MASTER

Stochastic Service Network Design with demand and travel time uncertainty

Orij, Rick

Award date:
2023

Stochastic Service Network Design with demand and travel time uncertainty

Master Thesis

Supervisors: Dr. A.H. Schrotenboer (TU/e)
Dr. L. Martin (TU/e)
Dr. R.N. van Lieshout (TU/e)
Maité Delannoy (DB Schenker)

Author
Rick Orij

Student ID
0999874

October 31, 2023

Abstract

This thesis proposes a Stochastic Service Network Design formulation that addresses both demand and travel time uncertainty, while also incorporating the possibility of outsourcing services. To account for these uncertainties, a multi-scenario formulation is used. The model is based on the well-known time-space network formulation, with the distinction that the time-space network is now scenario dependent, resulting in varying lengths of arcs for different scenarios. Recognizing the computational intensive nature of the problem, a solution method based on the Progressive Hedging method is developed, which decomposes the multi-scenario problem into single-scenario sub problems. The solutions of these single-scenario sub problems are aggregated, and by using information on a trend between these sub problems, costs are adjusted to guide the method to a consensus between the sub problems, to eventually get a good solution for the entire problem. This method comprises two strategies that have demonstrated their effectiveness in solving instances more efficiently compared to using a commercial solver for the original multi-scenario model formulation. To test the practical application of the solution method, a case study was conducted for a well-known LTL transport carrier. The method was used to find a new solution for parts of their operations to give insights for potential improvements.

Executive Summary

The Service Network Design problem has been widely used to model the operations of freight consolidation for less-than-load (LTL) carriers. In this problem freight from different origins is picked up and consolidated in the origin region at the origin hub to be transported to the destination hub, from which the freight is further distributed to the destination hubs. However within this problem some uncertainties might occur. Figure 1 illustrates these types of uncertainties with question marks. It can be seen that uncertainty might occur in the pick-up of freight from the supply-node to the hub node, and therefore the time the freight is ready to be transported to the destination terminal. Another factor of uncertainty in demand might occur in the volume of the freight. At last the duration of travel between two hub nodes may be uncertain. Within the literature uncertainty in demand and uncertainty in travel time have been discussed, however the combination of these two uncertainties is not well investigated. It is believed that modeling the combination of these uncertainties might improve the robustness of the operated schedule of a carrier. Therefore this research aims to contribute to the literature by modeling the uncertainty of demand and travel time in a Service Network Design, as well as a solution method to model the problem in a more efficient manner. The modeling is tested with a case study at the land-transport department of DB Schenker, a well known freight carrier. The contributions of this thesis are as follows:

1. Proposing a linear mixed-integer programming model of a Service Network Design, dealing with uncertainty in travel time and demand.

2. Proposing a new manner of using a time-space network in different scenarios and different lengths of arcs in the context of a Service Network Design problem.

3. Propose a Progressive-Hedging-based solution method that is able to deal with both demand and travel time uncertainty able to provide a close to optimal solution to the problem.

4. Show that the solution method is more efficient than using a commercial solver to solve the Service Network Design with demand and travel time uncertainty, without compromising solution quality

5. Use the solution method to give DB Schenker insights in their operations and give suggestions for improvements.

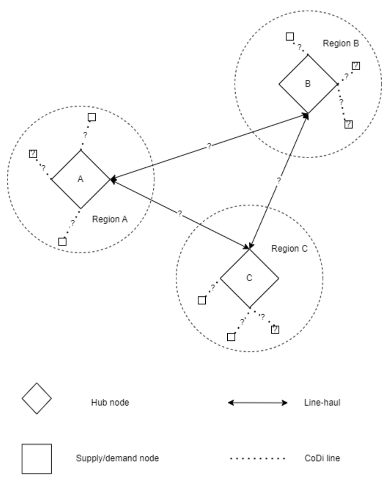


Figure 1: Schematic of Service Network Design with uncertainty

Method

In order to model the uncertainty of demand and travel time in a Service Network Design, a Mixed Integer Programming (MIP) model is proposed that models the uncertainty in the form of scenarios. The model takes the form of a multi-scenario deterministic formulation, where the underlying structure of the model is based on a time-space network. The model allows for the possibility of outsourcing when uncertainty in demand occurs, as well as recourse actions for allowing the possibility to delay shipments. In order to model the uncertainty in travel-time the structure of the time-space network is adjusted to allow for multiple scenarios. This has as result that arcs in the network between the same origin-destination pairs can have different lengths. In order to construct a schedule that can be used for the operations of a LTL carrier, a non-anticipativity constraint is added to the model so that within the multiple scenarios the same departure time at a origin destination takes place. This constraint thus enforces that delivery in every scenario will be on time.

Modeling this problem for increasing network size becomes increasingly complex, resulting in large computation times. In order to solve the problem more efficient without suffering much from solution quality, a Progressive-Hedging based matheuristic is developed to solve the Service Network Design problem with uncertainty and travel time. Introduced by Rockafellar and Wets (1991), progressive hedging decomposes the entire problem into single-scenario sub-problems. The entire problem is then solved iteratively, with each iteration solving the single-scenario problems. The idea is to investigate if there is some consensus between the sub-problems on which services to use. If there is a trend on either using or not using services, penalty factors in the single-scenario problems are adjusted to guide to solutions of the problems to a consensus between these problems. When the stopping criteria is met, either through reaching a maximum number of iterations, reaching a maximum number of iterations with an improvement or reaching a time limit, the services for which some amount of consensus is reached are fixed. The first phase of the method thus tries to find a consensus between the single-scenario problems. The second phase these fixed variables are used to solve the entire problem. For the Progressive-Hedging method two strategies can be used, the original *Lagrangian* strategy based on the method from Rockafellar and Wets (1991) and the *Heuristic* strategy based on the method from Crainic et al. (2011).

Results and conclusions

Both the MIP model and the Progressive-Hedging method are tested on their performance by testing them on solution quality and efficiency. For this evaluation test instances are generated loosely based on shipments and a part of the network of DB Schenker, where both factors of demand and travel time uncertainty are taken into account. First the MIP model is solved with a well-known commercial solver, Gurobi, that is also able to provide a lower bound for the instances. After that both the *Lagrangian* and *Heuristic* strategies of the Progressive-Hedging method are used to solve the instances. To solve the single-scenario sub-problems Gurobi is used, with a local gap-limit in solving the sub-problem to avoid getting stuck in a local optima. Different values for this gap-limit are also tested to see the difference in performance. From the experiments on the test instances it was concluded that the Progressive-Hedging method in general provided solutions close to the lower bound with considerable drop in computation time compared to the original MIP model solved with Gurobi. Of all the Progressive-Hedging methods, the *Heuristic* strategy with a gap-limit of 5 % for the single-scenario problems provided the best efficiency relative to the solution quality.

This method is then used for the case study at DB Schenker, with the goal to gain insights in some of the disrupted line-hauls and maybe find improvements on the current schedule. For the case study two sub cases were studied, one on the freight from the BeNeLux to Serris and on on the freight from the BeNeLux to Bologna. The method was used to find a new solution for the already operated line-hauls, as well as a greenfield solution with all freedom to optimize as much as possible. The method was able to find a solution

Line-haul	DB Schenker as-is planning	PH solution
EDE-SXR	Monday 20:00 - Tuesday 04:00 Tuesday 20:00 - Wednesday 04:00	Monday 16:00 - Tuesday 00:00 Tuesday 10:00 - Tuesday 18:00 Tuesday 23:00 - Wednesday 07:00
TLB-SXR	Monday 22:30 - Tuesday 05:30 Tuesday 22:30 - Wednesday 05:30	Monday 16:00 - Monday 23:00 Tuesday 07:00 - 14:00 Tuesday 23:00 - Wednesday 06:00

Table 1: Schedules of DB Schenker as-is planning and the PH solution for the case of Serris for $\frac{1}{10}c^k_{l_t,m_{t'}}$.

Line-haul	DB Schenker as-is planning	PH solution
TLB-BLQ	Monday 00:30 - Monday 21:00 Tuesday 00:30 - Tuesday 21:00 Wednesday 00:30 - Wednesday 21:00 Thursday 00:30 - Thursday 21:00	Tuesday 04:00 - Wednesday 01:00 Wednesday 00:00 - Wednesday 21:00 Wednesday 17:00 - Thursday 16:00

Table 2: Schedules of DB Schenker as-is planning and the PH solution for the case of Bologna for $\frac{1}{10}c^k_{l_t,m_{t'}}$.

with reduction in costs for every case. Table 1 and Table 2 show the current schedule and the solution of the Progressive-Hedging method for both the cases of Serris and Bologna respectively. It can be seen that the method is able to find a solution, in the case of Bologna close the current schedule with the exception of having a longer planned travel duration and needing one less line-haul. For the case of Serris it can be seen that an additional line-haul is being scheduled during the day. This will probably lead to additional costs in real life since the variance in volume of the trucks might lead to more empty miles. Moreover since this line-haul is scheduled during the day, more shipments will be stationary on the terminals while there may not be capacity for them.

These differences are mainly due to the cost structure. The solution method is based on optimization, so in the case of minimization it tries to minimize the costs. However, in this case a relatively large penalty is given to delayed shipments. Therefore the method is trying to deliver as much shipments, even though the data on the ready time and due time might not be accurate. Moreover, the costs for both using a service and for the flow of shipments in the network are approximated with a factor multiplied with the planned travel time and are not accurate. The method also does not consider other implications, such as concurrent line-hauls. When only looking at a small part of the network, the method does not consider other scheduled line-hauls in the network that have to be concurrent to the operated line-haul.

It is concluded that the method is able to deal with both uncertainty in demand and travel-time, and is able to provide a solution for potential improvement. The method however is not able to consider all implications, therefore the quality of the potential improvement is questionable.

Table of contents

1 Introduction

Freight consolidation for less-than-load (LTL) carriers has always been an important subject in literature, mainly focusing on the original hub-and-spoke network and the later, more extensive Service Network Design (SND). A hub-and-spoke network, first introduced by O'Kelly (1986), is a network in which freight is collected and distributed in local warehouses. Freight is consolidated in the local warehouses, or hubs, and sent to other local warehouses (hubs) with scheduled line-haul trucks. The freight is then further distributed from the destination hubs. Crainic (2000) describes the Service Network Design problem as a problem used for "intercity freight carriers that make intensive use of consolidation operations", where the carriers can mostly be classified as railways and LTL road transport carriers. The Service Network Design problem has been widely used for selection and scheduling of services and the routing of freight.

A company using a hub-and-spoke network for their LTL operations is DB Schenker. They are one of the world's leading global logistics providers, providing services to distribute goods through land transport, air and ocean freight. Their land transport network can be considered a hybrid hub-and-spoke network, where most of the services consists of LTL services, with exception of direct FTL (full-truckload) services if the volume of shipments is considered high enough to not need consolidation. Their network is currently managed by observing operations and fluctuations in demand and acting on triggers to be able to provide good service. Such a company observes a lot of uncertainties in their operations, so taking these uncertainties into account might be beneficial for the company to improve their operations.

While within the SND research much focus has been on constructing networks while considering factors such as travel time between hubs and demand between hubs are deterministic, more research is being done considering that those factors are usually stochastic. This is done since it is believed that taking uncertainty into consideration instead of solving a stochastic problem with deterministic methods will enhance solution quality and will provide a more robust solution. Often a network is constructed considering average demand; however, this constructed network might not prove very effective once uncertainty starts playing a role. Therefore with the stochasticity taken into account, a network is constructed, taking into account the uncertainty that may influence the solution, so that the solution is best suited for multiple scenarios. This is called two-stage stochastic programming, with the first stage fixing a certain set of services and the second stage dealing with uncertainty and arranging the flow over these selected services. Depending on the developed model, recourse actions can be taken to outsource excess flow for which no capacity is left. While much research within the Stochastic Service Network Design is done on either demand uncertainty or travel time uncertainty, not much research is done on combining these two types of uncertainty. Some research focusing on uncertainty in demand considers the possibility of outsourcing in case of excess demand; however, not much research focuses on modeling the resources for outsourcing.

To illustrate the Service Network Design problem with uncertainty, Figure 1.1 shows a network with three hub-nodes with either their demand/supply regions with their respective demand/supply nodes. Consolidation from the supply node of one region happens at the hub-node and is then transported to the region of demand, after which the demand is further distributed to the demand nodes from the hub-node. Within this network several uncertainties can play up, indicated in the figure by a question mark. At some demand/supply node a question mark is indicated, meaning we have uncertainty in the volume of that node. Then uncertainty in the CoDi line between a demand/supply node and a hub node can happen, indicating that there is uncertainty in travel time from the demand/supply node to the hub node and thus resulting in uncertainty in the readiness of a shipment. At last, a question mark can be seen at a line-haul, indicating the uncertainty in the travel time between hubs.

Within modeling a Service Network Design with uncertainty, we have mainly two types of approaches: Stochastic Programming and Robust Optimization. According to Xiang et al. (2022), for Stochastic Pro-

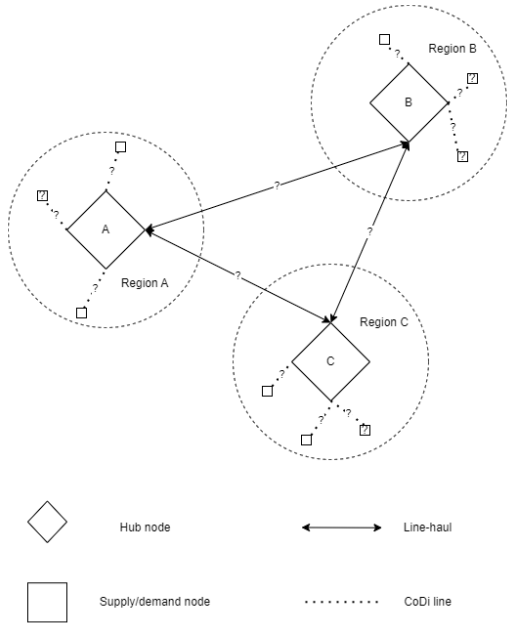


Figure 1.1: Schematic of Service Network Design with uncertainty

gramming, the exact probability distribution of the scenarios needs to be known, whereas Robust optimization aims to minimize the worst-case scenarios performance and does not require probability information. For this thesis, we use Stochastic Programming. The motivation for looking into Stochastic Programming rather than Robust Optimization is the fact that Stochastic Optimization is able to give better solutions, and it is assumed that information on the probability distribution of scenarios is known.

Since solving a Service Network Design, and a Stochastic Service Network Design in particular, can be difficult to solve due to high dimensionality (NP hard, Powell and Topaloglu (2003)) many solution methods, mainly (meta-) heuristics, have been proposed to be able to solve the problem more efficiently. Progressive Hedging is a method on which many meta-heuristics have been based to solve the Stochastic Service Network Design. The idea is to decompose the multi-scenario problem into single scenario sub problems and iteratively adjust the costs to guide the method to a consensus between the single scenario sub problems on an overall design.

