Centre interuniversitaire de recherche sur les réseaux d'entreprise, la logistique et le transport

Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation

Stochastic Scheduled Service Network Design: The Value of Deterministic Solutions

Xin Wang
Teodor Gabriel Crainic
Stein W. Wallace

September 2014

CIRRELT-2014-46

Bureaux de Montréal :
Université de Montréal
Pavillon André-Aisenstadt
C.P. 6128, succursale Centre-ville
Montréal (Québec)
Canada H3C 3J7
Téléphone : 514 343-7575
Télécopie : 514 343-7121

Bureaux de Québec :
Université Laval
Pavillon Palasis-Prince
2325, de la Terrasse, bureau 2642
Québec (Québec)
Canada G1V 0A6
Téléphone : 418 656-2073
Télécopie : 418 656-2624

www.cirrelt.ca

Stochastic Scheduled Service Network Design: The Value of Deterministic Solutions

Xin Wang[1], Teodor Gabriel Crainic[2,*], Stein Wallace[3]

[1] Department of Management Science,Lancaster University Management School, Bailrigg, Lancaster, LA1 4YX, United Kingdom

[2] Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation (CIRRELT) and Department of Management and Technology, Université du Québec à Montréal, P.O. Box 8888, Station Centre-Ville, Montréal, Canada H3C 3P8

[3] Department of Business and Management Science, Norwegian School of Economics, NHH, Hellevein 30, NO-5045, Bergen, Norway

Abstract. We study the value and quality of deterministic solutions to scheduled stochastic capacitated multi-commodity service network design problems. We study both the fixed and variable (integer and continuous) capacity cases, and investigate models with and without asset-balance constraints. For the deterministic cases, we replace the random variables in the stochastic model with the 50^{th} and the 75^{th} percentiles from the demand distributions, confirming the better quality of solutions from the latter case in all situations. We also investigate what makes the optimal stochastic solution better in the stochastic environment than other feasible solutions, particularly those obtained by addressing the deterministic versions of the problem. We do this by quantitatively analyzing the structures of different solutions. A measurement scheme is proposed to evaluate the level of potentially beneficial structural properties (multi-path usage and path-sharing) in different solutions. We show that these structural properties are important and correlated with the performance of a solution in the stochastic environment.

Keywords: Stochastic service network design, scheduled, consolidation, value of deterministic solution.

Acknowledgements. Partial funding for this project has been provided by the Natural Sciences and Engineering Council of Canada (NSERC), through its Industrial Research Chair and Discovery Grant programs.

---

* Corresponding author: Teodor-Gabriel.Crainic@cirrelt.ca

Dépôt légal – Bibliothèque et Archives nationales du Québec
Bibliothèque et Archives Canada, 2014

1 Introduction

In the freight transportation industry, service network design methods help provide solutions to a set of tactical problems: where, when and how to offer services for the delivery of the demands. The goal is to decide the selection, routing and scheduling of services, while balancing the operating costs and service quality. Discrete decision variables are normally involved in such processes and the resulting programming models are thus usually very complex.

This is even more so when some parameters of the model are uncertain. In freight transportation problems, the most commonly modeled uncertain phenomenon is demand. It is normally represented by a set of scenarios approximating a "known" demand distribution (multi-dimensional in the case of multi-commodity problems). The resulting stochastic programming model is, compared to its deterministic counterpart, much more difficult to solve due to its much larger size.

In most, if not all, situations, decisions are made under uncertainty. It is generally understood that in these situations, stochastic programming models are more appropriate than deterministic ones. In service network design, the optimal solution found with all random demands taking some fixed values, e.g. their means, (we call it a "deterministic solution" to a stochastic problem), represents the optimum for one single scenario of the uncertain demands. When the number of scenarios increases, the uncertainty of the demands is better represented (assuming the scenarios are well constructed), but the corresponding model will eventually become numerically unsolvable. One way out of this problem is to solve the stochastic program heuristically. Examples can be found in. e.g., Hoff et al. (2010) and Crainic et al. (2011).

The premise of this paper is that there are situations where an optimal (or near-optimal) deterministic solution can be found for a service network design problem, while the optimal stochastic solution cannot for numerical reasons; the problem is simply too large.

Of course, under normal circumstances, the expected behavior of a solution derived from a stochastic model should be much better than its deterministic counterpart when evaluated in the stochastic environment. The reason is that while it is optimal for one specific scenario, it might be very bad in those scenarios where it is not optimal. See for example Wallace (2000) and Higle and Wallace (2003) for discussions. This badness can be measured by "the Value of the Stochastic Solution" (Birge, 1982), or VSS, representing the expected gains obtained from using the stochastic rather than the deterministic solution in the stochastic environment. However, there are situations where the VSS is high, meaning that the deterministic solution behaves badly in the stochastic setting, yet the deterministic solution shares some properties with the corresponding stochastic solution. For example, Thapalia et al. (2012b, 2011, 2012a) show that for the single-commodity network design problem, certain structural patterns from the deterministic solutions re-emerge in the stochastic solutions. Similar observations are made in Maggioni and Wallace (2012) for a series of other problems.

So we ask: what makes a stochastic solution behave better than its deterministic counterpart for service network design problems? And when the deterministic solution is relatively easy to obtain, can we make any use of it even if it is bad in its own right? In particular, can we make use of information extracted from a deterministic solution, and construct a good solution for the stochastic case? If yes, which part of the deterministic solution should we extract? In this paper, we present a classical and a variant of the service network design model, and for both models, but in different ways, we are able to make effective use of parts of the deterministic solution.

A robust design for a transportation system resists the possible changes of the uncertain demands by providing operational flexibility. One way to achieve more flexibility is to provide services with higher capacity or at higher frequency, at the expense of increasing initial costs. Consolidation, however, can serve the purpose quite well in a highly dynamic environment where future demands are unknown, without requiring too much extra services. Traditionally, in consolidation-based freight transportation, consolidation is seen as a way to accommodate the fact that most vehicles would not be full with direct deliveries. In the stochastic setting, such consolidation can also be induced by the need to hedge against demand uncertainty. Studies by Lium et al. (2007, 2009) indicate that solutions produced with explicit consideration of stochastic demand are qualitatively different from those stemming from deterministic models. They show that, for their problems, multi-path usage and path-sharing offer better solutions when there are uncertainties in demand. In this paper we investigate whether this applies in our models as well. If yes, can we represent these structural differences *quantitatively*? If we can develop a measurement scheme to quantify the level of such structural properties for different solutions, then we can use these measurements to see how the level of the potentially important structural properties of a solution are related to its performance in the stochastic environment.

In addition, we investigate the possibility of extracting *partial* information from a deterministic solution and using this information to simplify the stochastic model. This will in practice mean fixing the values of certain variables before solving a simplified stochastic model. The solutions coming from this simplified stochastic model are then evaluated in the stochastic environment. This way we can measure the value (or quality) of the information we extracted from the deterministic solution.

