are controlled by the hub assignment variables z_h^k , while the routes may carry a theoretical oversupply. This does not affect the flow feasibility, since the forcing equations (3)-(4) still ensure that the demand is satisfied.

The Route and Hub formulation with route covers is obtained by replacing the equations (3)-(5) and (12) respectively with the next equations

$$\sum_{r \in R_{P,c}^{n,h}} b_P^n y_r + \sum_{r \in R_P^{n,h} \setminus R_c} x_r^n - \sum_{k \in K_O^n} z_h^k \geq 0, \quad \forall n \in N_P, h \in H, \quad (22)$$

$$\sum_{r \in R_{D,c}^{h,n}} b_D^n y_r + \sum_{r \in R_D^{h,n} \setminus R_c} x_r^n - \sum_{k \in K_D^n} z_h^k \geq 0, \quad \forall n \in N_D, h \in H, \quad (23)$$

$$\sum_{n \in N^r} x_r^n \leq u_r y_r, \quad \forall r \in R \setminus R_c, \quad (24)$$

$$x_r^n, z_h^k \in \mathbb{R}_+, \quad \forall r \in R \setminus R_c, n \in N, h \in H, k \in K, \quad (25)$$

where $R_c \subseteq R$ is the set of routes that are route covers, $R_{P,c}^{n,h} \subseteq R_c$ the set of pickup routes that are route covers for gateway n and end at hub h , $R_{D,c}^{h,n} \subseteq R_c$ the set of delivery routes that are route covers for ramp n and start at hub h , b_P^n the pickup demand for gateway n in pounds, and b_D^n the delivery demand for gateway n in pounds.

Since the capacity constraints (5) eliminated in the RHCC can never be violated in the RH, and any flow in the RH can be directly transformed into a flow in the RHCC and vice versa, the RH and RHCC are equivalent formulations. Typically, the proportion of routes that are route covers for the ESSND is not negligible. To the best of our knowledge, the route covers are new to the literature for the ESSND problem.

4.4. Route and Hub model with Cuts and Covers

Finally, we present our complete formulation, the Route and Hub model with Cuts and Covers (RHCC).

$$\min \sum_{r \in R} c_r y_r \quad (26)$$

$$\text{s.t.} \quad \sum_{h \in H} z_h^k \geq b^k, \quad \forall k \in K, \quad (27)$$

$$\sum_{r \in R_{P,c}^{n,h}} b_P^n y_r + \sum_{r \in R_P^{n,h} \setminus R_c} x_r^n - \sum_{k \in K_{\mathcal{O}}^n} z_h^k \geq 0, \quad \forall n \in N_P, h \in H, \quad (28)$$

$$\sum_{r \in R_{D,c}^{h,n}} b_D^n y_r + \sum_{r \in R_D^{h,n} \setminus R_c} x_r^n - \sum_{k \in K_{\mathcal{D}}^n} z_h^k \geq 0, \quad \forall n \in N_D, h \in H, \quad (29)$$

$$\sum_{n \in N^r} x_r^n \leq u_r y_r, \quad \forall r \in R \setminus R_c, \quad (30)$$

$$\sum_{r \in R^f: \mathcal{O}(r)=n} y_r - \sum_{r \in R^f: \mathcal{D}(r)=n} y_r = 0, \quad \forall n \in N, f \in F, \quad (31)$$

$$\sum_{r \in R^f: \mathcal{O}(r)=h} y_r - \sum_{r \in R^f: \mathcal{D}(r)=h} y_r = 0, \quad \forall h \in H, f \in F, \quad (32)$$

$$\sum_{r \in R_P^f} y_r \leq m^f, \quad \forall f \in F, \quad (33)$$

$$\sum_{f \in F} \sum_{r \in R_h^f} y_r \leq a_h, \quad \forall h \in H, \quad (34)$$

$$\sum_{k \in K} z_h^k \leq e_h, \quad \forall h \in H, \quad (35)$$

$$\sum_{r \in R_P^{n,h}} y_r \geq \frac{z_h^k}{b^k}, \quad \forall n \in N_P, k \in K: \mathcal{O}(k)=n, h \in H, \quad (36)$$