This thesis proposes a model that can be used for modeling uncertainty in travel time and uncertainty in demand, with the possibility of also modeling different types of resources. We believe that combining these two of uncertainties can result in a more robust solution and as result a better suited schedule to the operations of a LTL carrier. In order to be used for larger instances such as the operations of a company as DB Schenker, the model formulation is altered so that a Progressive Hedging matheuristic is able to provide a good solution in order to compare solutions to current operations planning. Two methods are constructed based on the Progressive Hedging principle, of which one is based on the original augmented Lagrangian principle of the adjustment of costs and the other adjusting costs through a heuristic. First, the methods are tested on small instances, from which results show that both the Lagrangian based and the Heuristic based Progressive Hedging methods prove computationally more efficient than using a commercial solver, yet suffering little from objective value performance.

After that the methods are tested on the network of DB Schenker to find new insights on operating their network. Results show that the method is able to provide a new schedule that has overall better objective values and can give insight in how to adjust the schedule. The proposed improvements however are in most cases difficult to implement due to many implications on the already existing network and due to implications of the solution, in particular with respect to how the costs are formulated. The method is therefore able to find a good objective value compared to the objective value of the current schedule due to the balance of the costs.

This contributions of this research are

1. Proposing a linear mixed-integer programming model of a Service Network Design, dealing with uncertainty in travel time and demand.
2. Proposing a new manner of using a time-space network in different scenarios and different lengths of arcs in the context of a Service Network Design problem.
3. Propose a Progressive-Hedging-based solution method that is able to deal with both demand and travel time uncertainty able to provide a close to optimal solution to the problem.
4. Show that the solution method is more efficient than using a commercial solver to solve the Service Network Design with demand and travel time uncertainty, without compromising solution quality.
5. Use the solution method to give DB Schenker insights in their operations and give suggestions for improvements.

The outline of this thesis is as follows. First a literature review on modeling the Service Network Design with taken into consideration uncertainty is given in Chapter 2. Chapter 3 describes the problem context followed by the model formulation used. Chapter 4 describes the framework of the solution method used to solve the model formulation. Then, the experimental study of the validation of the method is dealt with in Chapter 5, where the case study at DB Schenker is discussed in Chapter 6. Finally the conclusions and further improvements are discussed in Chapter 7.

2 Literature Review

For this thesis, the focus lays on the Stochastic Service Network Design problem where both demand uncertainty and travel time uncertainty is taken into consideration, while also considering the ability of using multiple resources. Service Network Design with multiple resources takes into consideration the ability to either operate with owned services, potentially long-term contract charters or spot contract services. Therefore literature is reviewed on the Stochastic Service Network Design problem, looking at both types of uncertainty or if possible combined. Moreover the option of using multiple resources in a (Stochastic) Service Network Design is reviewed. In this section thus the distinction is made between uncertainty in demand and uncertainty in travel time, while the problem of multiple resources is reviewed within these two types of uncertainty. This literature review shows that modeling both types of uncertainties has only been introduced in a model formulation, yet not in experimental studies. Therefore it is believed that is important to do experiments with a model formulation with uncertainty in demand and travel time. From the view point of Stochastic Programming, and the view of using a two-stage formulation, the construction of a network with owned services are considered first stage decisions, whereas the ability to outsource services with spot contracts and the ability to allow for the uncertainty of travel time are generally second stage decisions.

Lium et al. (2009) first introduced stochastic elements into service network design formulations by means of uncertainty in demand. They figured that structural differences with stochastic demand compared to deterministic demand provide a hedge against the uncertainty, by means of getting consolidation as a model output instead of consolidation as an a priori required property. This provides flexibility and robustness of the network design as a result. By using a procedure similar to the Monte Carlo simulation, they tested their formulation with a stochastic model. The comparison provided cost savings of 17% when using stochastic formulations instead of deterministic formulations. Hoff et al. (2010) were the first to use the formulation of a service network design with demand uncertainty on a real life instance. A metaheuristic using a neighborhood search was used for the research, where the required number of vehicles was tested for different instances. For small instances the objective was near optimal with the at least the same number of vehicles as optimal. For larger instances there was no comparison possible however.

Crainic et al. (2011) provided a two-stage stochastic programming formulation where the first stage deals with the design and where the second stage deals with the recourse decisions according to the commodity flow. For a solution method the authors took inspiration from Rockafellar and Wets (1991) for a progressive-hedging based meta-heuristic framework. The idea behind progressive-hedging is that the uncertainty of the problem is modeled by small numbers of sub problems derived from the underlying optimization problem. These sub-problems correspond to different scenarios, a limited representation of information on how the information on the problem might evolve. The idea is then that by looking at the sub problems and their particular optimal solutions there might becomes a trend visible between the different solutions. Acting on this trend by adjusting the costs of problems for which there is a trend on consensus between the sub problems, finally a "well hedged solution" to the original problem can be found. By using this method, scenario decomposition is used wherein each scenario is solved as a deterministic sub problems, after which the framework guides the solution of the initial problem by using some local information. The framework provided high quality solution quality for numerous instances.

Crainic et al. (2014) further improved the meta-heuristic based on progressive-hedging by dividing the problem into multi-scenario sub problems. They tried different strategies of grouping the scenarios by means of grouping similar scenarios, grouping dissimilar scenario's and evaluating whether it is more beneficial to create groups of scenarios that cover or partition the set of scenarios. These different strategies were compared against each other, while furthermore also against grouping the scenarios randomly and finally against the lower bound given by a state-of-the-art MIP solver. In the end they proved that using the different grouping strategies of the meta-heuristic based on multi-scenario sub problems can provide better solution quality and computational efficiency compared to only using single-scenario sub problems. Out of all the grouping strategies the covering strategy proved to lead to the best solution quality. Bai et al. (2014) considers the Service Network Design with demand uncertainty and takes the approach of rescheduling an existing network to improve costs for different scenarios. A two-stage model is proposed where the first stage consists of the network design and the second stage models the rerouting of the vehicles. The new model was found to take more computation time, however it proved to work more efficient while working with uncertainty. Hewitt et al. (2019) constructed a decision support tool by means of a scheduled network design model that addressed strategic decisions regarding resource acquisition and management under uncertainty. The idea was to optimize a repeatable transportation plan concerning the fleet size and potential outsourcing of services while having uncertain freight volumes or demand. They proposed a MIP model addressing the uncertainty aspect by means of using scenarios to model the fluctuating volumes. Since solving a stochastic program exactly imposes several difficulties, they propose a column-generation-based matheuristic that decomposes the problem. This matheuristic was used to develop potential cycles as an input for the model. Two sets of instances were used to evaluate the performance of the mathereustic, one generated to mimic the operations of a LTL freight carrier and one based on the existing network of an European postal carrier. For both sets the solution quality was considered high and the solution was obtained within acceptable computation times.

While in the literature the focus on the Stochastic Service Network Design with uncertainty in demand mainly lies on the outsourcing activity as a recourse action, Müller et al. (2021b) introduce a model where "ad-hoc modification of the planned schedule for the transportation assets is possible". Within the research a two-phase matheuristic is developed that is computationally much more efficient than using a commercial solver. The results prove that using the matheurisic the cost reduction is far better compared to using an expected value model or stochastic model without considering the option to modify the schedule. Very similar to the other literature Liu et al. (2023) tried to find fixed routes vehicles and flexible vehicles for the Stochastic Service Network Design. Different is that while in other literature the routes of vehicles are designed in the first stage, whereas Liu et al. (2023) generates fixed routes in the first stage and flexible routes in the second stage. Therefore a new type of solution technique was proposed called the learning-based multiple scenario approach to solve their two-stage MIP. Their method is based on the multiple scenario approach (MSA) first introduced by Bent and Van Hentenryck (2004), where the method uses deterministic scenario subproblems to come to a solution for the main problem. The learning-based multiple scenario approach extends the MSA by using scenario grouping and building multiple-scenario subproblems instead of single-scenario subproblems, and is thus somewhat similar to the method used by Crainic et al. (2014). From the evaluation of their results they found a respectable decrease in costs when fixing the routes in comparison with the case of no fixed routes.

Demir et al. (2016) proposes a model for intermodal transportation with both uncertainty in travel time and demand, although they only propose a solution method on the model with uncertainty in travel time. Within their model they take the approach of not only minimizing shipment costs and penalty costs for incurred delays, but also minimizing costs regarding CO2 emissions to deal with green house gas emissions. Each cost expression within the objective function than has their own weight. They state that within an intermodal transportation network the synchronization of time of flows is vital since this will affect the possibility of transshipping goods from one type of transport onto another. The model they propose is a combination of a classical Service Network Design combined with a Vehicle Routing Problem. The problem is solved by means of a sample average approximation method to deal with the uncertainty in travel times. This sample average approximation method is based on replacing the actual distribution with an empirical distribution by Monte Carlo sampling. That means that if the objective function corresponds to an expected value the objective function is approximated by its sample average estimate. The underlying problem is then solved by deterministic optimization methods. They concluded by stating that the proposed mathematical formulation provided robust solutions that were favorable compared to the solutions coming from a deterministic formulation.

Like Demir et al. (2016), Müller et al. (2021a) proposes a model for intermodal transportation combining a Service Network Design with an integrated Vehicle Routing Problem, although only for uncertainty in travel time. They combine a sample average approximation method with an iterated local search to solve the model. They consider three intermodal road-rail networks and model the uncertainty in the travel times of the road transportation. The conclusion of the research was that while the impact of delay reductions is an interesting subject, the impact was relatively low compared to the impact of changes in intermodal cost. Therefore whether to choose for a mode of transport depends more on the costs of the different modes rather than the costs of delay of a particular mode.

Whereas Demir et al. (2016) and Müller et al. (2021a) both addressed the problem of uncertainty in travel time for an intermodal transportation network, Lanza et al. (2021) provides a more generic model for a Stochastic Service Network Design without an integrated Vehicle Routing Problem. Furthermore the model is based on a space-time network rather than a directed network. Lanza et al. (2021) also define quality targets for delayed services and delayed delivery of volumes, where a penalty is given to every delayed service or demand. Like Crainic et al. (2011) and Crainic et al. (2014) a progressive-hedging based meta-heuristic is used to solve a two-stage model, wherein the first stage the decisions based on flow-distribution and the network design are made. Then a consensus is searched for which each solution for a particular scenario represents a similar design. On the basis of the augmented Lagrangian mechanism (from Rockafellar and Wets (1991)) costs of the sub-problems are changed to reach the consensus. The solution quality of the proposed metaheuristic proved very good when compared to CPLEX.

Of the literature on uncertainty in travel time, only Lanza et al. (2021) use a time-space network where time is modeled within the network structure, whereas Demir et al. (2016) and Müller et al. (2021a) use separate time constraints in their model formulation. Lanza et al. (2021) does not bother changing the structure of the time-space network and uses additional second stage variables to model the uncertainty. Boland et al. (2017) proposed a method where the duration of an arc in a time-space network can be altered in order to better deal with a continuous time deterministic Service Network Design. This alteration however might also be used to model differing arc duration's as if they are different in differing scenarios, therefore having the potential to be used for modeling uncertainty in travel time with multiple scenarios.

As can be stated from the literature review much research has been done on uncertainty in demand and uncertainty in travel time, yet few has been done on the combination of the two. Demir et al. (2016) has introduced a model formulation, however did not test their formulation in the experiments. Furthermore a time-space network has already been used to deal with uncertainty in travel time by Lanza et al. (2021), yet the uncertainty of the travel time is modelled by means of second stage variables and not as part of the structure of the time-space network. Moreover, little research has been done looking at deployment of multiple resources, let alone combining this with taking into consideration both types of uncertainty. Finally, heuristics based on Progressive-Hedging have never been used to tackle a problem with the combination of these problems. Combining these two of uncertainties however can result in a more robust solution and as result a better suited schedule to the operations of a LTL carrier. And although the problem becomes more complex by introducing these factors into a SND problem, a Progressive-Hedging based method might be efficient enough to deal with these factors. Therefore it is believed that this thesis is contributing to existing literature. To have a clear overview of the literature being handled in this section, Table 2.1 list every research and their contributions.

Source	TF	UD	UTT	MR	TSN	PB	Solution Method
Lium et al. (2009)	Road	✓	-	-	✓	✓	MIP
Hoff et al. (2010)	Not specified	✓	-	-	✓	✓	Neighborhood search metaheuristic
Crainic et al. (2011)	Not specified	✓	-	-	✓	✓	Two-stage matheuristic with PH
Bai et al. (2014)	Not specified	✓	-	∼	✓	✓	MIP
Crainic et al. (2014)	Not specified	✓	-	-	-	✓	Two-stage matheuristic with PH
Demir et al. (2016)	Intermodal	✓	✓	∼	-	✓	SAA method
Hewitt et al. (2019)	Not specified	✓	-	✓	✓	✓	CG matheuristic with local search
Müller et al. (2021a)	Intermodal	-	✓	-	-	-	SAA method combined with iterated local search
Müller et al. (2021b)	Intermodal	✓	-	∼	✓	✓	Two phase matheuristic with tabu-search
Lanza et al. (2021)	Not specified	-	✓	-	✓	✓	Two-stage matheuristic with PH
Liu et al. (2023)	Road	✓	-	∼	-	✓	Learning-based MSA

Table 2.1: Summary table for literature on Stochastic Service Network Design. TF=Transportation Focus, UD=Uncertain Demand, UTT=Uncertain Travel Time, MR=Multiple Resources, TSN=Time Space Network, PB=Periodic Behaviour, ✓=Included, ∼=somewhat similar to MR, SAA=Sample Average Approximation, PH=Progressive-Hedging

3 Model Formulation

This chapter elaborates on the methodology of the research. In Section 3.1 the problem statement is given, after which Section 3.2 deals with the model formulation. Chapter 4 then discusses how the model is adjusted to be useful for the solution method.

3.1 Problem Statement

The problem is based on the formulation used by Hewitt et al. (2019), that deals with uncertainty in a service network design problem. At first the tactical decision making is discussed, where a carrier might distribute shipments through a network of terminals denoted by Λ, with origin terminals Λ^+, destination terminals Λ^- and maybe transshipment terminals Λ^*, therefore $\Lambda = \Lambda^+ \cup \Lambda^- \cup \Lambda^*$. Services are defined to be used to distribute shipments over the network. These services will be selected and scheduled to be included in the plan. At the tactical-level decision making, services are scheduled over a certain length, which comprises of $\mathscr{T} = \{1, 2, ..., T\}$ time periods.