The contribution of this paper is to provide a complete analysis of the quality and upgradeability of deterministic solutions to stochastic scheduled service network design problems. We discuss fixed as well as variable (integer and continuous) capacity models, and we test both with and without asset-balance constraints. We compare, for each such case, deterministic solutions stemming from the 50^{th} and the 75^{th} percentile of the demand distributions, confirming the better quality of the latter.

This paper is organized as follows. In Section 2, some important issues on freight transportation and service network design are reviewed. Section 3 introduces the formulation of the scheduled stochastic service network design with fixed capacities. The value of the deterministic solutions in such a model is discussed in Section 4. Section 5 considers the variable capacity model with both integer and continuous capacities, while asset balance requirements are tested in Section 6. We conclude in Section 7.

2 Freight Transportation and Service Network Design

Transportation is an important domain of human activity. It supports and enables many other social and economic activities and exchanges. Freight transportation, in particular, is one of today's most important activities. Demand for freight transportation reflects the need to move goods between producers and consumers and requires a rather complex system which derives from the fact that the distances separating them are often significantly long. Crainic (2003) give a general presentation of freight transportation players, issues, and problem classes. In the increasingly competitive environment, carriers seek to offer reliable, high quality services to their customers at a lowest possible cost, and in the mean time make a profit.

An often encountered transportation system is consolidation, where one vehicle or convoy usually serves more than one customer. In a system where demand for transportation is represented by origin-destination (OD) pairs, freight of different OD pairs with different initial origins and final destinations are combined into common vehicles, e.g. railways, LTL motor carriers, container shipping lines and postal services.

The underlying structure of a consolidation transportation system normally consists of a large network of terminals and the transportation operations thus concerned are usually rather complex. This is in contrast to customized transportation which provide dedicated service for each OD pair. Consolidation-based transportation carriers usually engage into so-called hub-and-spoke networks to take advantage of economies of scale. In such systems, low-volume demands are first delivered to an intermediate terminal or a hub to be grouped and consolidated. High-frequency, high-capacity services are provided between the hubs, and can thus allow a much higher frequency of service between all the OD pairs. However, routing through several intermediate terminals and hubs would inevitably result in longer transport distance and more time spent at terminals and could sometimes cause serious delays. There is a great deal of literature on the subject.

Surveys are presented by Christiansen et al. (2004, 2007) for maritime transportation, Cordeau et al. (1998) for rail transportation, Crainic and Laporte (1997) and Crainic (2003) for land-based long-haul transportation, Crainic and Kim (2007) for intermodal transportation, and Crainic (2000) for service network design in freight transportation.

In order to satisfy the demand of customers more timely and reliably, consolidation carriers operates a series of *services*, each characterized by its own route, vehicle type, frequency, capacity etc. Internally, services are often collected in an operational plan (also referred to as load or transportation plan), generally accompanied by a schedule that indicates departure and arrival times at the terminals of the route (Crainic and Kim, 2007). Service network design formulations are used to build such a transportation plan (schedule) for the next operating period.

Service network problems address a set of major issues and decisions relevant for consolidation-based carriers: the selection and scheduling of the services to operate, the routing of freight for each OD pair and the consolidation operations at terminals. The goal is to achieve profitable operations while providing timely and reliable services according to customer expectations. The corresponding models usually take the form of network design formulations. With the complicated interactions among system components and decisions as well as the tradeoffs between operating costs and service quality, service network design models are very difficult to solve, and thus heuristics are usually the solution method of choice.

Reviews on the formulation of service network design models are presented by Crainic (2000, 2003), Delorme et al. (1988) and Cordeau et al. (1998). Efforts have been made towards both static and scheduled service network design formulations. The former assume a static demand throughout the whole planning period. The time dimension of the service network design is then implicitly considered through the definition of services and interservice operations at terminals. Such models have been proposed for multimodal transportation (Crainic and Rousseau, 1986; Crainic and Roy, 1988); LTL trucking (Roy and Delorme, 1989; Powell and Sheffi, 1983, 1986, 1989; Powell, 1986; Lamar et al., 1990), express courier services (Grünert et al., 1999; Grünert and Sebastian, 2000; Buedenbender et al., 2000; Barnhart and Schneur, 1996; Kim et al., 1999; Armacost et al., 2002), rail (Crainic et al., 1984; Keaton, 1989, 1991, 1992; Newton, 1996; Newton et al., 1998), and shipping (Christiansen et al., 2004) etc.

Scheduled service network design formulations include an explicit representation of movements of freight in time and usually target the planning of schedules to support decisions related to when services depart, either from origins or intermediate terminals. A *space-time network* with a scheduling time line is usually used to represent the operations of such scheduled service network systems. The representation of the physical network is replicated at each time point. Temporal arcs then connect the same or different terminals within two time-point representations to represent respectively holding activities at the same terminal or actual movement of freight between terminals. The resulting model formulations are similar to those of the static versions but on significantly larger networks due to the time dimension. The additional constraints related to scheduling also contribute to make this class of problems more difficult to solve than static versions. Such formulations have been proposed for, e.g., LTL trucking (Farvolden and Powell, 1991, 1994; Farvolden et al., 1992), express courier services (Smilowitz et al., 2003), rail (Haghani, 1989; Gorman, 1998a,b; Andersen et al., 2009a,b; Pedersen et al., 2009; Zhu et al., 2013) and navigation (Sharypova et al., 2012). Meta-heuristics were proposed in most cases.

Another noteworthy issue in any freight transportation system is the need to move empty vehicles. This follows from the imbalances between the freight supply and demand in different regions and points of the systems, resulting in imbalances between vehicle supplies and demands at the terminals of the system. To redress these imbalances, empty vehicles must be delivered to terminals where they will be needed to satisfy known or forecasted demand in the following time periods. These repositioning operations call for the most economic solutions and are normally left to be dealt with at the operational level of planning on a decided service network schedule (e.g., Dejax and Crainic, 1987; Cordeau et al., 1998; Crainic et al., 1989). Efforts have lately been dedicated to considering vehicle repositioning and other asset management requirements at the tactical design stage (e.g., Pedersen et al., 2009; Andersen et al., 2009a,b, 2011; Lium et al., 2007, 2009; Bai et al., 2014).

Service network design problems have mainly been studied under the assumption that all necessary information, particularly the demand as well as the cost and profit structure, is available before the design decisions are made. It is a general understanding that in most cases, at the time when the transportation plan is decided, the demand it will later face is uncertain. This is traditionally not explicitly taken into account during the design phase but postponed to be dealt with at the operational phase. Hence, most papers use deterministic models. Demand is usually set to some estimations of the future computed through various forecasting methods or based on historical data (e.g., the "regular" demand of a "normal" week obtained by adjusting last-year demand with this year input from the sales department).