$$\sum_{r \in R_D^{h,n}} y_r \geq \frac{z_h^k}{b^k}, \quad \forall n \in N_D, k \in K: \mathcal{D}(k)=n, h \in H, \quad (37)$$

$$\min(b_P^n, u_r) y_r \geq x_r^n, \quad \forall r \in R_P, n \in N_P^r, \quad (38)$$

$$\min(b_D^n, u_r) y_r \geq x_r^n, \quad \forall r \in R_D, n \in N_D^r, \quad (39)$$

$$\sum_{r \in R_P^S} \min(\sum_{n \in S} \varphi_r^n b_P^n, u_r) y_r \geq b_P^S, \quad \forall S \subseteq N_P, \quad (40)$$

$$\sum_{r \in R_D^S} \min(\sum_{n \in S} \varphi_r^n b_D^n, u_r) y_r \geq b_D^S, \quad \forall S \subseteq N_D, \quad (41)$$

$$y_r \in \mathbb{Z}_+, \quad \forall r \in R, \quad (42)$$

$$x_r^n, z_h^k \in \mathbb{R}_+, \quad \forall r \in R \setminus R_c, n \in N, h \in H, k \in K. \quad (43)$$

The objective (26) aims to minimize the total cost of the routes selected. It is not changed compared to the RH model. Inequalities (27)–(30) are the forcing and the capacity constraints. They now incorporate the route covers which reduce the number of capacity constraints (30) (see Section 4.3). Also notice that equations (27)–(29) are now inequalities (rather than equalities in (3)–(4)). Constraints (31)–(32) are not modified, they model the circulation problem. Inequalities (33)–(35) account for the counting, slots and sorting constraints. They remain unchanged. The commodity connectivity inequalities are included with (36)–(37) (see Section 4.1). Constraints

(38)–(41) add the strong linking and the strong cutset inequalities (see Sections 4.2). Finally, constraints (42)–(43) define the domain of the variables. Note that the number of flow variables x_r^n is reduced thanks to the route covers (see Section 4.3).

5. Computational Experiments

In this section, we present our computational experiments for testing the performance of the Route and Hub model with Cuts and Covers. Each instance was built by modifying some characteristics of a scenario with 2 hubs, 76 gateways and 7 equipment types for the European market. Multi-hub ESSND problems are particularly challenging for this region, because hubs are located at short distances between each other and each can serve almost all the gateways on time. This increases the combinatorial number of potential solutions. Three types of experiments were performed. First, in Section 5.1, we compare the solutions obtained by state of the art models and by the RHCC, with 20 cases of 20 ramps. These cases also allow us to analyze the impact of the valid inequalities we propose and of the route covers. Then, in Section 5.2, to show the reliability of our model the influence of various input parameters is studied: demands, sorting capacities, hub locations, number of equipment types, number of ramps and number of hubs. Finally, to illustrate the benefits of our approach on realistic size instances, Section 5.3 presents solutions of the RHCC on the full dataset.

The default characteristics of the experiments are the following. The solutions were calculated using Java/Gurobi 6.0.0, with a processor at 2.70 GHz and 16 GB in RAM. The solver was allowed to run for up to two hours using 3 threads out of 4. Two Gurobi configurations were defined. Configuration *c1* focuses on improving the lower bounds (the user defined parameters are: MIPfocus = 3, CutPasses = 5, SubMIP Cuts = 2), and configuration *c2* focuses on finding primal solutions (the user defined parameters are: MIPfocus = 1, SubMIPNodes = 1500, Heuristics = 0.95). Configuration *c1* is used for up to 110 minutes, and configuration *c2* for the remaining time. Regarding the route generation, the same algorithm is used for all the experiments, and loads 1-leg, 2-leg, 3-leg and ferry routes. About the strong cutset inequalities (see Section 4.2), they are included for $|S| \leq 2$ and, for $|S| = 2$, they are only generated for pairs of gateways that share at least one route.