In a usual time-expanded network, the selected plan consists of services that will be repeated on a schedule length basis. Knowing these time periods, a time space network is created, $\mathscr{G} = (\mathscr{N}, \mathscr{A})$, that models the transportation activities at different points in time with different nodes and arcs. The node set $\mathscr{N}$ models the terminals in different time periods, $\mathscr{N} = \Lambda \times \mathscr{T} = \{l_t | l \in \Lambda, t \in \mathscr{T}\}$, where l_t is terminal l at t. The arc set $\mathscr{A}$ consists of *service arcs* and *holding arcs*. *Service arcs* model the operation of a service between two terminals at a certain point in time. *Holding arcs* model the opportunity for a shipment or truck to be idle at a terminal for a certain amount of time. Service arcs are denoted by $\mathscr{S}$ and holding arcs are denoted by $\mathscr{H}$, therefore $\mathscr{A} = \mathscr{S} \cup \mathscr{H}$.

To deal with uncertainty, we have a set of scenarios, Ψ, with each scenario having a probability of $p_\psi, \psi \in \Psi$ of occurring. Therefore, we can say that every *service arc* may have a different length, depending on scenario ψ. As a consequence the set of service arcs becomes dependent on the set of scenario's, meaning we get $\mathscr{S}^\Psi$. To deal with the extra uncertainty of the service arcs we adjust the structure of the time-expanded network. A representation of the usual time-expanded network is given in Figure 3.1 for Scenario 0, since usually a time-expanded network only consist of deterministic arcs. As can be seen in the figure of Scenario 0, every arc in the network has a fixed length (e.g. a fixed duration from start to end node). However, with uncertainty in travel time, arcs may have different durations. Therefore it is shown that for Scenario 0 and Scenario 1 the duration to travel from l to m takes two time steps, however in Scenario 2 takes 3 time steps. $t^\psi_{l_t m_{t'}}$ denotes the travel time for given service from origin l to destination m at given time moments in scenario $\psi \in \Psi$

Ψ considers the three types of uncertainty are uncertainty in amount of volume, uncertainty in availability time of a shipment and uncertainty in travel time of a service. Shipments are modeled between terminal l and m, and have availability time a^ψ_k and a due time b_k. These shipments have an index k, with origin node $o(k) = l_a$ and destination node $d(k) = m_b$. The set of all shipments is denoted by $\mathscr{K}$. Then the uncertainty of the volume of these shipments is taken into account, where $q^{k\psi}$ denotes the volume of shipment $k \in \mathscr{K}$ in scenario $\psi \in \Psi$. Here $q^{k\psi}_i = q^{k\psi}$ when $i = o(k)$, $q^{k\psi}_i = -q^{k\psi}$ when $i = d(k)$ and $q^{k\psi}_i = 0$ otherwise. The decision variable $x^{k\psi}_{l_t m_{t'}}$ denotes the fraction of shipment k's demand that is transported with planned service $(l_t, m_{t'})$. Along with the volume of these shipments, the costs for shipping a shipment k is denoted by $c^k_{l_t m_{t'}}$.

Because of dimensionality, we therefore get that $\mathscr{A}^\Psi = \mathscr{S}^\Psi \cup \mathscr{H}^\Psi$. Regarding service arcs, we thus make a distinction between "planned" service arcs and "stochastic" service arcs. "planned" service arcs can be seen as the regular time a carrier would plan for a service and "stochastic" service arcs the time span it eventually might take to carry out a service. The aim of the model is that even though services in different scenarios

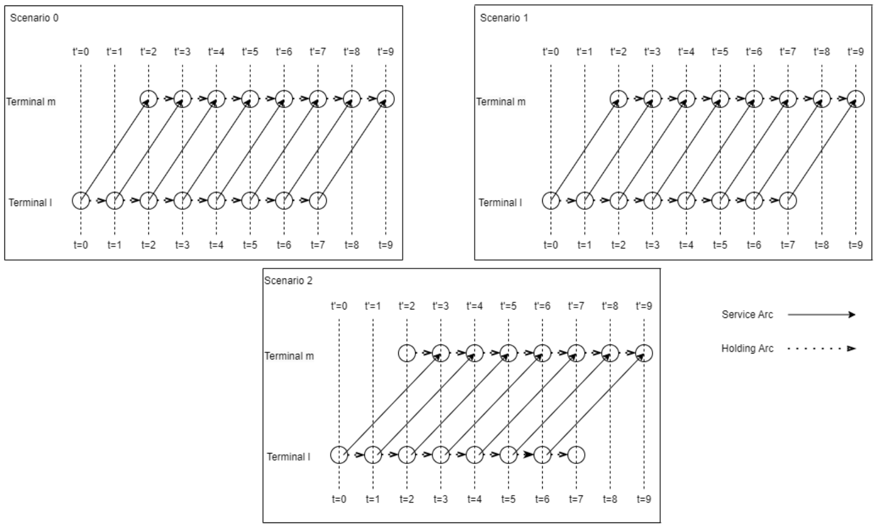


Figure 3.1: Schematic of uncertainty in time-space network.

might have a different duration between l and m, they should have the same departure time at origin l_t. By enforcing this, the "planned" service arc is used for non-anticipativity like purposes, such that the overall design remains similar. For each "planned" service $s = (l,m)$ between terminals $l,m \in \Lambda$ and time period $t \in \{1,...,T\}$, we have the arc $(l_t, m_{t'})$ where $t' = (t + t_{l_t m_{t'}}) mod T$ belonging to $\mathscr{S}$(with $t_{l_t m_{t'}}$ denoting the travel time for a service from l to m). For each "stochastic" service $s = (l,m)$ between terminals $l,m \in \Lambda$ and time period $t \in \{1,...,T\}$, we have the arc $(l_t, m_{t'})$ where $t' = (t + t^{\psi}_{l_t m_{t'}}) mod T$ belonging to $\mathscr{S}^{\psi}$ (with $t^{\psi}_{l_t m_{t'}}$ denoting the travel time for a service from l to m in scenario ψ). This is to force that schedules can be repeated. Capacity of services are set to $u_{l_t m_{t'}}$. Regarding the holding arcs, arcs of the form $(l_t, l_{t+1} mod T)$ are added. For this case holdover arcs are considered to be deterministic and take 1 time-step, however the length of holdover arcs might also depend on scenario ψ.

For a shipment to be distributed over a service arc $(l_t, m_{t'}) \in \mathscr{S} \subset \mathscr{A}$, that service must be executed by modeling this binary choice with a first stage variable $\bar{y}_{l_t m_{t'}}$. Executing this service incurs costs f_{lm}, that models all costs related to operating a service, labor, maintenance, trailer rental and transportation cost. Another option is that the service could be operated by outsourcing the service on the spot market. This decision is modeled by $w^{\sigma\psi}_{l_t m_{t'} k}$ with fixed costs f^{σ}_{lm}. These decisions are necessary to model the strategic behaviour in order to know how many resources are necessary to be able to transport all the shipments. Continuous variable $Z^{k\psi}$ is used to model the delay of a shipment with penalty costs related to a shipment c^{pen}_k.

3.2 Mathematical model

In this section the model used for the research is proposed. The reason that this research contributes to literature is that it aims to take into account resource acquisition, as well as multiple sources of uncertainty. Moreover, Stochastic Service Network Design problems are mainly formulated by a two-stage model, where first stage decisions consists of design decisions and second stage decisions of flow decisions. This research however takes the approach of Hewitt et al. (2019), where the model is formulated so that uncertainty is modeled by means of scenarios, therefore the model is formulated in a multi-scenario based, deterministic formulation. This model aims to minimize operating costs before and after uncertainty resolves by:

$$\text{minimize} \quad \sum_{(l_t,m_{t'})\in\mathscr{S}} f_{lm}\bar{y}_{l_t m_{t'}} \tag{3.1}$$

$$+\sum_{\psi\in\Psi} p_\psi \left(\sum_{(l_t,m_{t'})\in\mathscr{S}^\Psi} f^{\sigma}_{lm} w^{\sigma\psi}_{l_t m_{t'}} + \sum_{k\in\mathscr{K}} \sum_{(l_t,m_{t'})\in\mathscr{S}^\Psi} c^k_{l_t,m_{t'}} x^{k\psi}_{l_t,m_{t'}} + \sum_{k\in\mathscr{K}} c^{pen}_k Z^{k\psi} \right) \tag{3.2}$$

subject to

$$\sum_{(l_t,m_{t'})\in\mathscr{S}^\Psi} x^{k\psi}_{l_t,m_{t'}} - \sum_{(m_{t'},l_t)\in\mathscr{S}^\Psi} x^{k\psi}_{l_t,m_{t'}} = \begin{cases} q^{k\psi} - Z^{k\psi} & (l,t) = (o_k, e_{k\psi}) \\ -q^{k\psi} + Z^{k\psi} & (l,t) = (d_k, l_k) \\ 0 & \text{otherwise} \end{cases} \quad \forall k \in \mathscr{K}, \psi \in \Psi, (l,t) \in \mathscr{N} \tag{3.3}$$

$$\sum_{k\in\mathscr{K}} \sum_{(l_t,m_{t'})\in\mathscr{S}^\Psi} x^{k\psi}_{l_t,m_{t'}} \leq u_{l_t m_{t'}} (y^{\psi}_{l_t m_{t'}} + w^{\sigma\psi}_{l_t m_{t'}}) \quad \forall (l_t, m_{t'}) \in \mathscr{S}^\Psi, \tag{3.4}$$

$$\sum_{t'\in\mathscr{S}^\Psi} y^{\psi}_{l_t m_{t'}} = \bar{y}_{l_t m_{t'}} \quad \forall (l_t, m_{t'}) \in \mathscr{S}, \psi \in \Psi, \tag{3.5}$$

$$\bar{y}_{l_t m_{t'}}, y^{\psi}_{l_t m_{t'}}, w^{\sigma\psi}_{l_t m_{t'}} \in \{0,1\} \quad \forall (l_t, m_{t'}) \in \mathscr{S}, \psi \in \Psi, \tag{3.6}$$

$$x^{k\psi}_{l_t,m_{t'}}, Z^{k\psi} \geq 0 \quad \forall (l_t, m_{t'}) \in \mathscr{S}, k \in \mathscr{K}, \psi \in \Psi. \tag{3.7}$$

The objective of the model consists of two components: Constraint (3.1) describes the costs of operating costs before taking into consideration uncertainty, containing costs of operating owned services and services operated by long-term contract charters. Constraint (3.2) describes the costs with uncertainty taken into account, containing costs on services operated by the spot market, costs per shipments and costs per delayed shipments. Constraint (3.3) ensures that each shipment's volume is routed from its origin node to its destination node in scenario ψ. Rather than in models from Bai et al. (2014) and Müller et al. (2021b), where the excess variable is used to model potential outsourcing capabilities, the excess variable $Z^{k\psi}$ is used so that a shipment is able be late, bypassing the strong constraint that flow has to end at the destination terminal at the deadline. Constraint (3.4) ensures that in every scenario ψ, whenever a service is executed its associated capacity is sufficient to flow the total amount of demand that is transported via the paths that include the specific service. This includes all types of services, either owned stochastic services or spot market services. Constraint (3.5) ensures that an overall design is constructed, e.g. that even though every service arc in a different scenario ψ still starts at the same departure time. Therefore from a planning perspective there can be referred to the same service, even if the duration turns out to be different. This can be considered as a non-anticipativity constraint, a concept that rests on the theory that two or more variables should be equal to one another according to Rockafellar and Wets (1991). In this case the non-anticipativity constraint is used so that the departure time and origin-destination pair of a design variable is equal to that those of another design variable. Constraints (3.6) and (3.7) define the domains of the decision variables.

4 Progressive-Hedging framework

We use Progressive-Hedging, first introduced by Rockafellar and Wets (1991), to solve problem (3.1)-(3.7) in Section 3.2, using similar principles as Crainic et al. (2011), Crainic et al. (2014), Lanza et al. (2021), that use Progressive-Hedging for the Service Network Design problem. The idea is based on calling augmented Lagrangian techniques for decomposing the original multi-scenario problem into single-scenario sub-problems. This technique is an inexact approximation that might converge to the correct solution under certain assumptions. For this problem it is considered to be used as a heuristic. Then after iteratively solving the single-scenario sub-problems a solution is aggregated out of the single-scenario solutions, after which the solution is evaluated whether to have reached a consensus. If no consensus is met, or in the case of no identical sub-problem solutions, the decomposition parameters are adjusted so that the method leads the problem to a "well-hedged" solution. The Progressive-Hedging method is built on two different phases. The first phase is used to find consensus on using services. When some consensus is reached on variables, these variables will be included in a subset as input for the second phase. The second phase then tries to find a solution to the original problem constraint to the subset of variables. Crainic et al. (2011) and Crainic et al. (2014) have already shown that Progressive-Hedging works as an efficient method to reach good solution for a Service Network Design with uncertainty in demand, whereas Lanza et al. (2021) on its part proved it worked well for uncertain travel time in a SND.

This chapter is divided into different sections where crucial parts are elaborated on the progressive-hedging framework. Section 4.1 discusses how problem (3.1)-(3.7) is decomposed into single-scenario sub-problems. Section 4.2 gives an overview of all the crucial steps of the meta-heuristic. These steps consist of aggregating the solutions of the decomposed single-scenario sub-problems to find a global solution, discussed in Section 4.3. Section 4.4 discusses the other crucial step in the Progressive-Hedging method, namely reaching consensus, for which global and local information are taken into consideration to modify cost parameters of the sub-problems to lead the meta-heuristic towards a (near) consensus between the single-scenario solutions to a solution to the original problem. This is dealt with in.