Previous studies have shown that by introducing stochastic demand into a service network design model, the solutions produced can be qualitatively different from those stemming from deterministic models, see for example Wallace (2010). In Lium et al. (2007, 2009), the authors present a service network design model in which the decision variables also capture the *frequency* of each service, represented by the number of vehicles used for each service, and the repositioning of empty vehicles is then taken into account at the design stage. They show that compared to its deterministic counterpart, in a stochastic solution commodities use more paths; paths are shared by multiple commodities; and more hub-and-spoke structures are observed.

We first consider a classical version of a stochastic, fixed cost, capacitated, multi-commodity service network design model with fixed capacities (i.e., if the capacities are represented by the number of vehicles, each service may use a different number of vehicles, but this number is fixed if a certain service is selected). We also consider a variant, a model with variable capacities. For the two types of models, we investigate the situations both with and without asset balancing.

What we are interested in is firstly: for our models, are there any qualitative differences between the stochastic and the deterministic solutions? If yes, what are the characteristics of optimal stochastic solutions and do they project similar structural features as in the Lium et al. (2009) paper? These structural features are not obviously seen in classical service network design problems, therefore a method of measurement is needed to quantify such structural differences.

The second issue we are interested in is: in our models, how much value do the deterministic solutions have in the stochastic environment? For the classical model with *fixed capacities*, the "design" of the solution to the deterministic version carries only the information of the selection of the services, i.e., the {0,1} variables. Our experiments show that such information is important and most of the selected services are kept in the optimal solution to the stochastic problem as well. For the variant with *variable capacities*, the "design" also includes information about the respective capacities provided on all the services. We then define the "skeleton design" as the *service selection information*, i.e., where and when to set up the services. The results show that the deterministic solution is generally quite bad when used in the stochastic environment, so can we use part of it instead? More specifically, if we only keep the "skeleton design" of the deterministic solution and decide the capacities later, will the solution produced then be good?

We thus investigate the qualities of the deterministic solutions beyond just their absolute qualities, based on both variants of the service network design model. For each variant, the performance of the following three types of solutions are evaluated in the stochastic environment: 1. optimal stochastic solution (as a benchmark); 2. deterministic solution; 3. reconstructed solution based on the deterministic design.

Furthermore, in order to understand the structural differences between these different types of solutions, we propose a measurement scheme to quantify the level of the following structural properties: multi-path usage and path-sharing. We can thus examine the interaction between the level of these structural properties for each type of solution and its performance in the stochastic environment. If a certain quantitative pattern emerges, e.g. that better solutions have higher level of path-sharing, then reversely it may translates into a viable approach for finding better solution to the stochastic problems.

3 Stochastic Service Network Design

For the study presented in this paper, we will first address a classical stochastic, scheduled, capacitated, multi-commodity service network design model in which periodic, cyclic schedules are built for a number of commodities (OD pairs). All services are scheduled over a planning period, and the schedule is repeated for the duration of the planning horizon. See Andersen et al. (2009a,b); Pedersen et al. (2009); Zhu et al. (2013) for examples of circularly scheduled service networks.

3.1 Problem Setting and Notation

The service network design problem is set up on a space-time network consisting of nodes and arcs. The nodes in the space-time network diagram stand for terminals at different points in time and the arcs represent services for moving commodities between these terminals across time, as well as the possibility to hold vehicles and freight at a terminal between two consecutive periods.

We use {0,1} decision variables to capture the service selection choices, indicating whether or not the service leaves at the specified time point. Therefore, only one service, with specified characteristics, is allowed at a given time point from one terminal to another. When several departures are possible in the same time interval, general (non-negative) integer variables must be used. Note that by making the time intervals appropriately small, one can always use {0,1} variables to address multiple departures within a certain period of time. Therefore, the basic model we propose in this section includes only {0,1} variables.

Let $\mathcal{N}$ represent the set of terminals. In a space-time network, these terminals are replicated at each point in time. We denote by $\mathcal{T}$ the set of points in time. The scheduling period consisting of these time points is repeated in a cyclic fashion. We denote by $\mathcal{A}$ the set of arcs between the nodes. An arc $a = (i, j; t)$ represents the service departing from terminal i at time point t and arriving at terminal j. A service can be set up at any time point between any pair of terminals $(i, j), \forall i, j \in \mathcal{N}, i \neq j$, in either direction. It is assumed that a service can take one or more time intervals, depending on the physical distance between the two terminals. We use l_{ij} to represent the service length between terminals i and j, i.e., the number of time intervals required for transporting goods between the two terminals. Furthermore, it is assumed that the handling of freight at terminals happens instantaneously within the time intervals, which implies no time delay caused by terminal operations such as unloading, sorting, consolidation and loading activities.

Figure 1 shows the time-space network diagram with three terminals and a repetitive

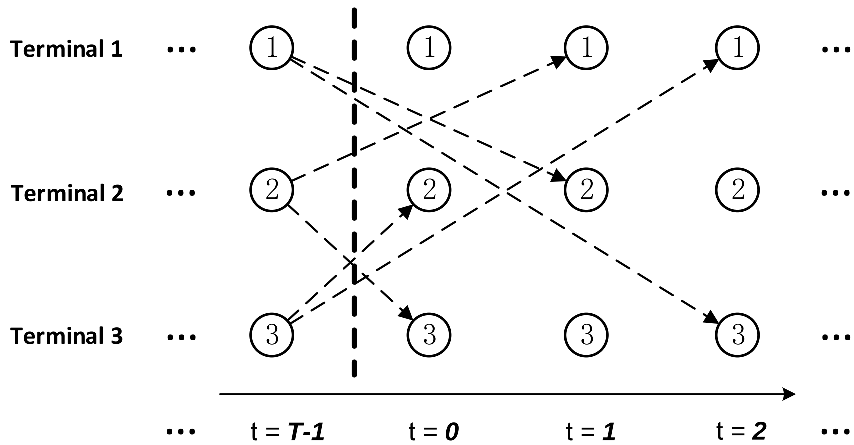

Figure 1: Time-space network diagram for cyclic service schedule.

scheduling period consisting of T time points. The dashed arrows represent the services that can be set up at $t = T - 1$, which is the last time point of the repetitive scheduling period. For example, a service departing from Terminal 1 at $t = T - 1$ takes three time intervals and arrives at Terminal 3 at $t = 2$ of the subsequent T-period scheduling period. Holding arcs, joining two representations of the same terminal in two consecutive periods, are not displayed for increased clarity. Note that the cyclic feature of the space-time network is illustrated by joining two sequential scheduling periods and the services arrive in the latter one by leaping over the bold division line in Figure 1.

The set of commodities (OD pairs) $\mathcal{K}$ represents the origin-to-destination demands for transporting a certain quantity of freight between the respective origin and destination terminals within a certain number of time intervals. For each $k \in \mathcal{K}$, the shipping requirements of commodity k are defined by: o_k, d_k, its origin and destination terminals; σ_k, τ_k, the time point it becomes available and the time point by which it must be delivered; and its demand.