5.1. Base cases

In a first series of experiments, we solve instances with reasonable size, i.e. 20 gateways, in order to compare our results with those of several other approaches. The size of these instances is similar to those from other works addressing the multi-commodity network design problem (e.g., Chouman and Crainic 2014, Crainic, Gendron, and Hernu 2004, Gendron, Crainic, and Frangioni 1999). Moreover, the instances are designed to capture most of the operative complexities of a real size case. To achieve this, we base the experiments on a realistic network in Europe with 76 gateways and

make two modifications (see Figure 2). First, conic regions containing 20 gateways are determined, so that there is a high proportion of multi-legs routes. Second, an additional fictitious gateway is included, linked to both hubs with a single free uncapacitated route. This gateway accounts for all the demands of the gateways that were not included in the region, which allows realistic demands and aircrafts for the 20 ramps. Four such regions have been defined, referred to as maps (see Figure 2). Table 1 details the general characteristics of the maps in terms of set sizes and route compositions. For each map, five scenarios have been defined by limiting the sorting capacities of the hubs to the following demand proportions respectively: 80%-20%, 65%-35%, 50%-50%, 35%-65% and 20%-80%. This leads to 20 cases for our experiments, that we call base cases.

We use these base cases to compare the performance of our model, the Route and Hub model with Cuts and Covers (RHCC), against the main approaches from the literature (see Section 2): the Route and Flow (RF) model introduced by Kim et al. (1999), the Route and Hub (RH) model proposed by Shen (2004), and the Gateway Cover Flow (GFC) formulation presented by Shen (2004). Additionally, to show the contribution of the valid inequalities (CCI, SLI and SCI) and of

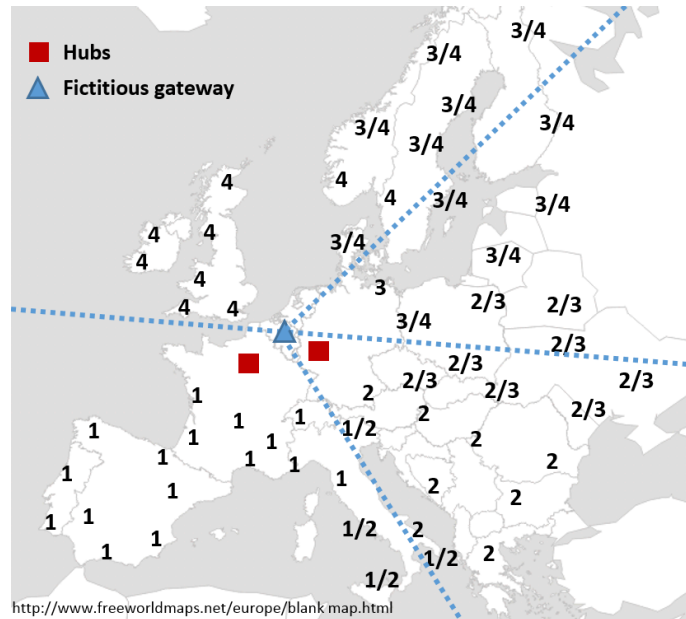


Figure 2 Exemplification of the maps' regions. The numbers give the gateways selected for each map.

Table 1 Statistics and routes compositions of the four maps.

	Maps statistics				Number of routes				
	$ N $	$ K $	$ H $	$ F $	1-leg	2-legs	3-legs	Ferry	Total
Map 1	21	348	2	6	409	990	8	2,696	4,103
Map 2	21	390	2	6	426	1,666	248	2,736	5,076
Map 3	21	392	2	6	439	1,805	296	2,714	5,254
Map 4	21	362	2	6	437	1,593	151	2,640	4,821

the route covers (RC), we also show the performance of the RF model when the three families of valid inequalities are included (RFC), and of the RHCC reduced models, i.e. when each of the valid inequalities or the route covers are removed independently. To refer to each reduced model, we use the notation RHCC-X (e.g. RHCC-CCI) where X is the removed feature (e.g the Commodity Connectivity Inequalities). All the models are solved directly with Gurobi except for the GFC which is solved with a column generation algorithm.

Table 2 displays the performance, gap and cost, of the RF, RH, RFC and RHCC models, as well as of the RHCC reduced models. The cost performance is measured as $(MBP/BP - 1)$, where MBP is the model's best primal solution, and BP is the best primal solution among the models for the same instance. We first compare the performance of the RHCC model against the RF and RH models. In terms of optimality gap, it appears clearly that the RHCC (1.24%) largely outperforms the RH (12.6%) and, even more, the RF (53.6%). Moreover, the RHCC obtained 8 optimal solutions, while the RF and RH models did not obtain any. The RHCC also obtained the best cost performance, improving by more than 5.2% the RH and by more than 78% the RF. The RHCC thus gets largely better lower bounds and primal solutions.