4.1 Decomposition

Due to the stochastic nature of the service arcs in problem (3.1)-(3.7), the problem has a different structure compared to Crainic et al. (2011). This is since the formulation already contains non-anticipativity constraints to reach an overall design vector that might have different travel times but still the same departure time, as opposed to normally introducing a non-anticipativity constraint when decomposing the problem. Therefore to formulate the problem into single-scenario form we need to adjust the objective function so that the overall design variable $\bar{y}_{l_t m_{t'}}$ is replaced by the stochastic design variable $w^{\psi}_{l_t m_{t'}}$. Note that Constraint (3.4) allows to use a spot service if beneficiary for a particular scenario in order not to force a costly first stage decision to be made that an arc is opened for every scenario while only being necessary for one scenario. This means that for a single-scenario problem a spot is not likely to be chosen, since a regular service is often cheaper. Therefore the spot design variable $y^{\sigma\psi}_{l_t m_{t'}}$ is not taken into account for the single-scenario problems of the progressive-hedging meta-heuristic. The exclusion of the spot variable is not relevant since decomposition is part of the first phase. The goal of the first phase of the method is to reach consensus on using services. When there is some consensus on using a service, this information will be used for the second phase. In the second phase services with a lot of consensus will probably be used as standard services and services with little consensus might be selected as spot variables

Problem (3.1)-(3.7) now is written as

$$\text{minimize} \quad \sum_{\psi\in\Psi} p_\psi \Big(\sum_{(l_t,m_{t'})\in\mathscr{S}^\Psi} f_{lm} y^{\psi}_{l_t m_{t'}} + \sum_{k\in\mathscr{K}} \sum_{(l_t,m_{t'})\in\mathscr{S}^\Psi} c^k_{l_t,m_{t'}} x^{k\psi}_{l_t,m_{t'}} + \sum_{k\in\mathscr{K}} c^{pen}_k Z^{k\psi} \Big) \tag{4.1}$$

subject to

$$\sum_{(l_t,m_{t'})\in\mathscr{S}} x^{k\psi}_{l_t,m_{t'}} - \sum_{(m_{t'},l_t)\in\mathscr{S}} x^{k\psi}_{l_t,m_{t'}} = \begin{cases} q^{k\psi} - Z^{k\psi} & (l,t) = (o_k, e_{k\psi}) \\ -q^{k\psi} + Z^{k\psi} & (l,t) = (d_k, l_k) \\ 0 & \text{otherwise} \end{cases} \quad \forall k\in\mathscr{K}, \psi\in\Psi, (l,t)\in\mathscr{N} \tag{4.2}$$

$$\sum_{k\in\mathscr{K}} \sum_{(l_t,m_{t'})\in\mathscr{S}} x^{k\psi}_{l_t,m_{t'}} \le u_{l_t m_{t'}} y^{\psi}_{l_t m_{t'}} \quad \forall (l_t, m_{t'}) \in \mathscr{S}^\Psi, \tag{4.3}$$

$$\sum_{t'\in\mathscr{S}^\Psi} y^{\psi}_{l_t m_{t'}} = \bar{y}_{l_t m_{t'}} \quad \forall (l_t, m_{t'}) \in \mathscr{S}, \psi\in\Psi, \tag{4.4}$$

$$\bar{y}_{l_t m_{t'}}, y^{\psi}_{l_t m_{t'}}, \in \{0,1\} \quad \forall (l_t, m_{t'}) \in \mathscr{S}, \psi\in\Psi, \tag{4.5}$$

$$x^{k\psi}_{l_t,m_{t'}}, Z^{k\psi} \ge 0 \quad \forall (l_t, m_{t'}) \in \mathscr{S}, k\in\mathscr{K}, \psi\in\Psi. \tag{4.6}$$

From using the non-anticipativity constraint, it may be concluded that if Constraints (4.4) are to be relaxed, the problem (4.1)-(4.6) becomes scenario separable. By using the formulation of an overall design vector, Lagrangian relaxation can be applied to the problem. By using this method, the difference between the single-scenario solutions and the overall design solution can be penalized. Constraints (4.4) therefore can be relaxed using the augmented Lagrangian strategy discussed by Rockafellar and Wets (1991). This introduces an additional quadratic term with an additional penalty parameters while relaxing the problem. The objective of the overall problem then changes to

$$\begin{aligned}\text{minimize} \sum_{\psi\in\Psi} p_\psi \Big(&\sum_{(l_t,m_{t'})\in\mathscr{S}^\Psi} f_{lm} y^{\psi}_{l_t m_{t'}} + \sum_{k\in\mathscr{K}} \sum_{(l_t,m_{t'})\in\mathscr{S}^\Psi} c^k_{l_t,m_{t'}} x^{k\psi}_{l_t,m_{t'}} + \sum_{k\in\mathscr{K}} c^{pen}_k Z^{k\psi} \\ &+ \sum_{(l_t,m_{t'})\in\mathscr{S}^\Psi} \lambda^{\psi}_{l_t m} (y^{\psi}_{l_t m_{t'}} - \bar{y}_{l_t m_{t'}}) + \frac{1}{2} \sum_{(l_t,m_{t'})\in\mathscr{S}^\Psi} \rho (y^{\psi}_{l_t m_{t'}} - \bar{y}_{l_t m_{t'}})^2 \Big)\end{aligned} \tag{4.7}$$

where $\lambda^{\psi}_{l_t m}, \forall (l_t, m_{t'}) \in \mathscr{S}^\Psi$ define the Lagrangian multipliers for the relaxed constraints and where ρ defines the penalty ratio. Looking at the quadratic term $\rho(y^{\psi}_{l_t m_{t'}} - \bar{y}_{l_t m_{t'}})^2$ in (4.7), this term becomes $\rho(y^{\psi}_{l_t m_{t'}})^2 - 2\rho\bar{y}_{l_t m_{t'}} y^{\psi}_{l_t m_{t'}} + \rho(\bar{y}_{l_t m_{t'}})^2$, which due to the binary requirements for the design variables can be expressed as $\rho y^{\psi}_{l_t m_{t'}} - 2\rho\bar{y}_{l_t m_{t'}} y^{\psi}_{l_t m_{t'}} + \rho\bar{y}_{l_t m_{t'}}$. The objective of the relaxed problem then becomes

$$\begin{aligned}\text{minimize} \sum_{\psi\in\Psi} p_\psi \Big(&\sum_{(l_t,m_{t'})\in\mathscr{S}^\Psi} (f_{lm} + \lambda^{\psi}_{l_t m} - \rho\bar{y}_{l_t m_{t'}} + \frac{\rho}{2}) y^{\psi}_{l_t m_{t'}} + \sum_{k\in\mathscr{K}} \sum_{(l_t,m_{t'})\in\mathscr{S}^\Psi} c^k_{l_t,m_{t'}} x^{k\psi}_{l_t,m_{t'}} + \sum_{k\in\mathscr{K}} c^{pen}_k Z^{k\psi} \Big) \\ &- \sum_{(l_t,m_{t'})\in\mathscr{S}^\Psi} \lambda^{\psi}_{l_t m} \bar{y}_{l_t m_{t'}} + \sum_{(l_t,m_{t'})\in\mathscr{S}^\Psi} \frac{1}{2} \rho\bar{y}_{l_t m_{t'}}\end{aligned} \tag{4.8}$$

This objective can further be rewritten since for a given overall design vector $\bar{y}_{l_t m_{t'}}$ the sub-problem decomposes per scenario so that each single-scenario sub-problem takes a deterministic form with modified costs, given by

$$\text{minimize} \sum_{(l_t,m_{t'})\in\mathscr{S}^\Psi} (f_{lm} + \lambda^{\psi}_{l_t m} - \rho\bar{y}_{l_t m_{t'}} + \frac{\rho}{2}) y^{\psi}_{l_t m_{t'}} + \sum_{k\in\mathscr{K}} \sum_{(l_t,m_{t'})\in\mathscr{S}^\Psi} c^k_{l_t,m_{t'}} x^{k\psi}_{l_t,m_{t'}} + \sum_{k\in\mathscr{K}} c^{pen}_k Z^{k\psi} \tag{4.9}$$

subject to

$$\sum_{(l_t,m_{t'})\in\mathscr{S}} x^{k\psi}_{l_t,m_{t'}} - \sum_{(m_{t'},l_t)\in\mathscr{S}} x^{k\psi}_{l_t,m_{t'}} = \begin{cases} q^{k\psi} - Z^{k\psi} & (l,t) = (o_k, e_{k\psi}) \\ -q^{k\psi} + Z^{k\psi} & (l,t) = (d_k, l_k) \\ 0 & \text{otherwise} \end{cases} \qquad \forall k\in\mathscr{K}, \psi\in\Psi, (l,t)\in\mathscr{N} \tag{4.10}$$

$$\sum_{k\in\mathscr{K}} \sum_{(l_t,m_{t'})\in\mathscr{S}} x^{k\psi}_{l_t,m_{t'}} \leq u_{l_t m_{t'}} y^{\psi}_{l_t m_{t'}} \qquad \forall (l_t, m_{t'}) \in \mathscr{S}^\Psi, \tag{4.11}$$

$$y^{\psi}_{l_t m_{t'}} \in \{0,1\} \qquad \forall (l_t, m_{t'}) \in \mathscr{S}, \psi\in\Psi, \tag{4.12}$$

$$x^{k\psi}_{l_t,m_{t'}}, Z^{k\psi} \geq 0, \qquad \forall (l_t, m_{t'}) \in \mathscr{S}, k\in\mathscr{K}, \psi\in\Psi \tag{4.13}$$

For every scenario or the sub-problem (4.9)-(4.13), the Lagrangian multipliers $\lambda^{\psi}_{l_t m}, \forall (l_t, m_{t'}) \in \mathscr{S}^\Psi$ and ρ are used to penalize the objective in order to guide the method in choosing a design variable in a sub-problem, by penalizing the difference between the overall design vector and single-scenario design variables. These multipliers can thus be used to potentially let all single-scenario solution converge to a single design. Moreover, by adjusting the fixed costs and the multipliers of the sub-problems can help to search for consensus between all single-scenario designs. The aggregation process on how a single design vector is obtained is discussed in Section 4.3 and how the adjustment of the cost parameters is used to find a consensus is discussed in Section 4.4.

4.2 Overview of Matheuristic

Algorithm 1 involves the complete matheuristic and involves the 3 steps of the first phase of Progressive-Hedging, namely solving the decomposed single-scenario sub-problems (discussed in Section 4.1), aggregating their solutions to an overall design (discussed in Section 4.3) and adjusting the fixed costs of the sub-problems according two strategies, that are the Lagrangian strategy (Subsection 4.4.1) and the Heuristic strategy (Subsection 4.4.2). Since the goal is to find consensus on a design among all scenarios, the framework consists of two phases. The first phase is properly seeking consensus until the stopping criteria is met. This stopping criteria consists of a certain amount of running time, achieving an amount of iterations without improvement on the design and a maximum amount of iterations. After one or more of the stopping criteria are met, the goal of the second phase is to find a solution to the original problem with a subset of arcs for which some consensus is achieved. In Crainic et al. (2011) and Lanza et al. (2021) consensus was easier to find, so only a small subset with strict thresholds was taken into account. For this problem however spot resources are also considered, and they might only be beneficial in a limited set of scenarios. Therefore arcs for which even a small amount of consensus is achieved are considered in the second phase, since these arcs might be useful as spot resources and might give additional penalties to not delivering a shipment when disregarding this opportunity of a service. In the first phase, the first step of the framework is to find initial solutions to the single-scenario problems. After the initial solutions are found, the program starts iterating and finds new solutions keeping track of the solutions of the previous iteration. This means that each single-scenario sub-problem represents a deterministic SND with modified costs according to adjustment strategies *Lagrangian* or *Heuristic*.

Several methods can be used to solve the single-scenario sub-problems, for example with heuristical methods such as a cycle-based Tabu-Search used by Crainic et al. (2011). For this research however a commercial solver is used, in particular Gurobi.

Algorithm 1 Progressive-Hedging based framework

procedure FIRST PHASE
  Initialization
  $\nu \leftarrow 0$;
  if Lagrangian = TRUE **then**
    $\lambda_{l_t m}^{\psi\nu} \leftarrow 0, \forall (l_t m_{t'}^{\psi} \in \mathscr{S}^{\Psi})$;
    $\rho^{\nu} \leftarrow \rho^{0}$;
  for $\psi \in \Psi$ **do**
    $f_{ltm}^{\psi\nu} \leftarrow f_{lm}, \forall (l_t, m_{t'}) \in \mathscr{S}$;
    Solve the corresponding single-scenario sub-problem;
  $\bar{y}_{l_t m_{t'}}^{\nu} \leftarrow \sum_{\psi\in\Psi} p_{\psi} y_{l_t m_{t'}}^{\psi\nu}, \forall (l_t, m_{t'}^{\psi}) \in \mathscr{S}^{\Psi}$;
  while stopping criterion not met **do**
    $\nu \leftarrow \nu + 1$;
    if Heuristic = TRUE **then**
      Adjust globally $f_{ltm}^{\psi\nu}, \forall (l_t, m_{t'}) \in \mathscr{S}$ using (4.17);
    for $\psi \in \Psi$ **do**
      if Lagrangian = TRUE **then**
        $f_{ltm}^{\psi\nu} \leftarrow f_{lm} + \lambda_{l_t m}^{\psi\nu-1} - \rho^{\nu-1}\bar{y}_{l_t m_{t'}}^{\nu-1} + \frac{\rho^{\nu-1}}{2}, \forall (l_t, m_{t'}) \in \mathscr{S}$;
      else if Heuristic = TRUE **then**
        Adjust locally $f_{ltm}^{\psi\nu}, \forall (l_t, m_{t'}) \in \mathscr{S}$ using (4.18);
      Fix some $y_{l_t m_{t'}}^{\psi\nu}$ if appropriate;
      Solve the corresponding single-scenario sub-problem;
    Update $\bar{y}_{l_t m_{t'}}^{\nu} \leftarrow \sum_{\psi\in\Psi} \phi_{\psi} y_{l_t m_{t'}}^{\psi\nu}, \forall (l_t, m_{t'}^{\psi}) \in \mathscr{S}^{\Psi}$
    if Lagrangian = TRUE **then**
      $\lambda_{l_t m}^{\psi\nu} \leftarrow \lambda_{l_t m}^{\psi\nu-1} + \rho^{\nu-1}(y_{l_t m_{t'}}^{\psi\nu} - \bar{y}_{l_t m_{t'}}^{\nu-1}), \forall (l_t, m_{t'}^{\psi}) \in \mathscr{S}^{\Psi}$;
      $\rho^{\nu} = \alpha\rho^{\nu-1}$;
procedure SECOND PHASE
  Fix design for which almost consensus is achieved;
  Solve the restricted multi-scenario problem;

4.3 Aggregation

The goal of this step of the progressive-hedging framework is to use the local information from the design solutions of the single-scenario sub-problems to generate an overall design that can be used for all scenarios. Section 4.2 discusses the framework that solves sub-problem (4.9)-(4.13) for every scenario over multiple iterations, so let ν be the index counting the number of iterations of the framework. Then $\bar{y}_{l_t m_{t'}}^{\nu}, \forall (l_t, m_{t'}) \in \mathscr{S}$

is the overall design that is obtained by aggregating $y^{\psi\nu}_{l_t m_{t'}} \forall \psi \in \Psi, (l_t, m_{t'}) \in \mathscr{S}^{\Psi}$, that are the solutions for every single-scenario sub-problem. Note that $\mathscr{S}$ is used for the overall design and that $\mathscr{S}^{\Psi}$ is used for the single-scenario sub-problems, meaning that we have a difference in the "planned" travel time between an origin-destination pair and the "stochastic" travel time in a certain scenario ψ. In order to aggregate services to a single design the departure at the origin terminal l_t and the destination terminal m are used as indexes since $m_{t'}$ might be different for a given scenario. The aggregation is defined by a function that combines the single-scenario solutions by using the weight (or probability) of a scenario ψ. Then the aggregation operator is given by

$$\bar{y}^{\nu}_{l_t m_{t'}} = \sum_{\psi \in \Psi} p_{\psi} y^{\psi\nu}_{l_t m_{t'}}, \forall (l_t, m_{t'}) \in \mathscr{S}^{\Psi}, \tag{4.14}$$

From (4.14) it should be noted that it may be possible that a non-feasible design is constructed. Feasibility can only be obtained if the design over all scenarios has reached consensus, and therefore $\bar{y}^{\nu}_{l_t m_{t'}} \in \{0,1\}$. In the case where there is no consensus it may be observed that $0 < \bar{y}^{\nu}_{l_t m_{t'}} < 1$, meaning that due to the integrality constraints on the design variables that the problem is infeasible. As Rockafellar and Wets (1991) also noted the overall solution strategy thus may not converge to an optimal solution, and therefore it cannot be guaranteed that a good solution will be obtained. However Crainic et al. (2011) and Lanza et al. (2021) mentioned that if the value of $\bar{y}^{\nu}_{l_t m_{t'}}$ is close to 0 or close to 1, there might be a trend toward either not selecting or selecting the overall design variable. This means that (4.14) will be used in the first phase of the framework where it is investigated if a subset of service arcs can reach consensus. Crainic et al. (2011) and Lanza et al. (2021) then fix these service arcs for the restricted problem. However for this research the arcs for which consensus has not been reached might be useful for the second phase, since non zero design variables might not be useful for the overall design but can indicate that the design can be used for a spot departure. Therefore for this research all $\bar{y}^{\nu}_{l_t m_{t'}} > 0$ will be used for the second phase, having the confidence that all unnecessary arcs will be equal to zero after enough iterations of the first phase of the framework.