We take explicitly into account the demand stochasticity at the design phase, describing the demand for each commodity by a continuous distribution. To be able to analyze the stochastic problem with exact methods, the multi-dimensional demand distribution is then represented by a set of scenarios $\mathcal{S}$. A probability p^s is assigned to each scenario $s \in \mathcal{S}$, with $\sum p^s = 1$. We use δ_k^s to denote the demand for commodity k in scenario s; thus a scenario is |K|-dimensional and contains one demand realization for each commodity.

There is a fixed set up cost $f_{ij;t}$ associated with opening an arc $(i, j; t) \in \mathcal{A}$ and providing the related fixed capacity $h_{ij;t}$. Also we need to pay for commodity flows, that is, the transportation and storage of the commodities. Thus costs $e_{ij;t}$ associated to each arc $(i, j; t)$ represent the unit flow costs incurred to move commodities on services or have them wait at terminals on holding arcs. Additionally, to account for demand not satisfied by the services, we denote by b_k the unit ad hoc handling cost of moving commodity k whenever part (or all) of its demand cannot be satisfied by regular services.

The goal is to solve the stochastic optimization problem in order to find a good, if not optimal, solution such that a periodic schedule is designed to minimize the expected total system costs. This corresponds to a two-stage structure in the decision model. The first stage decisions, i.e., the selection of services or "the design", are made before the realization of the random demands. A fixed cost must be paid whenever a service is selected (set up), representing its make up or maintenance costs. Once these decisions are made, the design is used to satisfy the observed realization of random demands. So the second stage is characterized by routing commodity flows using the selected services and the "extra" capacity described by the ad hoc arcs. The overall objective is thus to minimize the cost of the first stage design plus and expected operational and ad-hoc handling costs when applying such a design to the demand realizations.

3.2 The Fixed Capacity Model

Let $V_{ij;t}$ represent the {0,1} service selection decision variables, and $Y^s_{a;k} = Y^s_{ij;t;k}$ be the flow variables, representing the continuous flow of commodity k on arc $(i, j; t)$ in scenario s. Furthermore, let Z^s_k represent the continuous volume of commodity k that uses ad hoc handling in scenario s.

Due to the cyclic nature of the network, the m^{th} time point prior to time t can be denoted as:

$$t \ominus m = (t - m + T) \bmod T \tag{1}$$

The two-stage stochastic formulation of the scheduled service network design problem can then be written as:

$$\min \sum_{i\in\mathcal{N}} \sum_{j\in\mathcal{N}} \sum_{t\in\mathcal{T}} f_{ij;t} V_{ij;t} + \sum_{s\in\mathcal{S}} p^s \Big(\sum_{i\in\mathcal{N}} \sum_{j\in\mathcal{N}} \sum_{t\in\mathcal{T}} \sum_{k\in\mathcal{K}} e_{ij;t;k} Y^s_{ij;t;k} + \sum_{k\in\mathcal{K}} b_k Z^s_k \Big) \tag{2}$$

$$\sum_{i\in\mathcal{N}} Y^s_{ij;t\ominus l_{ij};k} - \sum_{i\in\mathcal{N}} Y^s_{ji;t;k} = \begin{cases} \delta^s_k - Z^s_k, & \text{if } j = d_k \text{ and } t = \tau_k \\ -\delta^s_k + Z^s_k, & \text{if } j = o_k \text{ and } t = \sigma_k \\ 0, & \text{other} \end{cases}$$
$$, \forall j \in \mathcal{N}, \forall t \in \mathcal{T}, \forall k \in \mathcal{K}, \forall s \in \mathcal{S} \tag{3}$$

$$\sum_{k\in\mathcal{K}} Y^s_{ij;t;k} \leq h_{ij;t} V_{ij;t} \qquad \forall i, j \in \mathcal{N}, i \neq j, \forall t \in \mathcal{T}, \forall s \in \mathcal{S} \tag{4}$$

$$0 \leq Z^s_k \leq \delta^s_k \qquad \forall k \in \mathcal{K}, \forall s \in \mathcal{S} \tag{5}$$

$$0 \leq Y^s_{ij;t;k} \leq \delta^s_k \qquad \forall i, j \in \mathcal{N}, \forall t \in \mathcal{T}, \forall k \in \mathcal{K}, \forall s \in \mathcal{S} \tag{6}$$

The objective function (2) minimizes the costs for opening services plus the expected costs for moving and holding commodities, as well as using ad hoc capacities. Constraints (3) represent the conservation of flow for commodities. Constraints (4) make sure the total flow on each arc does not exceed its capacity. Constraints (5) limit the ad hoc capacity to the observed actual scenario demand and (6) limit the flow of every commodity to its corresponding demand on all arcs.

4 Methods and Results for the Fixed Capacity Model

In the Introduction, we posed these two questions which correspond to the general objective of this paper: what makes a stochastic solution behave better than its deterministic counterpart? And when the deterministic solution is relatively easy to obtain, can we make any use of it even if it is bad in its own right?

Therefore, we start by evaluating the absolute qualities of the deterministic solution in the stochastic environment. We then construct another solution, using parts of the deterministic solution, to see if the performance can be improved. The optimal stochastic solution is used as benchmark.

We then proceed to the question: what structural properties make one solution better than another in the stochastic environment? We propose a quantitative scheme (based on counting multi-path usage and path-sharing) to measure the level of the potentially relevant structural properties in different solutions. We then relate these counts to the quality of the solutions.

4.1 Performance Comparison

For a given scenario tree, we use the following comparison tests inspired by Thapalia et al. (2012b) to compare the performances of different designs in the stochastic case:

1. Stochastic solution: (optimal solution)

   We solve the stochastic problem. This is used as a benchmark for other solutions.

2. Deterministic solution used in the stochastic model: (**Deter**)

   This is the standard "Value of the Stochastic Solution" evaluation. We first solve the deterministic version of the problem and observe which arcs are open. We keep these arcs open and close all other arcs in the network, i.e., we fix the first stage decision variables $V_{ij;t}$, and then run an LP to set the flow variables of the stochastic model.

3. Deterministic design with extra services in the stochastic case: (**Deter Plus**)

   Again we start by solving the deterministic problem. Now we keep those arcs obtained from the deterministic solution open, but do NOT close other arcs in the network. We then run the stochastic problem again to allow services to be set up in addition to those opened in the deterministic solution. This is again a MIP due to the service selection process on the left-to-be-decided arcs.

The tests of **Deter** and **Deter Plus** are performed to check the **absolute performance** and **upgradeability** (elaborated later in the paper) of the deterministic solution in the stochastic setting. So the **Deter Plus** test is not included for its numerical efficiency, but in order to learn about the problem we are investigating: can the deterministic solution be upgraded to a good solution in the stochastic environment, or are we already lost when implementing the deterministic solution? Both conclusions are possible as demonstrated by (Maggioni and Wallace, 2012).

4.2 Parameter Setting and Instance Generation

To evaluate the qualities of the deterministic solution and its upgradeability, we should preferably use the "true" stochastic solution as benchmark. For this reason we only study cases where the stochastic programs can be solved numerically.