The RFC shows the impact of the valid inequalities even when they are added to the RF, achieving an improvement of almost 40% in the optimality gap and of almost 70% in the cost. However, the performance of the RFC is still worse than that of the RH (and thus of the RH with the valid inequalities) both in optimality gap and costs.

We may conclude from Table 2 that the three valid inequalities contribute to the performance improvement, as removing any one of them deteriorates the performance. We see that the worst results are obtained when the CCI are removed (RHCC-CCI), lowering the optimality gap by 2.6%. Even though the SLI yield stronger LP relaxations than the SCI, the final results are slightly worse when the SCI are not present (RHCC-SCI) than when the SLI are missing (RHCC-SLI). As mentioned in Section 4.2, Gurobi exploits the knapsack structure of the SCI for generating cover

**Table 2 Average optimality gap and cost performance,
in percent, for the base cases with different models.**

	Optimality gap (%)							
	RF	RH	RFC	RHCC				full model
				-CCI	-SLI	-SCI	-RC	
Map 1	52.2	7.66	4.70	0.63	0.0	0.0	0.0	0.0
Map 2	54.0	15.0	16.9	5.87	3.14	3.69	2.15	1.78
Map 3	55.2	14.1	16.6	4.77	3.45	3.72	2.25	2.28
Map 4	52.9	13.8	17.6	4.08	2.34	2.65	0.91	0.91
Average optimality gap (%)	53.6	12.6	13.9	3.84	2.23	2.52	1.33	1.24
Average cost performance (%)	78.7	5.44	10.2	1.36	0.77	0.84	0.13	0.16

cuts (e.g., it generated on average almost 32 times more cover cuts with the RHCC-SLI than with the RHCC-SCI). For the base cases, the difference between the RHCC-RC and the RHCC is not significant, with less than 0.1% for both the average optimality gap and cost performance. This will not be true for larger size instances (see Section 5.3).

To further test the efficiency of our formulation, we compare it with the GFC column generation algorithm proposed by Shen (2004) which, to the best of our knowledge, is the most recent algorithm from the literature for solving the multi-hub ESSND with flexible hub assignment. Since we do not share the same testing instances, we adjust some characteristics of the algorithm as follows. The maximum running time for the column generation is set to two hours, and the minimum improvement between two consecutive reduced master problems to .0001%. For the final master problem, we established a maximum running time of 10 minutes, which was never reached. The GFC also requires an initial solution for starting the column generation process. We tested three different sets of initial solutions provided by our model, the RHCC, when it runs for 5, 20 and 60 seconds, using a Gurobi configuration *c3*. This configuration finds faster solutions than configuration *c2* when running for short periods of time (the user defined parameters are: MIPfocus = 1, Heuristic = .9 and Cuts = 0).

From these experiments, it appears that our approach, RHCC, gets better results in less than 20 seconds than GFC in two hours, as we see that GFC does not improve the solutions found by the RHCC when it is run 20 seconds. It can also be seen that the GFC was only able to slightly improve the results when the RHCC ran for five seconds to provide an initial solution. The difficulty for the GFC algorithm to find new primal solutions was also reported by Shen (2004). In his case study (85 gateways for UPS in the US), the GFC algorithm was not able to improve the provided initial solution after 100 hours of generating columns and other 100 hours of running the final master problem.

As a conclusion, we see that our approach compares favorably in terms of efficiency and effectiveness to the existing exact approaches in the literature for solving the ESSND with flexible hub assignment.

Table 3 Average results of the column generation algorithm (GFC) for the base cases.

	GFC-5s	GFC-20s	GFC-60s
Initial solution running time (s)	5	20	60
Sub-problems	23,126	25,708	24,697
Columns generated	19,964	21,942	21,148
Master iterations	4.5	4.9	7.9
Provided solution optimality gap* (%)	20.50	6.11	3.65
GFC optimality gap* (%)	20.10	6.11	3.65

*Calculated with respect to the best lower bound known for each problem.