4.4 Search for consensus

After the original problem (3.1)-(3.7) is decomposed into single-scenario sub-problems (4.9)-(4.13), and using (4.14), one can search for consensus among the single-scenario sub-problems by adjusting the fixed costs of each sub-problem to penalize the difference between the single-scenario designs and the overall design. For the adjustment of the fixed costs two strategies are discussed, namely a Lagrangian strategy in Subsection 4.4.1 and a Heuristic strategy in Subsection 4.4.2.

4.4.1 Lagrangian strategy

The Lagrangian strategy is the basic strategy for progressive-hedging, based on the augmented Lagrangian method described by Rockafellar and Wets (1991) and furthermore used by Crainic et al. (2011), Lanza et al. (2021) to solve Stochastic Service Network Design problems. The Lagrangian strategy modifies the Lagrangian multipliers and the ρ penalty parameter associated to the relaxed non-anticipativity constraints in (4.9). Because of the integrality of the decision variables, the local convexity and global convergence properties of the augmented Lagrangian method do not apply, meaning that the Langrangian strategy can be considered as a heuristic strategy. For this strategy the modifications aim to open and close a service arc when its status is different than that of the overall design, by slightly increasing the penalty factor to gain a stronger incentive to open or close the arc. The modifications of the multipliers rest on the fact that there may be three possible adjustment for a service arc in a scenario ψ for at iteration ν, that are:

- $y^{\psi\nu}_{l_t m_{t'}} < \bar{y}^{\nu-1}_{l_t m_{t'}}$. Here the service arc is closed in the single-scenario case but the value of the overall design is not zero. Therefore the trend is towards opening this service arc, thus reducing the fixed costs of this arc in the single-scenario sub-problem and thus give the incentive to open the arc.

- $y^{\psi\nu}_{l_t m_{t'}} > \bar{y}^{\nu-1}_{l_t m_{t'}}$. Here the service arc is opened in the single-scenario case but the value of the overall design is lower, meaning that overall the service arc is opened less in other scenarios. Therefore the trend is towards closing this service arc, thus increasing the fixed costs of this arc in the single-scenario sub-problem and thus give the incentive to close the arc.
- $y^{\psi\nu}_{l_t m_{t'}} = \bar{y}^{\nu-1}_{l_t m_{t'}}$. This means that there is consensus among all single-scenario designs with regards to opening or closing this service arc and therefore the fixed costs remain the same.

Regarding the cases in which adjustment is possible, the Lagrangian multiplier and penalty ratio parameter ρ are modified by

$$\lambda^{\psi\nu}_{l_t m} \leftarrow \lambda^{\psi\nu-1}_{l_t m} + \rho^{\nu-1}(y^{\psi\nu}_{l_t m_{t'}} - \bar{y}^{\nu-1}_{l_t m_{t'}}), \forall (l_t, m_{t'}) \in \mathscr{S}^{\Psi}, \tag{4.15}$$

$$\rho^{\nu} = \alpha\rho^{\nu-1}, \tag{4.16}$$

where ρ^0 is set to a small positive value, where ν is the iteration counter and where $\alpha > 1$ which makes for a slow increase in the penalty term. ρ^0 should not be chosen too large since the framework might converge to a possibly far from optimal design. Again, because of the stochastic implications of the duration of the service arcs the modification of (4.15) is done by looking at the origin-destination pair and the departure time at the origin.

4.4.2 Heuristic strategy

The Heuristic strategy is based on the Heuristic strategies used by Crainic et al. (2011) and Lanza et al. (2021), where at each iteration the fixed costs of arcs are modified with a status different from what a large amount of the other arcs agree upon at the current iteration. Similar to the Lagrangian strategy the goal is to get consensus on a design between all single-scenario sub-problems. Within the Heuristic strategy, two types of adjustments are used, one at a global level where the overall design is evaluated and at a local level for each scenario. For the global adjustment the overall design at iteration $\nu, \bar{y}^{\nu}_{l_t m_{t'}}$ aims to favor the current trend for opening or closing service arcs in the single-scenario designs. A low value of $\bar{y}^{\nu}_{l_t m_{t'}}$ might indicate that an arc is not favorable to be used, whereas a high value might indicate that an arc is useful in the design. This then can be approached as the relation that fixed costs either have to be increased when the value of $\bar{y}^{\nu}_{l_t m_{t'}}$ is lower than a given c^{low} or have to be decreased when the value of $\bar{y}^{\nu}_{l_t m_{t'}}$ is higher than a given c^{high}. This is to either distract a single-scenario sub-problem from using a service arc or attract a sub-problem to use a service arc. This results in a global adjustment strategy given by

$$f^{\nu}_{ltm} = \begin{cases} \beta f^{\nu-1}_{ltm} & \text{if} \bar{y}^{\nu-1}_{l_t m_{t'}} < c^{low}, \\ \frac{1}{\beta} f^{\nu-1}_{ltm} & \text{if} \bar{y}^{\nu-1}_{l_t m_{t'}} > c^{high}, \\ f^{\nu-1}_{ltm} & \text{otherwise}, \end{cases} \tag{4.17}$$

where $\beta > 1, 0 < c^{low} < 0.5$, and $0.5 < c^{high} < 1$, and f^{ν}_{ltm} the modified costs of an arc from l to m at departure time t in all scenarios. Due to the stochastic behaviour of the "stochastic" service arcs the departure time t is taken as a reference next to the origin-destination pair.

For the local adjustment of the fixed costs the difference between the values of the local design $y^{\psi\nu}_{l_t m_{t'}}$ and the current overall design $\bar{y}^{\nu}_{l_t m_{t'}}$ is considered. This adjustment is only made if there is a large enough difference

(c^{far}) between the local design and the overall design, given by

$$f_{ltm}^{\psi\nu} = \begin{cases} \beta f_{lm}^{\nu} & \text{if} |y_{l_t m_{t'}}^{\psi\nu-1} - \bar{y}_{l_t m_{t'}}^{\nu-1}| \geq c^{far} \text{and} y_{l_t m_{t'}}^{\psi\nu-1} = 1, \\ \frac{1}{\beta} f_{ltm}^{\psi\nu} & \text{if} |y_{l_t m_{t'}}^{\psi\nu-1} - \bar{y}_{l_t m_{t'}}^{\nu-1}| \geq c^{far} \text{and} y_{l_t m_{t'}}^{\psi\nu-1} = 0, \\ f_{lm}^{\nu} & \text{otherwise,} \end{cases} \tag{4.18}$$

where $\beta > 1$ and $0.5 < c^{far} < 1$, represents the threshold defining a local adjustment is to be applied. $f_{ltm}^{\psi\nu}$ stands for the modified local cost of an arc from l to m departing at time t in scenario ψ at iteration ν. Same as done by Crainic et al. (2011), the idea of local adjustment can be further extended to speed up the algorithm by looking at possibilities of "fixing" part of the structure. Looking at $|y_{l_t m_{t'}}^{\psi\nu-1} - \bar{y}_{l_t m_{t'}}^{\nu-1}| \leq c^{near}$, meaning that the difference between the local design and the overall design is significantly low, e.g. below a certain low threshold $0 < c^{near} < 0.5$, it may be stated that one does not want to try and modify the design arc in the next iteration. Therefore, in order to take advantage of this information, the local design variable of the following iteration can be fixed by setting $y_{l_t m_{t'}}^{\psi\nu} = y_{l_t m_{t'}}^{\psi\nu-1}$, so that a smaller sub-problem is solved and thus may be solved more quickly.

5 Computational Experiments

In order to validate the model formulation proposed in Chapter 3 and compare its performance to the Progressive Hedging method discussed in Chapter 4, experiments are done that are discussed in this section. First the method on generating the small test instances are discussed in Section 5.1 after which the results of the methods on these instances are shown in Section 5.2 followed by a discussion on the results in Section 5.3.

5.1 Generating Instances

In order to generate instances two different types of sets were generated, the first being generated with a relative small amount of shipments in the instances yet a relative high amount of nodes and scenarios. The second set consists of instances where the amount of nodes and scenarios are relatively small and the amount of shipments is relatively large compared to the other instances set. The different instances are based on a network with 5 or 8 nodes, which can be considered as a relevant amount considering the size of many logistical clusters observed from DB Schenker. Arcs between the nodes are generated with the Euclidean distances between the coordinates of the nodes and are converted to travel time by using an average speed. For simplicity, this expected speed is for every arc 60 km/h. The actual speed differs between scenarios, following a Gamma distribution with the travel time as mean, in line with Chen and Fan (2020) and Lanza et al. (2021). The capacity of an arc is the capacity of a box trailer used by DB Schenker, 25150 kg. Costs of opening a service-arc and costs of shipments are taken as a factor of the travel duration that is in such balance that the costs of shipments are not too expensive so that consolidation still occurs (when costs of a shipment are too high, delivering a shipment through the least amount of transshipment proves favorable). For the time frame 10 days are considered, with time steps of 1 hour. Shipments are based upon original data from DB Schenker. These shipments are taken from the first two days of a week, so that consolidation is still possible with the little amount of shipments being used for the instances. This is because if an instance with only small amount of shipments divided over a whole week there is low possibility to test consolidation. Scenarios are taken mostly as the same shipments, with some alteration in volume and ready/due time. Costs of opening an arc in the network and costs of shipments are factors of the planned travel times between two terminals.

With all these considerations taken into account 6 instances are generated for instance set **A** with 5 and 8 nodes, with 8, 12 and 16 shipments and for 5 and 10 scenarios. The results on different instances of set **A** are listed in Table 5.1, where the name of the instance A_n_k_ψ refers the the instance with $n \in \mathscr{N}$ nodes, $k \in \mathscr{K}$ shipments and $\psi \in \Psi$ scenarios. A different instance set **B**, also consisting of 6 instances is generated with 3, 5 and 6 nodes, with 10, 20, 30, 40 and 60 shipments, where most of these instances consist of 5 different scenarios and 2 instances that consist of 10 different scenarios. The results on different instances of set **B** are listed in Table 5.4, where where the name of the instance B_n_k_ψ refers the the instance with $n \in \mathscr{N}$ nodes, $k \in \mathscr{K}$ shipments and $\psi \in \Psi$ scenarios.

5.2 Algorithmic performance

For the experiments on instance set **A** from Section 5.1 the experimental setup and results are discussed in this section. Both the mathematical model and the Progressive Hedging method are coded using Python (version 3.9.13), with the optimization being performed by Gurobi (version 10.0.1). The experiments were conducted using a Lenovo Thinkpad T14 Gen2 with Intel i5 2.4 GHz 4 core processor with 16 GB RAM.

A time limit is set for both methods for 10 hours. For the Progressive Hedging the other stopping criteria are set to a limit of 10 iterations without improvement and a maximum number of iterations of 50. Furthermore,

	Gurobi				PH-Lagrangian (gap 1 %)			PH-Heuristic (gap 1 %)		
Instance	**L.B.**	**Val.**	**Gap**	**Time [min:sec]**	**Val.**	**Gap**	**Time [min:sec]**	**Val.**	**Gap**	**Time [min:sec]**
A_5_8_5	13730.3	13730.3	0.00 %	5:37	13730.3	0.00 %	20:49	13768.7	0.00 %	4:06
A_5_8_10	24885.8	24885.8	0.00 %	31:19	24885.5	0.00 %	58:09	24885.8	0.00 %	10:42
A_8_12_5	27106.5	27106.5	0.00 %	119:28	27176.4	0.25 %	145:44	27133.5	0.09 %	18:03
A_8_12_10	38177.6	38262.0	0.22 %	601:04	38331.9	0.40 %	364:27	38299.3	0.31 %	38:23
A_8_16_5	39275.9	39292.3	0.04 %	603:20	39398.4	0.31 %	195:07	39295.5	0.04 %	20:42
A_8_16_10	50389.9	50447.8	0.11 %	601:19	50553.9	0.32 %	389:40	50478.0	0.17 %	146:30

Table 5.1: Results on instance set **A**.

	PH-Lagrangian (gap 1 %)			PH-Heuristic (gap 1 %)			PH-Lagrangian (gap 5 %)			PH-Heuristic (gap 5 %)		
Instance	**Val.**	**Gap**	**Time [min:sec]**	**Val.**	**Gap**	**Time [min:sec]**	**Val.**	**Gap**	**Time [min:sec]**	**Val.**	**Gap**	**Time [min:sec]**
A_5_8_5	13730.3	0.00 %	20:49	13768.7	0.00 %	4:06	13730.3	0.00 %	13:09	13730.3	0.00 %	3:19
A_5_8_10	24885.5	0.00 %	58:09	24885.8	0.00 %	10:42	24885.8	0.00 %	38:02	24885.8	0.00 %	8:50
A_8_12_5	27176.4	0.25 %	145:44	27133.5	0.09 %	18:03	27176.4	0.25 %	44:40	27176.4	0.25 %	9:30
A_8_12_10	38331.9	0.40 %	364:27	38299.3	0.31 %	38:23	38331.9	0.40 %	127:19	38331.9	0.40 %	26:17
A_8_16_5	39398.4	0.31%	195:7	39295.5	0.04 %	20:42	39398.4	0.31 %	53:41	39398.4	0.31 %	8:15
A_8_16_10	50553.9	0.32 %	389:04	50478.0	0.17 %	146:30	50553.9	0.32 %	160:23	50553.9	0.32 %	27:02

Table 5.2: Difference between Progressive Hedging on sub problem gap of 1 % and 5 % on instance set **A**.