Instances are built using randomly generated parameters. To start with, we generate values for the coordinates of all the terminals, evenly positioned inside a square-shaped area. Direct services are allowed between *any* two terminals, which indicates a potentially complete service network. The service lengths are decided according to the physical distances between the two terminals that are associated with the considered service, such that for any $i, j \in \mathcal{N}$ and $i \neq j$, service length l_{ij} has three possible values: 1, 2 or 3. The values for the service unit flow costs $e_{ij;t}$ and of the unit ad-hoc handling cost b_k associated with commodity k are set proportional to the distances between the terminals i and j, the latter being ten times higher than the former. Service fixed costs and capacity are fixed at 25 and 6, respectively.

For every commodity, its origin and destination terminals are both selected randomly. The time span (from the time point it becomes available to the time point it has to be delivered) ranges from 2 to 5. In the stochastic versions of our test instances, the stochastic demands of all commodities are subject to symmetric triangular distributions with a standard deviation equal to 40% of the mean.

We discretize the demand distributions by generating scenarios with equal probabilities to represent the stochasticity. The scenario generation process is performed using the moment-matching method introduced by Høyland et al. (2003). The demand correlation matrix needed to generate the scenarios for every instance is created as follows: (a) the commodities are equally (or almost equally, if there is an odd number of commodities) divided into two groups; (b) if two commodities are in the same group, their demands are assumed to be positively correlated with a correlation value randomly chosen within the range [0.00,0.50); otherwise, if the two commodities are in different groups, their demands are negatively correlated with a correlation value randomly chosen within (-0.50,0.00]; (c) the resulting matrix has to be positive semi-definite to ensure its validity to be used as a correlation matrix; if not, step (b) is repeated. This way of constructing correlation matrices normally leads directly to positive semi-definiteness.

The more scenarios, the better the representation of the demand distribution. But as we increase the number of scenarios, the difficulty to obtain an optimal solution becomes severer as well. Thus there is a trade-off between the stability of the stochastic solution and the problem growing too large. In our experiments, we use 30 scenarios to represent the stochasticity. The in-sample stability tests (Kaut and Wallace, 2007) give a difference of less than 5%, which is acceptable, between the highest and lowest optimal objective function values on a large number of different scenario trees.

4.3 Experiments and Results

We use ten test instances randomly generated following the procedure mentioned earlier, each based on the space-time network of 7 time points, 6 terminals and 16 commodities. We then generate 30 scenarios for each instance to represent the demand stochasticity. The three comparison tests are performed for all instances and the results are shown in Figure 2. The bars show the losses produced in the **Deter** and **Deter Plus** tests, relative to the optimal solutions of the stochastic program.

The ***Min.Loss*** and ***Max.Loss*** show the best and worst cases for the two tests respectively out of the ten instances. The ***Avg.Loss*** are the mean losses the two tests produce across all the instances.

4.3.1 Value of Deterministic Solution

From the comparisons in Figure 2 we can see that although losses can go as high as nearly 55%, the deterministic solutions are generally rather good in the stochastic setting, producing an average loss of around 20%, which is rather small compared with some other stochastic network design problems (see Thapalia et al. 2012a; Maggioni and Wallace 2012). However, with extra services the deterministic solutions can still be greatly improved. The **Deter Plus** test shows that adding extra services to the deterministic design is beneficial and effective in most circumstances (loss is under 10% even for the

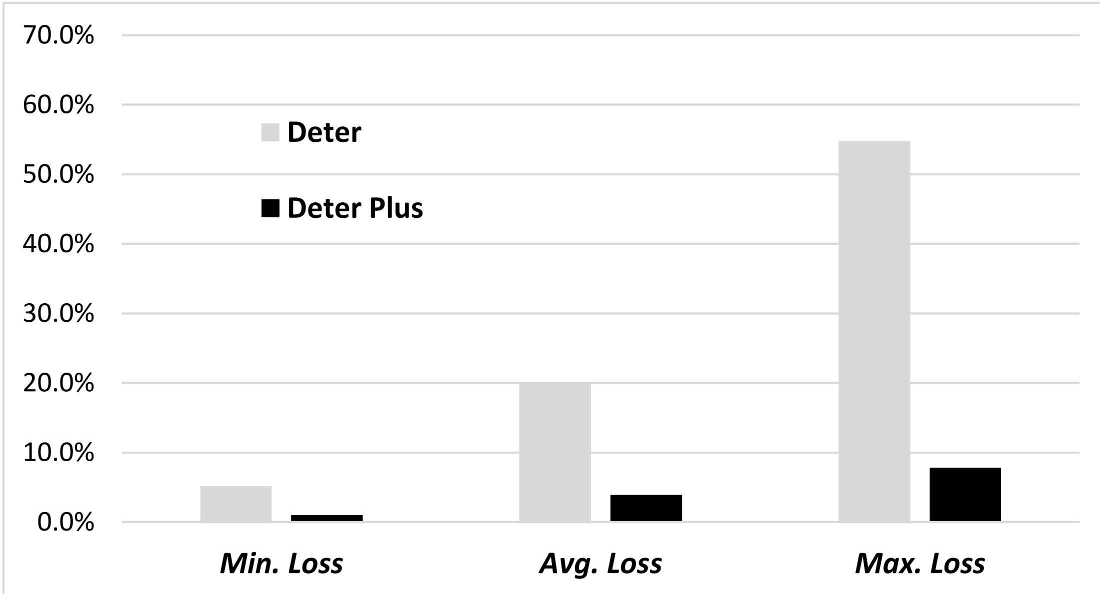

Figure 2: Comparison of performances of the **Deter** and **Deter Plus** designs in the fixed capacity model. Results are measured by minimum, average and maximum losses relative to the stochastic (optimal) solution.

worst case).

In the above tests we characterize demand stochasticity for each commodity using a symmetric distribution, which is replaced by its mean (i.e. the 50^{th} percentile of the distribution) in the deterministic case. So, vaguely stated, there is a chance that only 50% of the cases can be handled by the design; in the other 50% expensive ad-hoc capacity may be needed. However, depending on the ratio of total demand to total capacity provided in the deterministic solution, more demands in the stochastic case can be delivered using the deterministic design. In those scenarios where some demands cannot be satisfied with the deterministic design, the more expensive ad hoc capacity must be used, which translates into the losses reflected in Figure 2 for the **Deter** bars: about 20% on average and 55% at the highest. However, if the deterministic design is allowed to be expanded with extra services, these unmet demands may use the relatively cheaper extra services instead of ad hoc capacities, hence the lower losses for **Deter Plus**. The fact that all losses for **Deter Plus** are extremely small therefore shows the upgradeability of deterministic solutions.