5.2. Base cases variations

For a more comprehensive testing of the performance of the RHCC, we analyze 15 variations of the base cases explained in Section 5.1 (for the four maps, see Table 1), and present the results for 304 additional experiments. Each variation is obtained by modifying one important feature from an economic and operative perspective, while keeping the other parameters to their default values. The input parameters modified for obtaining each variation are the following:

Volume of commodity demands: It is doubled or cut by half.

Hub sorting capacities: They are changed by increasing/decreasing by 0.5 hours the time window between the pickup due time and the delivery release time (the larger the time window, the smaller the number of feasible routes).

Number of aircraft types: It is reduced from six to five, and then four, which reduces the number of feasible routes.

Distance between hubs: It is increased by a factor of 1.5, or 2. This changes the network structure in many ways (e.g., traveling distances of the routes, number of feasible routes, feasibility to connect a gateway with a hub).

Number of gateways: It is increased to 40. This is done by joining two of the maps from the base cases.

Number of hubs: It is increased to three and then to four. By adding more hubs, the five hub capacity scenarios used for the two-hub tests cannot be applied. For the variation with three hubs, six scenarios have been defined with the following hub sorting capacities with respect to the total demand: 50%-35%-15%, 35%-15%-50%, 15%-50%-35%, 70%-20%-10%, 20%-10%-70% and 10%-70%-20%. Since they are applied for the four maps, they account for 24 experiments. For the variation with four hubs, five scenarios have been defined with the following hub sorting capacities with respect to the total demand: 25%-25%-25%-25%, 50%-25%-15%-10%, 25%-15%-10%-50%, 15%-10%-50%-25% and 10%-50%-25%-15%.

From the 11 variations explained so far, we selected four (base cases, 40 ramps, 3 hubs and 4 hubs) to create additional variations in which we include a business rule used by some express carriers, referred to as the *major hub connectivity condition* (MHCC) (Barnhart et al. 2002, Kim et al. 1999, Shen 2004). The MHCC states that each gateway must be connected by at least one pickup and one delivery route to their major hub, i.e. the one with the largest sorting capacity. Therefore, the MHCC enforces the existence of an origin-hub-destination path for each commodity through the main hub. It is thus a special case of the commodity connectivity inequalities.

Table 4 shows the average results of the base cases and its 15 variations in terms of optimality gap and of cost change compared to the best primal solutions known for the corresponding base cases. Since the variants with 3 and 4 hubs have different hub sorting capacities, their cost change can

only be calculated on the average costs and not for each instance. The results show the consistency of the RHCC model on the 15 variations. The optimality gap remains low with an average of 2.93%, and the RHCC found the optimal solutions in 113 out of 304 instances. For the 8 variations with the same number of hubs and gateways as the base cases, the average optimality gap (1.26%) is practically equal to the gap of the base cases (1.24%). Larger optimality gaps are naturally observed in variations that make the problem harder to solve and increase the solution space. Augmenting the demand eliminates feasible solutions, making the optimality harder to prove. Extending the sorting time reduces the solution space. Doubling the number of gateways leads to triple the number of routes and more than quadruple the number of commodities. Augmenting the number of hubs increases the number of routes and of possible hub assignments. Unsurprisingly, the worst gaps are obtained with more gateways or hubs.

For these cases, we tested the addition of the major hub connectivity condition (MHCC), leading to better optimality gaps. This condition facilitates the bounding while branching as many routes serving the secondary hubs become inefficient once a ramp is connected to the major hub. The average optimality gap decreases to 2.97%, compared to 8.78% without the MHCC. The cost

Table 4 Average results for the base case variations solved with RHCC.

	RHCC	
	Optimality gap (%)	Cost change (%)
Base cases	1.24	0.16
Base cases with MHCC	0.34	-0.14
Double demand	2.87	42.4
Half demand	0.03	-30.5
Hub sorting time -0.5 hours	2.45	-2.17
Hub sorting time +0.5 hours	0.02	14.1
5 aircraft types	0.84	0.45
4 aircraft types	0.46	1.21
Hubs distance X 1.5	1.60	3.15
Hubs distance X 2	1.78	3.80
Average	1.26	3.6
40 gateways	8.31	111.3
3 hubs	6.72	0.40 [†]
4 hubs	11.31	-1.46 [†]
Average	8.78	36.8
40 gateways with MHCC	2.49	104.6
3 hubs with MHCC	2.59	1.09 [†]
4 hubs with MHCC	3.83	0.63 [†]
Average	2.97	35.4
Overall average RHCC	2.93	15.5
Overall average RHCC-RC	3.22	15.8

[†] Calculated on the average costs and not for each instance.

performance is improved for the case with 40 gateways (as the solutions procedure is more efficient) but deteriorated when the number of hubs is increases (as the MHCC rules is more limiting). It is worth noting that even the results on the base cases are improved, with an optimality gap of only 0.34%.