$\beta = 1.1$, $c_{high} = 0.8$, $c_{low} = 0.2$, $c_{far} = 0.6$, $c_{near} = 0.2$ and $\rho_0 = 5$. These values are similar to the values set by Crainic et al. (2011) except for c_{far}, which is set a little lower to be able to not exclude arcs that might be useful in small amount of scenarios. For the Progressive Hedging method both the Lagrangian and the Heuristic are tested, where first a comparison is made between the performance of Gurobi and the performance of both the Progressive Hedging methods. Gurobi is mainly used to test the gap of the solution methods to a lower bound, since Gurobi is able to find a lower bound for every instance. For the first comparison the gap for the single-scenario sub-problems is set to 1 %. It proved that when trying to solve the single scenario sub problems to optimality, the algorithm got stuck in local optima. For phase two of the Progressive Hedging framework no gap was set, yet the second phase proved to find a good solution (0 % gap to the restricted multi-scenario problem) for every instance within short amount of time of around 1 seconds for every instance. A second comparison is made where the gap value of the single scenario sub problems is changed to a higher value to test the objective value compared to the lower bound and the run time to reach the objective value.

Table 5.1 shows the performance results of the solution methods applied to instance set **A**, with for Gurobi the L.B., Val., Gap and Time values represent respectively the best found lower bound, upper bound, gap between these values and the total computation time in minutes. For the PH-Lagrangian and PH-Heuristic methods both the Val., Gap and Time values represent respectively the best objective value found by the method, the gap value to the lower bound found by Gurobi and the total run time of the method. Another comparison between the two PH methods where a difference in gap parameter for Phase 1 of 1 % and 5% is shown in Table 5.2. At last Table 5.3 shows the results for the PH methods with a gap of 10 % on the single scenarios, as well as the EEV solutions of the instances from instance set **A**, as stated by Birge and Louveaux (1997), EEV being the *expected result of using the EV solution*, where EV on itself is the *Expected Value problem*.

For the experiments on instance set **B** the same setup is used as for experiments on instance set **A**. From the results of the experiments done on instance set **A** the PH methods with a gap allowance of 5 % for the single scenario sub problems proved to be very time efficient with little drop in objective performance compared to the other methods, therefore PH-Lagrangian (gap 5 %) and PH-Heuristic (gap 5 %) are used for a performance comparison to Gurobi. For the experiments done on instance set **B** a time limit of 3 hours is set instead of the 10 hours set for instance set **A**. Similar to the results from instance set **A**, Table 5.4 shows the performance results of the solution methods applied to instance set **B**, with for Gurobi the L.B., Val., Gap and Time values represent respectively the best found lower bound, upper bound, gap between these

	PH-Lagrangian (gap 10 %)			PH-Heuristic (gap 10 %)			EEV
Instance	**Val.**	**Gap**	**Time [min:sec]**	**Val.**	**Gap**	**Time [min:sec]**	**Val.**
A_5_8_5	13730.3	0.00 %	20:49	13730.3	0.00 %	4:38	13828.7
A_5_8_10	24885.5	0.00 %	54:54	24885.8	0.00 %	13:31	24950.2
A_8_12_5	27176.4	0.25 %	58:30	27176.4	0.25 %	14:11	27153.9
A_8_12_10	38331.9	0.40 %	164:02	38331.9	0.40 %	33:32	38302.6
A_8_16_5	39398.4	0.31%	73:45	39398.4	0.31 %	14:33	39332.1
A_8_16_10	50553.9	0.32 %	202:18	50533.9	0.32 %	36:33	50494.4

Table 5.3: Progressive Hedging on sub problem gap of 10 % and EEV on instance set **A**.

	Gurobi				PH-Lagrangian (gap 5 %)			PH-Heuristic (gap 5 %)			EEV
Instance	**L.B.**	**Val.**	**Gap**	**Time [min:sec]**	**Val.**	**Gap**	**Time [min:sec]**	**Val.**	**Gap**	**Time [min:sec]**	**Val.**
B_3_10_5	18013.6	18013.6	0.00 %	0:14	18013.6	0.00 %	1:36	18013.6	0.00 %	1:05	18013.6
B_3_20_5	40369.3	40369.3	0.00 %	1:15	40369.3	0.00 %	10:28	40369.3	0.00 %	1:46	40376.1
B_5_20_5	72889.3	72940.7	0.07 %	180:25	72968.1	0.10 %	25:27	72968.1	0.10 %	4:12	72997.7
B_5_40_5	149553.4	149602.7	0.03 %	185:50	149602.7	0.03 %	47:01	149602.7	0.03 %	7:06	149689.1
B_6_30_5	126095.7	126158.0	0.04 %	181:06	126158.0	0.04 %	54:25	126158.0	0.04 %	8:11	126174.4
B_6_30_10	126093.4	126158	0.05 %	183:27	126158.0	0.05 %	156:09	126158.0	0.05 %	25:43	126174.4
B_6_60_5	140769.5	140831.9	0.04 %	181:19	140831.9	0.04 %	73:28	140831.9	0.04 %	11:51	140859.1
B_6_60_10	140660.8	140831.9	0.01 %	189:16	140859.1	0.01 %	190:15	140831.9	0.01 %	32:26	140859.1

Table 5.4: Results on instance set **B**.

values and the total computation time in minutes. For the PH-Lagrangian (gap 5 %) and PH-Heuristic (gap 5 %) methods both the Val., Gap and Time values represent respectively the best objective value found by the method, the gap value to the lower bound found by Gurobi and the total run time of the method. Next to that the EEV value for the instances is shown.

5.3 Discussion on Algorithmic comparison

Of all the methods, PH-Heuristic (5 %) performs the best when looking at efficiency and quality. From the results on both instance sets **A** and **B** it can be seen that the more scenarios and more nodes are considered, the problem takes significantly more time to solve for all methods. Gurobi hits the time limit from instances with 8 nodes and from 12 shipments and 10 scenarios. Both Progressive Hedging methods do not reach the time limit, however reach their stopping condition of 50 iterations. For the smaller instances, Gurobi still performs well compared to the Progressive Hedging methods, able to find a Gap of 0 % with relatively small run time compared to PH-Lagrangian method. PH-Heuristic is already able to find solutions in a smaller amount of time. For the larger instances, both PH methods significantly outperformed Gurobi regarding run time, while both PH methods are still able to find a good objective value close to the lower bound. Of these two PH methods however, PH-Heuristic significantly outperforms PH-Lagrangian when it comes to run time and objective value. This is because both strategies adjust the cost structure, however within the Heuristic strategy we also have the possibility to fix variables if the trend is strong towards either opening or closing the service arc. With fixing variables the problem is restricted and therefore a solution can be found more quickly. The EEV solutions for both instance sets are relatively close to the solutions of the PH methods. One reason could be that we have too much correlation in the instances and too little variation between the scenarios. This could happen due to the rounding of the travel time to hours, leading to less variation in travel time between scenarios.

6 Case Study

This chapter discusses the case study of this thesis. For the case study the solution method described in Chapter 4 will be tested on a (a part of the) real life network of DB Schenker. Section 6.1 describes the context of the company. Section 6.2 describes the scope of the case study. Section 6.3 describes all necessary data preparation steps taken. Section 6.4 discusses the experimental study performed on the company data and the results of the case study.

6.1 Company context

DB Schenker is one of the world's leading global logistics providers, by providing service to distribute goods through land transport, worldwide air and ocean freight. For this thesis only the land transport is considered. The company makes use of a hybrid hub-and-spoke network for their distribution. In this hybrid network, LTL operations are used between hubs and FTL operations can be used either between hubs, or between pick-up location and delivery location, or between pick-up location and destination hub or between origin hub and delivery location. The network of DB Schenker can be considered as a hybrid hub-and-spoke network, since they use both system services and direct services, with additional part-load services. For their system freight service they consolidate demand at the hub, ship the demand to the destination hub and further distribute from the destination hub. For direct services they consider large enough shipments or valuable shipments that need to be routed directly to the destination. They also work with part-load options, here a small amount of large size shipments is combined and picked up at the client and is routed directly to the destination hub, but do not have to be consolidated at the origin hub. DB Schenker uses a weekly reoccurring cyclic schedule wherein they plan their "line-haul" operations. In order to investigate how the company operates their line-haul planning, we first need to define the context of the research.

Line-hauls are planned at strategic level. These line-hauls often operate at daily basis when there is high demand between two hubs, weekly when there is less demand. Line-haul schedules are kept in the ENP (European Network Plan) of DB Schenker. This plan consist of the DLE's (Daily Line-haul Execution). The ENP stays fixed except for when there is a holiday within a destination or drive-through area, or in case of severe road-works impacting the planned duration of a trip. This means that there is cyclic behaviour in the line-haul planning. Once there is a disruption, this often impacts the concurrent line-haul. Line-hauls at strategic level are managed on a weekly basis. Every week the statistics of the shipment volumes are analysed. It can happen that for example there is a steady growth between two areas so it might be necessary to open a line-haul. It might also happen that due to a new client there is sudden growth in volume, therefore resulting in a new line-haul having to be opened. Line-hauls at operational level are managed on daily basis. Status of a line-haul is being measured via LiNeS, a system that captures the departure times, arrival times, planned duration of a trip, type of truck, whether the truck has GPS. A line-haul can be operated by either an owned resource, a long-term contract charter or a spot (resource acquired from spot market).

Demand being shipped with line-hauls is first picked at the customers and then consolidated at the hubs, before being transported by means of the line-hauls. This problem is called CoDi at DB Schenker. Each branch within the BeNeLux cluster has its own CoDi area, with the exception of a few areas that are shared by often 2 branches. It also happens that there are direct shipments when a shipment has high value or when the size of the shipment is already sufficient to fill a (full) truck. It can also happen that a shipment is shared with another shipment when they have a sufficient combined volume to fill a (full) truck. Demand is updated throughout the day, and are scheduled on a line-haul shipment by a line-haul planner through the ITMS (Integrated Transportation Management System). This system is used to keep track of every shipment once it is being ordered, through arriving at the hub to be loaded into a truck, all the way to the their destination. It can happen that a shipment is being ordered on time and is being scheduled to be transported on the same day, yet arriving late on the hub to be loaded into a truck. It depends on the priority of the shipment if this late CoDi shipment is being postponed to the next day or that something is done to still be able to ship this order.

Planning of drivers/trucks is dealt with on a daily basis. The goal is to find a schedule with chartered transport companies so that they are able to reach a certain amount of monthly kilometers in order to reach a discount on the kilometer rate DB Schenker is being charged. It has be taken into account that drivers need their mandatory breaks after a certain amount of hours driving. This may lead to unwanted complexities in the planning. Trucks can be equipped with either one or two drivers. The difference is that with two drivers either driver is able to take a mandatory break while the other driver is driving.

There are several grades within shipments, with each having a higher priority compared to the others. Premium shipments not being delivered on time will result in the shipment being free of charge for the customer. Standard shipments will have an extra day slack for it to be delivered. Once a line-haul is delayed by more than 15 minutes, the branch operating the line-haul will be fined to the destination branch of the line-haul. The fine depends on the delivery day of the week, being more expensive for a late line-haul on Mondays than for a late line-haul for the rest of the week. This difference is due to more activity in the warehouses on Mondays and due to disruptions on Mondays leading to disruptions further down the week. For the context of this research, we are only interested in the export line-hauls operated from the BeNeLux cluster from DB Schenker.

The literature on the design of a hub-and-spoke network largely coincides with the literature on a Service Network Design, where design is made of a network with as most important constraint that the demand must be fulfilled by flowing from the origin to the destination. Moreover, the demand is consolidated at the origin and destination terminals. This means that the Service Network Design problem with taking into account uncertainty in demand and uncertainty in travel time can be used to help gain insights in potential improvements in the operations of DB Schenker, and might be able to help DB Schenker to develop a more robust line-haul schedule.

6.2 Scope

The domain of the research is the Land Freight Transportation department of DB Schenker, by focusing on the export line-hauls of the BeNeLux cluster. Due to complexity however the choice is made to only consider export line-hauls from the three Dutch terminals, Oldenzaal, Ede and Tilburg. A very large proportion of DB Schenker's operations consists of LTL (Less than Truckload) shipments, where shipments are consolidated at a hub before it is being shipped to another hub. DB Schenker also carries out FTL (Full Truckload) shipments, where large enough shipments are being picked up at costumers and are directly being shipped to a hub or customer without involving a hub to consolidate shipments with other shipments. There are some uncertain factors in the optimization of the planning process, some of which are included in the research: uncertainty in volume, uncertainty in travel time. Some uncertain factors are excluded in the research: uncertainty in availability of drivers, uncertainty in availability of trailers, uncertainty in availability of shipments being loaded on trucks at the hubs, uncertainty in service time at the hubs. Line-hauls can be operated by drivers of DB Schenker, long term contract charters or spots. Spot departure are considered to be part of the scope, since operating a line-haul with different type of operators result in a difference in cost. Long term contract charters are out of scope. Variability in these costs are however out of scope. Handling of goods at the hubs is considered out of scope, however the time an order is ready to be transported and the due time an order must be delivered at a hub are taken into consideration. There is a difference in types of trailers used, being swap-body, box, and mega trailers. This difference results in different capacity of the trailers. However the significantly largest part of trailers used are box-trailers, so for the research only the box type is considered for the capacity. Moreover, shipments can be premium or normal. This difference in order type can result into a difference in costs when a line-haul is delayed and a difference in when a shipment has to be delivered, so the order type is considered to be part of the scope. Factors that are considered to be part of the scope:

- LTL (Less than Truckload) operations.
- Type of order (either premium or standard).
- Type of operator (either self or spot)
- Uncertainty in volume.
- Uncertainty in travel time.

6.3 Data Preparation

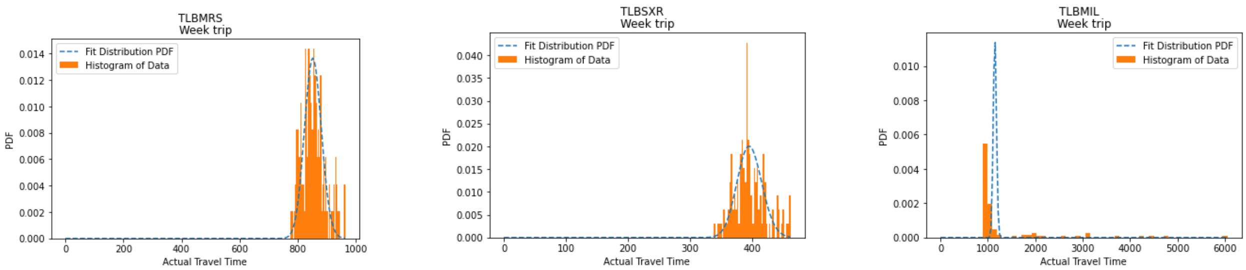
In order to use the method proposed in Chapter 4 to help find improvements in the operations of DB Schenker data from existing shipments and line-haul operations have to be used. In order for the data to be useful, first some preparation has to be done. In the subsections of this section the additional preparation steps apart from the standard steps are elaborated.