In our experiments, losses of the deterministic designs in the stochastic environment are primarily caused by insufficient capacities. We therefore test for the deterministic designs produced with the demand of each commodity taking the value of the 75^{th} percentile of its corresponding distribution. This is common practice in many industries. Our results show that when using the 75^{th} percentile of every uncertain demand as its deterministic value, the average loss of the corresponding deterministic solution used in the stochastic environment (test **Deter**) drops from 19.96% to 11.23%. The **Deter Plus** tests are also performed for the 75^{th} percentile cases, and an average loss of 2.55% is observed, which also indicate the upgradeability of the deterministic solutions. The detailed results are reported in Table 1 at the end of Section 5.

Mathematically speaking, the difficulty of performing the **Deter Plus** test in our model is on par with solving the original stochastic problem to optimality. However, the actual difficulty depends on the specific instance. But on a complete service network, its complexity is not reduced much by fixing a relatively small number of $\{0,1\}$ decision variables, as it is still a big MIP when we allow other services to be opened. However, the fact that the deterministic design can be upgraded into an extremely good solution shows that the investments in the deterministic design are not wasted. In a highly dynamic transportation industry, it means that decision makers can sometimes safely invest on some services well ahead of time, especially if a discount is applicable by doing so. This is also a good way to reduce risks when the cost of setting up services is expected to be highly uncertain in the future or even go up closer to the time when one has to make the final plan.

Similar observations are made with the expected value approach in some other types of problems as well, we refer the interested readers to Maggioni and Wallace (2012) for more details. Note though, that it is not at all obvious that deterministic solutions are upgradeable. Also that is illustrated by Maggioni and Wallace.

4.3.2 Structural Differences

In consolidation-based freight transportation, consolidation is traditionally considered to be a way to accommodate the fact that most vehicles would not be full should each pair of terminals be linked by a direct service. Commodities are grouped, consolidated and then shipped together to avail of services with higher capacity and frequency between local hubs. In a stochastic setting, consolidation can also be induced by the need to hedge against demand uncertainties. This is shown in Lium et al. (2009) where, in particular, more hub-and-spoke structures are observed after demand stochasticity is explicitly considered. Also, there is usually more than one path for each commodity and more commodities are sharing paths with each other. This is also true in our model. In the stochastic (optimal) solution, we observe more consolidation activities compared to the deterministic solution.

Therefore, when used in the stochastic environment, we wonder: is the consolidation level of the deterministic solution (observed when used to satisfy the second stage scenarios) also lower than for the optimal solution? If the consolidation levels of **Deter**, **Deter Plus** and the optimal solutions can be quantitatively measured, we may find a correlation between the level of consolidation and the performance of the associated solution. In other words, we want to investigate the effect of the following structural feature of a design on its performance in the stochastic environment: the potential to allow a higher level of consolidation.

However, to precisely define the “the level of consolidation” allowed by a design is difficult. It is thus hard to find a straightforward way to quantitatively determine the potential of a design to allow a higher level of consolidation. Therefore, we propose a scheme to measure two substitute phenomena: the levels of *multi-path usage* and *path-sharing* when the design is used in the stochastic environment. If more commodities are using multiple paths to reach their respective destinations, and more services in the network are shared by several commodities, then potentially more consolidation activities should take place.

For a given solution, we count the number of paths each commodity is using and then produce a histogram to display the frequencies (in terms of number of commodities) *with all instances added up*. We use small instances so that we can obtain the true optimal solution for each of the instances, we then add up the counting results of all instances. For example, if we have 10 instances, each with 16 commodities, we count this as 160 commodities in the statistics. We then count how many of these commodities travel on one, two, three, and so on paths.

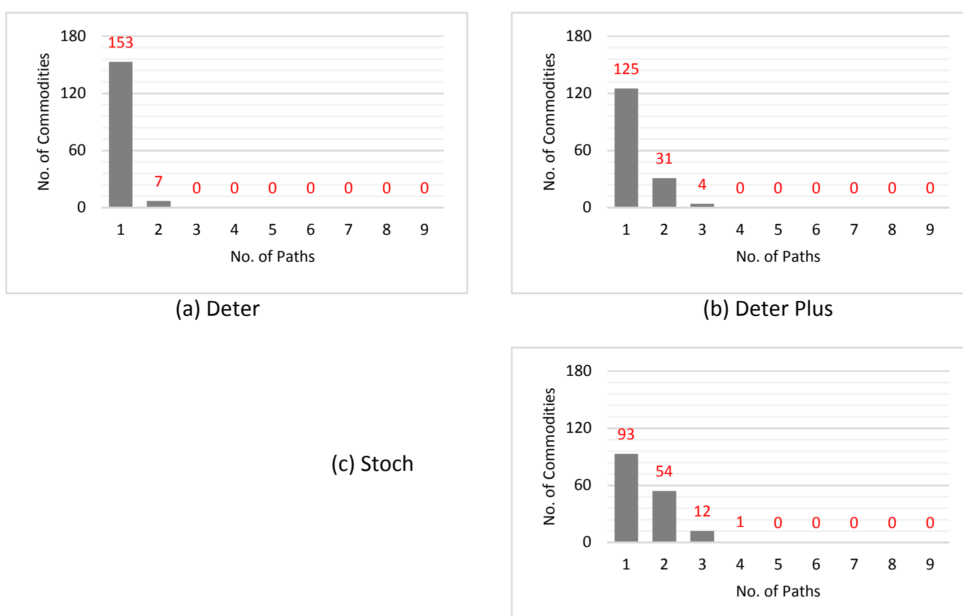

Figure 3: Measuring the level of multi-path usage for the fixed capacity model.

Figure 3 presents the level of multi-path usage measured by commodity counts, in the **Deter**, the **Deter Plus** (deterministic design with extra services) and the stochastic (optimal) cases. In the first case, there are 153 commodities using only one path and 7 commodities which use 2 paths to reach their destinations. The number of commodities using 2 paths rises to 54 in the stochastic case, and there are even 12 commodities using 3 paths and 1 commodity using 4 paths while the number for only one path has dropped from 153 to 93. From Figure 3 (b) we can also see a significant increase in the number of commodities using multiple paths compared to the **Deter** case, yet lower compared to the stochastic case.

If we define the levels of multiple-paths usage measured for **Deter**, **Deter Plus** and **Stoch** as low, medium and high, then when compared against their performances in the stochastic environment we see a trend. (See Figure 2 for the performances of **Deter** and **Deter Plus**; **Stoch**, as the optimal solution, will of course produce 0% losses). That is, the better the solution performance, the higher the level of multiple-path usage. Considering the great improvement in performance from **Deter** to **Deter Plus**, this also indicates that with some new arcs opened, the deterministic design is able to evolve to a structurally different design that allow a higher level of multi-path usage and become very competitive for the stochastic problem.

Similar insights can be drawn when measuring the level of path-sharing. We do this by counting the number of commodities routed through each opened service. Note that a commodity may be routed though a number of services to reach its destination. We thus say that if two commodities have at least one service in common, they are sharing paths.