As for the cost changes and economical results in general, we can see the effect of the economies of scales when the demand changes (the cost increases only by 45% when the demand doubles). We also observe an increase of the cost when the solution space is reduced by decreasing the number of routes (e.g., by increasing the hub sorting times or reducing the number of aircraft types).

Finally, to continue the analysis on the contribution of the route covers (RC), we also performed the 15 variations with RHCC-RC reduced model. From the averages, we see that the RHCC tends to perform better than the RHCC-RC, and indeed in 12 out of 15 variations it gives a better gap. The usefulness of the route covers is confirmed on lareger size instances next.

5.3. Full size instance for an Express carrier

The last set of experiments is based on a realistic dual-hub case within the European region, with 76 gateways and 7 equipment types. Table 5 shows its characteristics in terms of size and routes composition. Eleven variations of the case were defined by changing the hub sorting capacities. The extreme variations correspond to one hub with no sort capacity and the other receiving the total demand (0%-100%). The variations in between change these percentages in steps of 10%, where both hub sort capacities sum up to the total demand. As previously, we compare the results of the existing models (RF and RH) with our model (RHCC), and we evaluate the contributions of the valid inequalities and of the route covers by removing them independently from. Each model is run for up to two hours for each variation, using the Gurobi configuration *c2* (see introduction of Section 5), with 7 threads. Also, for all the variations, we included the Main Hub Connectivity Condition (MHCC) (see Section 5.2).

Table 6 shows the results of the experiments for the 11 variations of the realistic dataset. Comparing to the state of the art models RF and RH, the RHCC models, the performance of the RHCC shows again drastic improvements, with an average optimality gap of 5.08%, while the RH and RF get gaps of 27.1% and 65.6% respectively. As for the cost performance, the RHCC always obtained a better primal solution, with a cost improvement of almost 9% compared to the RH

Table 5 Full size instance characteristics.

Statistics				Number of routes				
$ N $	$ k $	$ H $	$ F $	1-leg routes	2-leg routes	3-leg routes	Ferry routes	Total
76	5,266	2	7	1,460	7,544	849	24,578	34,431

Table 6 Results for variations on full size instances.

Sort capacity (%)	Optimality gap (%)						
H1 - H2	RF	RH	RHCC				full model
			-CCI	-SLI	-SCI	-RC	
0 - 100	44.0	9.3	1.08	1.32	1.61	1.05	0.65
10 - 90	63.4	24.5	2.51	2.89	8.70	2.86	2.60
20 - 80	64.7	26.5	5.21	5.07	9.02	5.02	3.99
30 - 70	71.0	33.1	8.72	7.48	10.3	8.92	6.67
40 - 60	72.3	37.7	8.88	9.37	11.5	9.57	8.60
50 - 50	83.8	39.8	8.97	8.39	9.62	8.90	8.69
60 - 40	83.9	36.9	8.12	8.89	12.1	9.42	8.46
70 - 30	70.0	31.1	8.71	8.40	12.8	10.1	7.48
80 - 20	63.9	27.6	5.03	6.49	10.8	6.55	4.82
90 - 10	62.6	21.6	3.73	3.75	7.82	3.06	3.28
100 - 0	42.4	9.9	0.74	1.31	1.38	0.85	0.63
Average optimality gap	65.6	27.1	5.61	5.76	8.70	6.03	5.08
Average cost performance	100.8	9.22	0.43	0.72	2.34	1.27	0.26