Data considering the line-haul operations contain information about planned departure/arrival of a line-haul, actual departure/arrival of a line-haul, planned duration and distance of a trip, actual duration and distance of a trip, origin/destination of a line-haul. From this information on the trips an analysis of the travel times between hubs can be performed to potentially find a distribution of the travel time. This is further elaborated in Subsection 6.3.1. However, there is only information on travel time of existing planned line-hauls, meaning that information on travel time between hubs where no line-haul is being operated is unknown. Therefore an approximation of this unknown travel time is elaborated in Subsection 6.3.2.

When looking at the shipment data, information like origin/destination postal-code, origin/destination hub, origin/destination country, type of shipment (premium or standard) and different measures of size are kept (gross weight, cubic meters, loading meters or tax weight). The difficulty of using different types of size measures is that sometimes a truck can be fully loaded according to a specific size measure, however in reality still has capacity left. Moreover, of several shipments not every size measure is kept. Therefore for simplicity only one size measure is taken into account, and since most of the data entries contain the size measure tax weight this is the size measure is used for determining the shipment size. Events such as when a shipment is registered to be picked up at a client and when a shipment is collected at a hub are also tracked, from which the collection event is useful since this essentially describes when a shipment is ready to be transported to the destination hub. The due time of the shipment however is not precisely known, so therefore an approximation is made in Subsection 6.3.3.

6.3.1 Existing travel time information

For the analyses of existing line-hauls to analyze the travel time of these line-hauls a fit is made of the distributions. According to Chen and Fan (2020), often the distribution of travel times of trips can be fitted with a gamma distribution, since it is often difficult to arrive much earlier than the average travel time but it can happen that the "delayed" trips have a duration far from the average but do not have that much similar occurrences. To test this theory every line-haul is fitted with the gamma distribution of the travel time in minutes, of which 3 representations are shown in Figure 6.1. Here a histogram of recent trips is shown as well as the fitted gamma distribution. Many of the line-hauls fitted with the gamma distribution have similar characteristics as the line-hauls **TLBMRS** and **TLBSXR** shown in Figure 6.1a and Figure 6.1b, where it can be seen that the fitted distribution represents the histogram of the travel time data well. At last some



(a) TLBMRS week trip **(b)** TLBSXR week trip **(c)** TLBMIL week trip

Figure 6.1: Fitted gamma distributions for existing line-hauls.

line-hauls show similar characteristics like line-haul **TLBMIL** shown in Figure 6.1c, where there clearly is a similar behaviour to a gamma distribution, however not fitted well due to a lot of other observations that have a much longer travel time than the mean. Therefore the mean of the fitted distribution is a bit more than the seemingly mean of the histogram. As can be seen from the three figures, there are some arguments to either accept or reject using the gamma distributions to fit the travel times of line-hauls. For this case study however, it is assumed that the gamma distribution can be used for all existing travel time data.

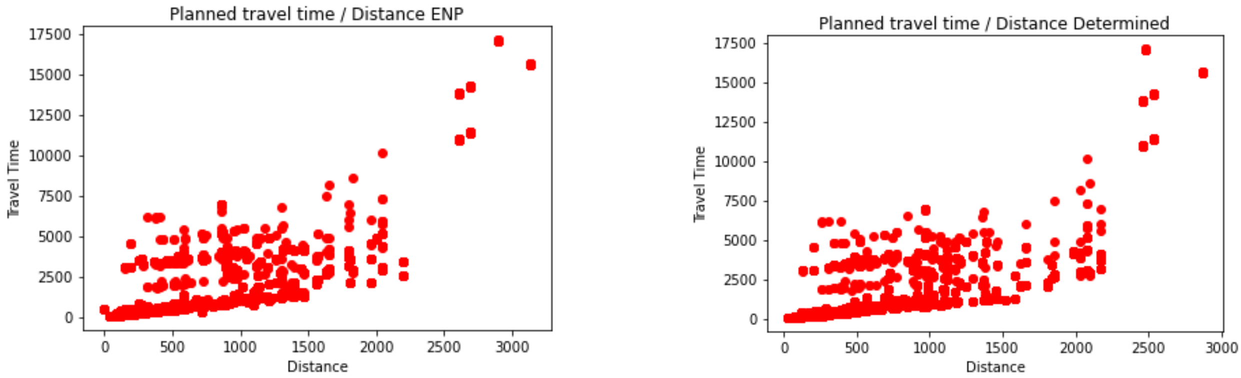
6.3.2 Approximating missing travel time information

Since data is only available on the operated line-hauls, information on travel time between hubs where there is currently no line-haul being operated is missing. In order to design a new network that potentially shows that other options might yield better operations these possible connections also need to be taken into account. Moreover, in the case of DB Schenker there is combined line-haul operated from Tilburg to Portugal, that first stops in Porto and ends in Lisbon. In the data however this combined line-haul is listed as both a line-haul from Tilburg to Porto and a line-haul from Tilburg to Lisbon. Therefore there is no general information on a line-haul between Porto and Lisbon.

Because of these cases an attempt is done to determine the mean of non-existing line-haul possibilities, and use this mean to generate a gamma distribution from which the stochastic travel time can be generated. This attempt is done by first determining the Euclidean distance similar to the method done in Section 5.1. Using an average speed to determine the travel time however proves not very efficient. Figure 6.2a shows a scattered plot of the planned travel time against the distance from the ENP (European Network Plan), where it can be seen that there is no real relation between only travel time and distance. Figure 6.2b however is loosely similar to Figure 6.2a, from which it can be said that the distance determined is similar to the distances from the ENP.

This determined distance might then be used to determine the travel time by means of Multiple Linear Regression. In order to use Multiple Linear Regression three different cases are used to best be able to find the travel time. For every case the existing line-haul data is used to fit a Multiple Linear Regression model and test its accuracy, so that these models can be used on predicting travel time of non-existing line-hauls.

First case is determining the travel time while knowing the Euclidean distance determined, the country of origin, country of destination and if the country of destination is either Turkey, Ireland, Finland, Romania and Sweden. The countries have a relative large travel time in relation to distance compared to other countries due to ferry crossings or border controls. The idea is that for these countries using a separate regression a more reliable travel time can be determined. For this case, the R^2 score of the regression model is **0.91**. A caveat to this model is that it can only be used for non-existing line-hauls where the country of origin and country of destination are equal to that of an existing line-haul, and that the country of destination is either



(a) ENP distance

(b) Euclidean distance

Figure 6.2: Travel time against distance.

Turkey, Ireland, Finland, Romania and Sweden. For non-existing line-hauls where this is not the case either the second case or third case applies.

For the second case a regression model is made to determine the travel time with features such as the Euclidean distance determined, country of origin and country of destination. For this case, the R^2 score of the regression model is **0.85**. Similar to the first case this model can only be used for non-existing line-hauls where the country of origin and country of destination are equal to that of an existing line-haul. Motivation for still including this model rather than only using the third case is that this model is better suited to specific line-hauls and by excluding these cases for the third case the model of the third case is better suited for general line-hauls.

At last for the third case a model is made to determine the travel time with the Euclidean distance. The idea of this model is that by excluding all specific line-haul operations where the travel time is relatively higher than general line-hauls, this model better predicts the general cases where their are no special conditions to the operations of a line-haul. For this case, the R^2 score of the regression model is **0.83**.

It can be said that for all 3 cases the model still has some inaccuracy in finding the travel time. In the end however by using using the gamma distribution to generate a number from using the output of the model as the mean of the gamma distribution, there is some stochasticity in the generated stochastic value of the travel time. This aspect loosely takes care of the slight uncertainty in determining the mean by means of a wider "confidence interval".

6.3.3 Approximating due time of shipments

Due times of shipments of DB Schenker are determined according to the existing CoDi schedule and existing line-haul schedule. To determine the due time of a premium shipment the quickest opportunity to deliver the shipment is determined. First the quickest opportunity of scheduled collection routes from the origin terminal is selected. From the planned arrival of this collection route at the origin hub, the first concurrent line-haul option is chosen. From the planned arrival of this line-haul option the next concurrent distribution option is chosen. This results in a schedule for a shipment that takes a certain amount of time, the quickest lead time it can take for a shipment to be delivered after it has been registered. This is also the lead time for a premium shipments. For standard system shipments the lead time is allowed to be one day more than the lead time of a premium shipment, or in the case where there is no delivery option on the next day the next delivery option.

When designing a new schedule, it is therefore difficult to determine the due times of shipments when the

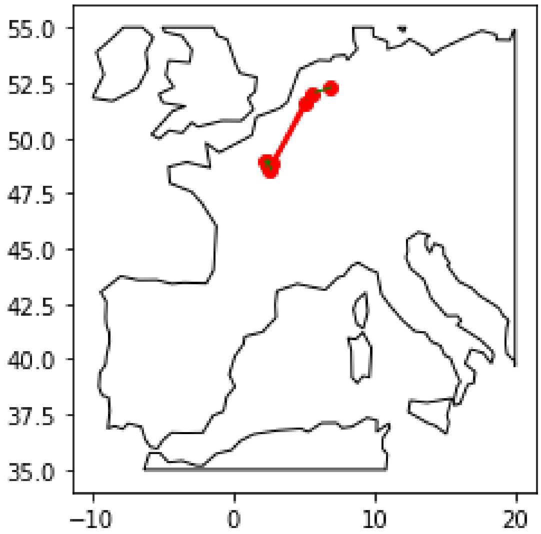
Distance hub-hub	Latest arrival time system	Latest arrival time spot departure
1 - 100 km	01:00 am	00:00 am
101 - 200 km	02:00 am	01:00 am
201 - 300 km	03:00 am	02:00 am
301 - 400 km	04:00 am	03:00 am
> 400	05:00 am	04:00 am

Table 6.1: Arrival guidelines according to distance (handled by DB Schenker).

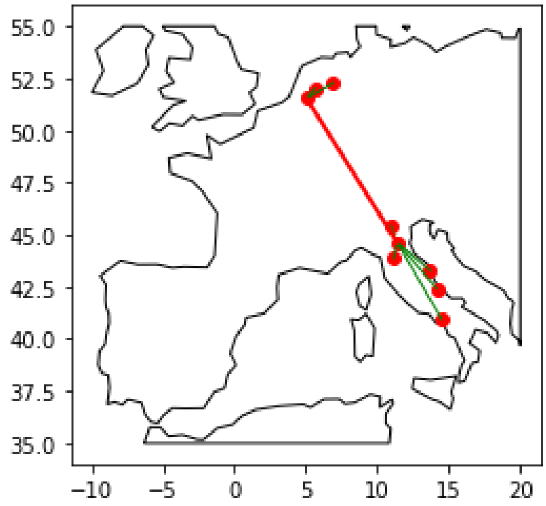
schedule on which the due times are based is not known. Several assumptions therefore have to be made in order to make the approximation. First, DB Schenker lists its expected delivery dates of shipments at the client according to the current schedule. Then the time on that expected delivery date is determined according a guideline according the hub to hub distance in Table 6.1. This guideline is originally based on hub-hub distances for line-hauls, however this distance could also be used as a distance between the origin and destination hub. For this calculation it is assumed that the due time of a shipment with a hub-hub distance larger than 400 km is always 05:00 am. In practice the trip from the origin to destination hub can be performed by multiple line-hauls with the last line-haul only having a distance of e.g. 150 km and therefore having a latest arrival time of 02:00. The decision on how many line-hauls are needed to perform the trip however is taken after the due time is known, so the assumption is made that the lead time of a trip with only one line-haul is the eventual lead time and therefore leads to the determined due time. Moreover, since the decision of using either a standard departure or a spot departure is taken after the due time is determined, the assumption is made that the due time of a shipment is not depending on the type of design decision (standard or spot departure). Since DB Schenker determines the lead-time according to an existing schedule, and moves the due time to the next option of there is a too tight lead time schedule, an assumption will be made that a day will be added to the lead time so that the due time will be moved to the next day. It is assumed that when a shipment is delivered in the morning and expected to be delivered at a client, these two events often can occur on the same day. Therefore the the expected delivery date at the client is also the due date of the shipment and according to the arrival guidelines the due time is determined on that date. Sometimes however this guideline is not accurate as can be seen in the case of BeNeLux-Serris in Subsection 6.4.1, where the due times of line-hauls are later than the supposed deadlines stated in Table 6.1. Therefore for some special cases of the case study the due time of the concurrent domestic line-haul is looked up and this due time is stated as the due time of the shipments going to that terminal.

6.4 Computational Experiments - Case Study

The experimental setup of the case study has remained the same compared to the experimental setup of Chapter 5, so for the experimental setup and parameters look at Section 5.2. The maximum number of iterations of the PH-Heruisitic method (gap 5%) however changed to 25. Since looking at the entire network of DB Schenker requires large amounts of computation power, the idea of the case study is to look at small parts of the network where disruptions occur frequently. Then export shipments with the Netherlands as origin (so from terminals Tilburg, Ede and Oldenzaal) are taken from a range of week 9 till week 13, with only looking at the first 2, 3 or 4 days of the week. This period is chosen since there is little influence of bank holidays to the existing schedule, and the choice for only two or three days rests on the fact that it is computationally intensive. Moreover the schedule used by DB Schenker is often a replication of one day for the rest of the week. For the timeframe timesteps of 1 hour are taken, meaning that every due time, ready time and travel time is rounded to the nearest hour. The goal is to take the original line-haul schedule of these smaller parts of the network and compare them to the solutions from the PH method.



(a) Serris



(b) Bologna

Figure 6.3: Illustration of network for case studies on Serris and Bologna.

This is done by looking at the objective values when taking different proportions of costs of shipments compared to the costs of operating a (spot) line-haul, potentially having an effect on the solution. After looking at the objective values the interpretation of the network is elaborated. Often only a look is taken at the export line-haul, not the domestic line-hauls (line-hauls operated to deliver shipments from the export line-haul destination to the eventual destination of the shipment). These domestic line-hauls are however part of determining the entire costs of the operations, since these are still an important part of delivering the shipments to the destination of the shipments. From this comparison a managerial insight is given. First a comparison is made on a part of the network between the Netherlands and Serris (France) and a part of the network between the Netherlands and the Italian city of Bologna, wheere these parts of the network are compared with the Progressive-Hedging method with respect to costs in Subsection 6.4.1. Then the results of this comparison are discussed in Subsection 6.4.2 For the comparison only options for improvements on the current operated line-hauls are taken into consideration, however the method could also be used to give insights in a whole new line-haul schedule with some direct line-hauls instead of the current line-hauls where a lot of consolidation takes place. Illustrations of both networks are shown in Figure 6.3a and Figure 6.3b for the case of Serris and Bologna respectively.