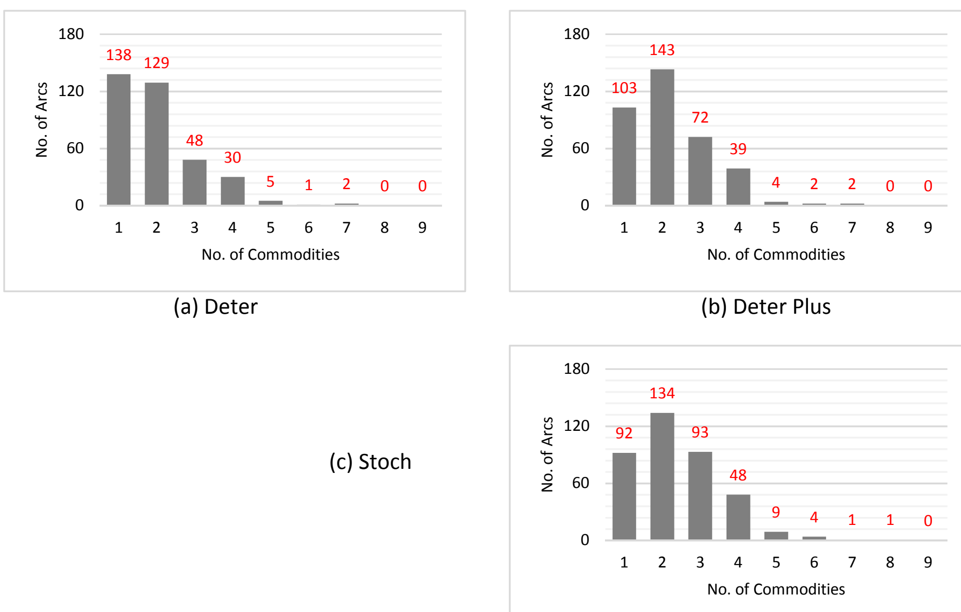

Figure 4: Measuring the level of path-sharing for the fixed capacity model.

The results of path-sharing measurements are displayed in Figure 4. In the **Deter** case, 129 arcs are shared by 2 commodities and 48 arcs are shared by 3 commodities.

These two counts increase to 143 and 72 in the **Deter Plus** case. In the **Stoch** case, the number of arcs shared by 2 commodities stays at a similar level (134) while the number of arcs shared by 3 commodities increase further to 93, and the number of arcs shared by 4 commodities reaches 48. In general, we can see a right shift of the frequency curve, from **Deter** to **Deter Plus** and then to the stochastic case, while the performance of the corresponding solution improves in the stochastic environment, indicating that the better the solution performs the higher level of path-sharing it has.

The above results confirm two structural features for our model: it is potentially beneficial to have a design structure that allows high levels of multi-path usage and path-sharing. Furthermore, with some extra services, the deterministic solution can be structurally changed in terms of its potential to allow higher levels of these two phenomena, and become much better suited to handle the stochastic demands. So how many extra services are required to make the change?

First of all, our results show that, based on the ten test instances, there are 70%-90% overlap of arcs between the stochastic and deterministic solutions. It means that most of the service selection decisions of the stochastic solution are shared with the deterministic counterpart, but it includes additional arcs (services) to obtain a structure with much higher flexibility to handle demand variations through higher levels of multi-path usage and path-sharing.

Our numbers also show that, on average, around 15% extra arcs are added to the deterministic design in the **Deter Plus** test. Therefore, by adding a limited number of extra arcs, the deterministic design can become structurally different, and much better suited for the stochastic environment. So what can we do to find the right extra arcs? As mentioned earlier, on a complete network, the difficulty to find these extra arcs numerically can be on par with solving the original stochastic program. However, if a heuristic approach is used to obtain the solution to the deterministic version of the problem, we may already have some potentially useful information to start with. For example, one can target those arcs which are not part of the final solution but had the longest stay inside the incumbent solutions during the process of the heuristic, or, have entered the candidate list with the highest frequencies.

5 The Model with Variable Capacities

In the previous model, where the capacity for each service is fixed, the deterministic solutions themselves are rather good at handling demand uncertainty. Moreover, we see further improvements using the **Deter Plus** approach. Yet, in terms of optimization complexity, it is still at the same level as solving the original stochastic program to optimality. In this section, we present a variant of the classical service network design formulation, for which we introduce a way of using *parts* of the deterministic design to produce acceptable solutions for the stochastic problem with a computational effort significantly reduced compared to solving the original stochastic program.

5.1 Variable Capacity Service Network Design

We now consider a scheduled service network design model where the maximum service capacity, $h_{ij;t}, (i,j;t) \in \mathcal{A}$, is built of a number of units to be determined when the plan is built. This concerns, e.g., rail cars making up a block or train, trailers in a multi-trailer trucking service, barges in a barge-train. For simplicity's sake, we assume all units making up a service have equal capacity, $u_{ij;t}$ (set to 1 in our experiments). The cost of adding one unit of service capacity is represented by $c_{ij;t}, (i,j;t) \in \mathcal{A}$. To our best knowledge, this problem setting has not been studied before.

We define the integer decision variables $X_{ij;t}$ to represent the number of units of capacity for service $(i,j;t) \in \mathcal{A}$. The other decision variables are the same as in the classical model capturing the service selection choices, indicating whether or not the service leaves at the specified time point ($V_{ij;t}$), the continuous flow of commodity k on arc $(i,j;t)$ in scenario s ($Y^s_{ij;t;k}$), and the continuous volume of commodity k that uses ad hoc handling in scenario s (Z^s_k). The formulation then becomes:

$$\min \sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{N}}\sum_{t\in\mathcal{T}} f_{ij;t}V_{ij;t} + \sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{N}}\sum_{t\in\mathcal{T}} c_{ij;t}X_{ij;t} + \sum_{s\in\mathcal{S}} p^s\Big(\sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{N}}\sum_{t\in\mathcal{T}}\sum_{k\in\mathcal{K}} e_{ij;t;k}Y^s_{ij;t;k} + \sum_{k\in\mathcal{K}} b_k Z^s_k\Big) \tag{7}$$

$$\sum_{i\in\mathcal{N}} Y^s_{ij;t\ominus l_{ij};k} - \sum_{i\in\mathcal{N}} Y^s_{ji;t;k} = \begin{cases} \delta^s_k - Z^s_k, & \text{if } j = d_k \text{ and } t = \tau_k \\ -\delta^s_k + Z^s_k, & \text{if } j = o_k \text{ and } t = \sigma_k \\ 0, & \text{other} \end{cases}$$
$$, \forall j \in \mathcal{N}, \forall t \in \mathcal{T}, \forall k \in \mathcal{K}, \forall s \in \mathcal{S} \tag{8}$$

$$\sum_{k\in\mathcal{K}} Y^s_{ij;t;k} \le u_{ij;t}X_{ij;t} \qquad \forall i,j \in \mathcal{N}, i \neq j, \forall t \in \mathcal{T}, \forall s \in \mathcal{S} \tag{9}$$

$$0 \le u_{ij;t}X_{ij;t} \le h_{ij;t}V_{ij;t} \qquad \forall i,j \in \mathcal{N}, \forall t \in \mathcal{T} \tag{10}$$

$$0 \le Z^s_k \le \delta^s_k \qquad \forall k \in \mathcal{K}, \forall s \in \mathcal{S} \tag{11}$$

$$0 \le Y^s_{ij;t;k} \le \delta^s_k \qquad \forall i,j \in \mathcal{N}, \forall t \in \mathcal{T}, \forall k \in \mathcal{K}, \forall s \in \mathcal{S} \tag{12}$$

The objective function (7) minimizes the total cost for offering services and providing service capacities, plus the expected cost for moving or holding commodities and using ad hoc handling. Constraints (4) in the fixed capacity model was replaced by constraints (9) and (10). Constraints (9) make sure the total flow on each arc does not exceed the provided capacity, which is now also a decision variable. Constraints (10) ensure the capacity limit on each arc is not exceeded.