(100% compared to the RF). When looking at the RFC model, i.e. the valid inequalities are added to the RF, the root relaxation did not solve for any instance (in two hours). A similar behavior occurred by adding only the SLI to the RF. This is because, with disaggregated flows, the number of SLI is too high, thus showing the computational advantage of using aggregated flows (RH) for large instances of the ESSND. Furthermore, Table 6 allows us to compare the results of the RHCC and all its reduced versions. While in the limited size base cases removing the CCI from the RHCC caused the largest change, for these larger cases this change is the smallest. This is because the MHCC rule, which is applied to these instances, has the same effect as the CCI for the major hub (not for the second hub). Comparing the RHCC-SLI and the RHCC-SCI, the later worsens sensibly more for both the optimality gap and cost performance (almost 3% and 1.6% respectively), which confirms that the structure of the SCI better exploited than the structure of the SLI in the branch and cut procedure. Finally, for these large instances, removing the RC worsens both the cost and optimality gap by almost 1%, showing their contribution. For express integrators even a one percent cost reduction usually represent millions of dollars per year. Overall, when looking at Tables 2 and 6, it appears clearly that the proposed valid inequalities and covers bring a large improvement both in terms of optimality gap and cost performance when applied together, and that each of them contributes individually.

FedEx Express Europe ran independently a series of tests with our RHCC and confirmed achieving similar performance. The tests included a comparison between the RHCC and the algorithms they use for solving the multi-hub ESSND problem with flexible hub assignment. The comparison was performed using the 9 instances of our large experiments that are properly multi-hub cases (2 are single-hub instances). In the 9 cases, the RHCC got better primal solutions, showing an average solution improvement of 4.4%.

6. Conclusion

In this paper, we study the multi-hub Express Service Network Design problem, whose goal is to define a network of flights that, at minimum cost, enables the overnight flow of packages from their origins to their destinations. The case where the hub assignment is included as a decision, the ESSND with *flexible hub assignment*, has been scarcely addressed in the literature. In this paper, we address this problem and develop the Route and Hub model with Cuts and Covers (RHCC), aiming to improve the optimality gap and the primal solutions compared to the state of the art models. We propose three families of valid inequalities – commodity connectivity inequalities, strong linking inequalities and strong cutset inequalities. The commodity connectivity inequalities ensure there is a valid origin-hub destination path for each commodity. The strong linking inequalities disaggregate the aircraft capacity constraints. The strong cutset inequalities ensure that there is enough capacity to move the demand between two subsets of nodes, strengthened by eliminating the idle capacity of the routes. Then, we reformulate the forcing constraints with the *route covers* which are routes that have more capacity than the demand of the gateways they visit. Thus, their related capacity constraints and flow variables can be omitted without altering the nature of the problem. To the best of our knowledge, the commodity connectivity constraints are new to the multi-hub ESSND problem, and the strong cutset inequalities and the route covers are new to the ESSND problem.

The numerical experiments show that, for instances with limited size (20 gateways, a size commonly found in the literature for the multi-commodity network design problem), the Route and Hub model with Cuts and Covers lead to an average optimality gap of 1.24%, showing a drastic improvement compared to the exact state of the art approaches (Kim et al. 1999, Shen 2004). For a more comprehensive testing of the performance of our RHCC, we analyzed 15 variations of the instances with limited size by changing six important input parameters (demands, sorting capacities, hub locations, number of equipment types, number of gateways and number of hubs). For all the variations, the average optimality gaps remain at low levels, obtaining a total average of 2.93% (1.26% on same size instances and 2.97% on instances with a larger number of gateways or hubs using the MHCC rule). In the case of large realistic instances (76 gateways), the improvement of our formulation is even more significant. The RHCC obtained the best performance with an average optimality gap of 5.08%, improving by 22.02% and the cost by almost 9% compared to the best performing state of the art model. As a conclusion, we see that our approach compares favorably in terms of efficiency and effectiveness to the existing exact approaches in the literature for solving the multi-hub ESSND with flexible hub assignment.

A close research collaboration with FedEx Express Europe grounded the model presented in the actual planning challenges faced by the express air carriers. Our research partner built the testing

instances used in this paper to ensure that they are representative of real instances. They confirmed that the performance exhibited by our numerical results is similar to the one they obtained with their own testing, and ultimately confirmed the value of our approach for the industry.

Nevertheless, this research could be further developed. First, a cutting plane methodology could be designed to dynamically include the most useful valid inequalities to the model. Second, one could search new families of valid inequalities to further strengthen the LP relaxation. Finally, more types of routes could be added to the problem, such as direct routes between gateways, inter-hub routes that move consolidated volumes between hubs, and trans-load routes that allow the transshipment of packages between vehicles at any node in the network.