6.4.1 Comparison between network and solution method

For the case of BeNeLux to Serris we have 485 shipments over a period of 3 days for 4 weeks that need to be delivered to 6 other terminals next to Serris. Table 6.2 shows the objective values on the sub network between the Netherlands and Serris, where the objective of the network as it is right now is given, as well the solution of the Heuristic PH method and a Greenfield solution when the PH method is allowed to most optimal solution according to the current costs. These objectives are determined for three different cases where the shipment costs are altered. This is to find out if with a lower shipment costs a lower amount of transshipment takes place. For every cost setting the PH method provides a better solution seen from the objective.

For the case of BeNeLux to Bologna, we have 312 shipments over a period of 4 days for 4 weeks, having to be delivered to 6 other terminals next to Bologna that will be connected through the terminal of Bologna.

	DB Schenker network (as-is)	PH-Heuristic (gap 5%)	Greenfield solution
Cost factor	**Val.**	**Val.**	**Val.**
$c^k_{l_t,m_{t'}}$	225924	201894.4	183081.9
$\frac{1}{5}c^k_{l_t,m_{t'}}$	193278	167090.5	152674.8
$\frac{1}{10}c^k_{l_t,m_{t'}}$	184498	162441.3	148949.3

Table 6.2: Difference in objective values between DB Schenker as-is planning and new planning for the case of Serris.

	DB Schenker network (as-is)	PH-Heuristic (gap 5%)	Greenfield solution
Cost factor	**Val.**	**Val.**	**Val.**
$c^k_{l_t,m_{t'}}$	197992.7	189281.2	173656.4
$\frac{1}{5}c^k_{l_t,m_{t'}}$	76742.9	65036.6	59264.1
$\frac{1}{10}c^k_{l_t,m_{t'}}$	61586.7	49810.3	44950.4

Table 6.3: Difference in objective values between DB Schenker as-is planning and new planning for the case of Bologna.

Table 6.3 shows again the objectives for the case of a line-haul between Tilburg and Bologna, where the difference is shown in objective values for the as-is planning of DB Schenker, solution of the PH method and again a Greenfield solution. Here again it can be seen that the objective values of the PH solution are significantly lower compared to the objective values of DB Schenker planning.

Costs however do not tell the whole story, so Table 6.4 and Table 6.5 show the current schedule next to the solution obtained from the PH method for the cases of Serris and Bologna respectively.

6.4.2 Results and Discussion

From the comparison in Subsection 6.4.1 it can be said that in both the cases of Serris and Bologna the solution method is able to find a solution with lower costs for every cost setting. Moreover it can be said that when trying to obtain a greenfield solution, that is by neglecting the constraints that only the same arcs have to be used as the export line-haul in the as-is planning of DB Schenker, a even lower objective value is reached. To have a better understanding of the differences between the solutions, Table 6.4 and Table 6.5 show the current schedule being used and the suggested schedule.

From Table 6.4 it can be seen that in the case of Serris, compared to the original schedule for both **EDE-SXR** and **TLB-SXR** the PH method selects a departure on Monday at 16:00 and a departure on Tuesday at 23:00. The departure on Monday at 16:00 replaces the scheduled departures during the evening of both

Line-haul	**DB Schenker as-is planning**	**PH solution**
EDE-SXR	Monday 20:00 - Tuesday 04:00 Tuesday 20:00 - Wednesday 04:00	Monday 16:00 - Tuesday 00:00 Tuesday 10:00 - Tuesday 18:00 Tuesday 23:00 - Wednesday 07:00
TLB-SXR	Monday 22:30 - Tuesday 05:30 Tuesday 22:30 - Wednesday 05:30	Monday 16:00 - Monday 23:00 Tuesday 07:00 - 14:00 Tuesday 23:00 - Wednesday 06:00

Table 6.4: Schedules of DB Schenker as-is planning and the PH solution for the case of Serris for $\frac{1}{10}c^k_{l_t,m_{t'}}$.

Line-haul	DB Schenker as-is planning	PH solution
TLB-BLQ	Monday 00:30 - Monday 21:00 Tuesday 00:30 - Tuesday 21:00 Wednesday 00:30 - Wednesday 21:00 Thursday 00:30 - Thursday 21:00	Tuesday 04:00 - Wednesday 01:00 Wednesday 00:00 - Wednesday 21:00 Wednesday 17:00 - Thursday 16:00

Table 6.5: Schedules of DB Schenker as-is planning and the PH solution for the case of Bologna for $\frac{1}{10}c^k_{l_t,m_{t'}}$.

line-hauls. The departure at 23:00 on Tuesday merely resembles the departure of the original schedule for both lines, with the exception of **EDE-SXR** that departs earlier in the evening. Besides these differences, the method also chooses to select a third option on Tuesday in the morning for both line-hauls. Looking into the log it does seem that the loads are evenly spread among the three options for each line-haul, meaning there is some reason to support using them all. Moreover, there are less delayed shipments when using the solution from the PH method compared to using the original schedule, which results in a lower objective. A schedule of 3 line-haul options for two days however has some drawbacks that the method does not take into consideration, at first that the line-hauls during the day will interfere with the schedule of concurrent operations, and therefore might lead to shipments being stationary on terminals for too long, resulting in overcapacity at the terminals. Moreover, when driving with an extra truck results in a higher variability of the load during the execution of the line-haul, that might result in more empty miles.

One thing that directly can be noticed for the case of Bologna in Table 6.5 is that the planned duration of the current scheduled line-haul is 20 hours and that the planned duration of the scheduled line-hauls from the solution method are 21 hours. This duration is the average duration from data, therefore it may be said that the current duration is not entirely accurate. It can also be seen that as a result 2 line-hauls arrive roughly at the same time of the current line-hauls, but have to depart earlier. The second suggested line-haul however does not follow the similar time frames of the other the other two, but has a departure late in the evening and arrives somewhat later than the trend early on Thursday instead of late on Wednesday. This may be due to a lot of shipments being ready according to the old schedule and not ready early enough to depart earlier. Somehow there is still connection between the arrival at Bologna and the other domestic line-hauls. But it can be said that with the uncertainty in these ready times and maybe by not looking at enough scenarios there may be a difference in this departure time.

Another thing to mention is that the solutions from the PH method are based upon optimization. Therefore the method always chooses the best solution with for a minimization problem the lowest costs. This is also seen in the Greenfield solutions where suggestions were given different to the line-hauls currently being operated. Based on the current cost structure the method selects some direct connections, in order to deliver shipments under any circumstance. This is however unlikely to be a better solution in real life, since this will lead to more empty miles and a schedule with less cyclic behaviour. Due to a relatively high penalty costs for delayed shipments however, the method prefers to operate more line-hauls rather than having a shipment delayed. This extra amount of line-hauls being operated also has to do with the ready times of shipments, e.g. the time a shipment arrives at the origin terminal, since the model tries to get as many shipments onto a line-haul as possible. When the costs are therefore somewhat out of balance, it prioritises delivering as many shipments as possible, even though the ready times and due times of shipments may not be accurate. Therefore if the costs are not accurate, the solution might also not be accurate. It can thus be said that the method is able to provide a solution to the problem. With more realistic costs parameters the solution will eventually be more accurate.

7 Conclusion

This thesis addresses the Stochastic Service Network Design problem with uncertainty in demand and travel time while also considering multiple resource types. It proposes a MIP model in multi-scenario deterministic form. In this model, uncertainty in travel time is accounted for by making the time-space network scenario-dependent, and uncertainty in demand is handled by allowing additional resources to cope with excess demand. Additionally, recourse in the delay of shipments is considered. The model formulation was then adjusted to be used with a Progressive Hedging framework from existing literature on a Service Network Design with uncertainty in demand alone. This was adjusted so that consensus was sought only on the origin-destination pair and the departure time of a service due to the stochastic travel time and thus stochastic arrival. The framework has proven to be very efficient in solving the problem while problem size increased, without compromising solution quality compared to a commercial solver like Gurobi. Of the two Progressive Hedging strategies, the Heuristic strategy proved to be favorable. It also provided DB Schenker insights into a more robust schedule for their operations by suggesting new schedules for two cases where a part of the entire network was analyzed. The results included a reduction in objective values and a different schedule compared to the already existing schedule. However, due to the complex nature of the DB Schenker network, applying the newly designed schedule proved challenging. Focusing on a small part of the network often overlooks implications of the larger network. Hence, it might be beneficial to apply this method to a larger portion of the network, although it would require more computational power. Finally, due to the many assumptions made related to travel times and therefore also to costs a further look has to be given to the cost structure and how this influences the solution of the method. Since the solution method aims to find the optimal objective value, it prioritizes decisions that will lead to a lower objective value. Therefore the proposed schedule might be off by introducing more options since some shipments might not be delivered on time while waiting for other shipments that are not ready in time. By allowing other options for these shipments that are ready at a later time and do not interfere with the due time of the other shipments, the solution method thinks it has found a good solution that however in real life might prove very costly. It can thus be said that the method is able to provide a solution to the problem. With more realistic costs parameters the solution will eventually be more accurate. With more time invested, the method might be able to find a solution with a more cyclic behaviour, so that the solution of the method will be more useful in real life.

Implications and further research

From this thesis on, there are several opportunities for further research. First, this research proposes a new formulation in which the time-space network is used for different scenarios. This makes the problem more computationally intensive. It needs to be tested however if this makes the problem unnecessarily complex, and how this formulation compares to other formulations. The modeling of uncertainty in a SND is on itself already complex, the question however may be if this formulation makes it even more complex. To further extent the model formulation, the model might also include the use of long term contract charters such as done by Hewitt et al. (2019) or the model might include the influence of delays and the propagation of these delays, like done by Lanza et al. (2021). Right now the single-scenario problems are solved with a commercial solver, however a Tabu heuristic might be even more efficient in finding solutions, and may be able to find solutions for larger instances. Moreover more thorough experiments may be conducted, with less correlation in instances between the shipments in the different scenarios. Lastly, since for the time-frame of experiments time-steps of 1 hour are used, travel times also have to be rounded. Therefore the effect uncertainty of the travel time uncertainty is reduced. It might be interesting to test the effect of travel-time uncertainty by rounding to parts of an hour. Furthermore, this thesis only investigates whether it would be beneficial to model both uncertainty in demand and travel time, but it does investigate what type of uncertainty is more beneficial to consider.

Acknowledgements

I would like to give my mentor dr. Albert Schrotenboer a special thanks for his guidance on my thesis, great feedback on the process and the interesting discussions we had. I would also like to thank dr. Layla Martin, second supervisor from the TU/e for her help in providing feedback to my work and good suggestions for improvement.

I would also like to thank the department of land-transport of the BeNeLux cluster of DB Schenker, with a special thanks to my company supervisor Maité Delannoy for her patience during the weekly meetings and guiding me through the operations of the company. It really helped to have this opportunity of doing my thesis at DB Schenker to get more insight in how to operate within a company. At last I would like to thank Alexander van der Poll for helping me gain access to all the data necessary to do this thesis.

Bibliography

Bai, R., Wallace, S. W., Li, J., and Chong, A. Y.-L. (2014). Stochastic service network design with rerouting. *Transportation Research Part B: Methodological*, 60:50–65.

Bent, R. W. and Van Hentenryck, P. (2004). Scenario-based planning for partially dynamic vehicle routing with stochastic customers. *Operations Research*, 52(6):977–987.

Birge, J. R. and Louveaux, F. (1997). *Introduction to Stochastic Programming*. Springer Series in Operations Research and Financial Engineering. Springer.

Boland, N. L., Hewitt, M., Marshall, L., and Savelsbergh, M. W. P. (2017). The continuous-time service network design problem. *Operations Research*, 65:1303–1321.

Chen, Z. and Fan, W. D. (2020). Analyzing travel time distribution based on different travel time reliability patterns using probe vehicle data. *International Journal of Transportation Science and Technology*, 9(1):64–75.

Crainic, T. G. (2000). Service network design in freight transportation. *European Journal of Operational Research*, 122(2):272–288.

Crainic, T. G., Fu, X., Gendreau, M., Rei, W., and Wallace, S. W. (2011). Progressive hedging-based metaheuristics for stochastic network design. *Networks*, 58(2):114–124.

Crainic, T. G., Hewitt, M., and Rei, W. (2014). Scenario grouping in a progressive hedging-based metaheuristic for stochastic network design. *Computers & Operations Research*, 43:90–99.

Demir, E., Burgholzer, W., Hrušovský, M., Arıkan, E., Jammernegg, W., and Woensel, T. V. (2016). A green intermodal service network design problem with travel time uncertainty. *Transportation Research Part B: Methodological*, 93:789–807. Maritime Logistics.

Hewitt, M., Crainic, T. G., Nowak, M., and Rei, W. (2019). Scheduled service network design with resource acquisition and management under uncertainty. *Transportation Research Part B: Methodological*, 128:324–343.

Hoff, A., Lium, A.-G., Løkketangen, A., and Crainic, T. G. (2010). A metaheuristic for stochastic service network design. *Journal of Heuristics*, 16:653–679.

Lanza, G., Crainic, T. G., Rei, W., and Ricciardi, N. (2021). Scheduled service network design with quality targets and stochastic travel times. *European Journal of Operational Research*, 288(1):30–46.

Liu, C., Lin, S., Shen, Z.-J. M., and Zhang, J. (2023). Stochastic service network design: The value of fixed routes. *Transportation Research Part E: Logistics and Transportation Review*, 174:103118.

Lium, A.-G., Crainic, T. G., and Wallace, S. W. (2009). A study of demand stochasticity in service network design. *Transportation Science*, 43(2):144–157.

Müller, J. P., Elbert, R., and Emde, S. (2021a). Integrating vehicle routing into intermodal service network design with stochastic transit times. *EURO Journal on Transportation and Logistics*, 10:100046.

Müller, J. P., Elbert, R., and Emde, S. (2021b). Intermodal service network design with stochastic demand and short-term schedule modifications. *Computers & Industrial Engineering*, 159:107514.

O'Kelly, M. E. (1986). The location of interacting hub facilities. *Transportation Science*, 20(2):92–106.

Powell, W. B. and Topaloglu, H. (2003). Stochastic programming in transportation and logistics. In *Stochastic Programming*, volume 10 of *Handbooks in Operations Research and Management Science*, pages 555–635. Elsevier.

Rockafellar, R. T. and Wets, R. J.-B. (1991). Scenarios and policy aggregation in optimization under uncertainty. *Mathematics of Operations Research*, 16(1):119–147.

Xiang, X., Fang, T., Liu, C., and Pei, Z. (2022). Robust service network design problem under uncertain demand. *Computers & Industrial Engineering*, 172:108615.