Acknowledgments

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

References

- Agarwal YK (2002) Design of capacitated multicommodity networks with multiple facilities. *Operations Research* 50(2):333–344.
- Armacost AP, Barnhart C, Ware KA (2002) Composite variable formulations for express shipment service network design. *Transportation science* 36(1):1–20.
- Armacost AP, Barnhart C, Ware KA, Wilson AM (2004) Ups optimizes its air network. *Interfaces* 34(1):15–25.
- Barnhart C, Krishnan N, Kim D, Ware K (2002) Network design for express shipment delivery. *Computational Optimization and Applications* 21(3):239–262.
- Barnhart C, Schneur RR (1996) Air network design for express shipment service. *Operations Research* 44(6):852–863.
- Belotti P, Malucelli F, Brunetta L (2007) Multicommodity network design with discrete node costs. *Networks* 49(1):90–99.
- Boeing Commercial Airplanes (2014) World air cargo forecast 2014-2015.
- Chouman M, Crainic T, Gendron B (2003) A cutting-plane algorithm based on cutset inequalities for multicommodity capacitated fixed charge network design. Technical report, Centre for Research on Transportation.
- Chouman M, Crainic TG (2014) Cutting-plane matheuristic for service network design with design-balanced requirements. *Transportation Science* 49(1):99–113.
- Crainic TG (2000) Service network design in freight transportation. *European Journal of Operational Research* 122(2):272–288.
- Crainic TG, Frangioni A, Gendron B (2001) Bundle-based relaxation methods for multicommodity capacitated fixed charge network design. *Discrete Applied Mathematics* 112(1):73–99.

-
- Crainic TG, Gendron B, Hernu G (2004) A slope scaling/lagrangian perturbation heuristic with long-term memory for multicommodity capacitated fixed-charge network design. *Journal of Heuristics* 10(5):525–545.
- Crainic TG, Rousseau JM (1986) Multicommodity, multimode freight transportation: A general modeling and algorithmic framework for the service network design problem. *Transportation Research Part B: Methodological* 20(3):225–242.
- Fleuren H, Goossens C, Hendriks M, Lombard MC, Meuffels I, Poppelaars J (2013) Supply chain-wide optimization at tnt express. *Interfaces* 43(1):5–20.
- Gendron B, Crainic TG (1994) Relaxations for multicommodity capacitated network design problems. Technical report, Centre for Research on Transportation.
- Gendron B, Crainic TG, Frangioni A (1999) *Multicommodity capacitated network design* (Springer).
- Kim D, Barnhart C, Ware K, Reinhardt G (1999) Multimodal express package delivery: A service network design application. *Transportation Science* 33(4):391–407.
- Kuby MJ, Gray RG (1993) The hub network design problem with stopovers and feeders: The case of federal express. *Transportation Research Part A: Policy and Practice* 27(1):1–12.
- Louwerse I, Mijnaerends J, Meuffels I, Huisman D, Fleuren H (2014) Scheduling movements in the network of an express service provider. *Flexible Services and Manufacturing Journal* 26(4):565–584.
- Magnanti TL, Wong RT (1984) Network design and transportation planning: Models and algorithms. *Transportation science* 18(1):1–55.
- Meuffels WJM (2015) The design of road and air networks for express service providers. Ph.D. thesis, Tilburg University, School of Economics and Management.
- Minoux M (1989) Networks synthesis and optimum network design problems: Models, solution methods and applications. *Networks* 19(3):313–360.
- Pedersen MB, Crainic TG, Madsen OB (2009) Models and tabu search metaheuristics for service network design with asset-balance requirements. *Transportation Science* 43(2):158–177.
- Shen S (2004) Logistics service network design: models, algorithms, and applications. Ph.D. thesis, Massachusetts Institute of Technology.
- Sherali HD, Bae KH, Haouari M (2013) An integrated approach for airline flight selection and timing, fleet assignment, and aircraft routing. *Transportation Science* 47(4):455–476.
- Wieberneit N (2008) Service network design for freight transportation: a review. *OR spectrum* 30(1):77–112.