5.2 Experiments and Results

For the variant model considered in this section, and different from the classical version, the information of a particular solution contains two components: the *skeleton*, that is, the service selection information, and the capacities provided on the services selected. We thus want to find out: (a) How good (or bad) will the performance be if we fully use the deterministic solution (skeleton & capacities) in the stochastic environment? Will it be the same as for the first model? (b) Can we use only parts of the deterministic design, in particular, the deterministic skeleton and produce a new and well-performing solution by setting the capacities separately? (c) How much extra effort is needed to determine the capacities (in terms of computational complexity)? In other words, what is the value and viability of such an approach?

We therefore introduce another test to value the deterministic skeleton, named the **Deter Skl** test: Start by solving the deterministic problem; Fix the service selection variables $V_{ij;t}$ only; Then run a MIP to set the capacities $X_{ij;t}$ and the flows.

The same ten instances of the previous section are used to perform the tests here. We first compare the performances of the following three types of solutions in the stochastic environment: the deterministic solution (**Deter**), the deterministic skeleton with updated capacities (**Deter Skl**) and the deterministic design with extra services (**Deter Plus**). The alternative tests with the deterministic demand taking the value of the 75^{th} percentile of the associated distribution are also performed. The results are reported in Table 1 in the next subsection. Note that in the **Deter Plus** test, for the variable capacity model, we see both services set up and capacities provided on these services as "invested". We however still allow more capacity to be offered on these selected services, as long as their corresponding capacity limits are respected. We also allow extra services to be set up apart from the selected services.

Figure 5 shows that the deterministic solution (**Deter**) is quite bad in the stochastic setting, while **Deter Skl** behaves much better. Although the maximum loss of **Deter Skl** is still high (over 40%), its average loss (around 15%) is quite acceptable. On the other hand, in the **Deter** tests, the average loss goes over 55%, and even the minimum loss is nearly 40%, when the deterministic solution is directly used. Of course, the **Deter Plus** approach offers the best performance but its computational effort could be high.

The results show that, in general, the deterministic solution does not handle well demand uncertainty when the capacity is not fixed a priori, but inheriting the skeleton of the deterministic solution is beneficial in most circumstances. This is very well illustrated by comparing the **Deter** performances in Figures 2 and 5. This may be explained by the possibility in the variable-capacity model to closely adjust the supplied capacity to the demand. This capability is very useful for a deterministic setting but not when evaluating the deterministic solution in a stochastic setting. Indeed, adjusting the capacity to the

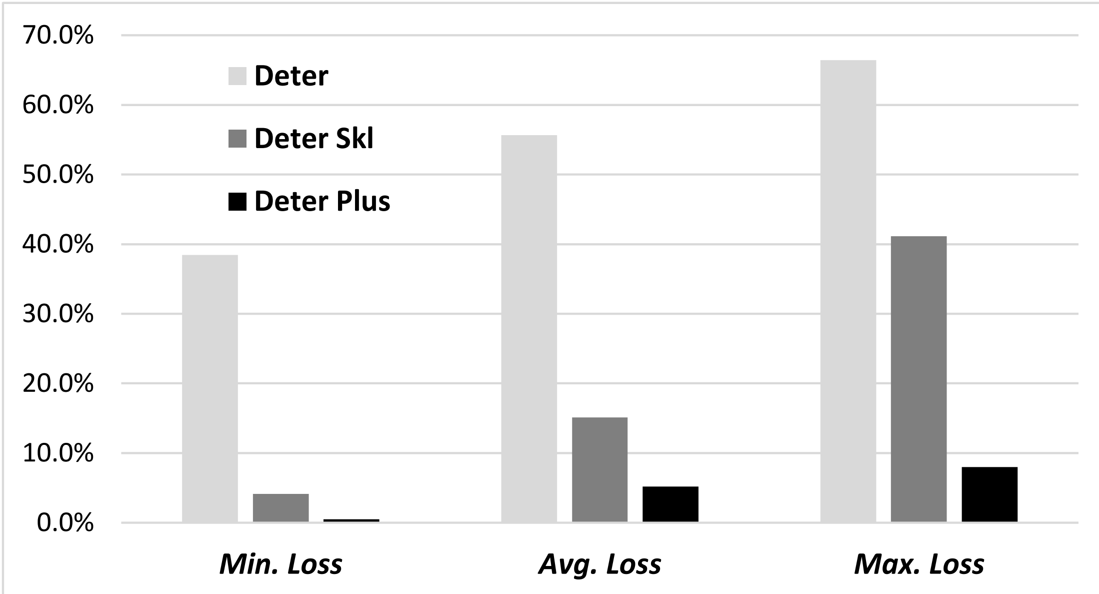

Figure 5: Comparison of **Deter**, **Deter Skl** and **Deter Plus** designs in the variable capacity model. Results are measured by minimum, average and maximum losses relative to the stochastic (optimal) solution.

estimated demand results in little extra capacity available when the observed demand is higher than the prediction, which come at the price of much ad-hoc capacity used. The results reported in Table 1 are extremely telling in this context, the performance of the **Deter** approach improving dramatically (threefold) when the 75^{th} percentile of the demand distribution is used as forecast. The performance of the skeleton-based solution is still better, but the two approaches are more at par in that situation, as the improvement of **Deter Skl** is less important. Notice that the last observation points to the fact that this approach could be more "forgiving" of a bad demand estimation. On the other hand, the performance of **Deter Plus** is fundamentally constant.

Following a similar thinking as in the previous section to attempt to further explain the performance improvement of the **Deter Skl** and **Deter Plus** methods compared to directly using the deterministic solution, we use the measurement scheme proposed earlier to observe the quantitative structural changes.

We may draw similar conclusions from the results displayed in Figures 6 and 7 as in the fixed capacity model: the better the solution performs, the higher level of multiple-paths usage and path-sharing it has. This is clearly visible from the numbers and the charts.

But if we consider the changes from **Deter** to **Deter Skl** (they have the same service selection decisions, but provide different capacities), we can see some interesting similarities, in contrast to the updates from **Deter** to **Deter Plus** in the fixed capacity

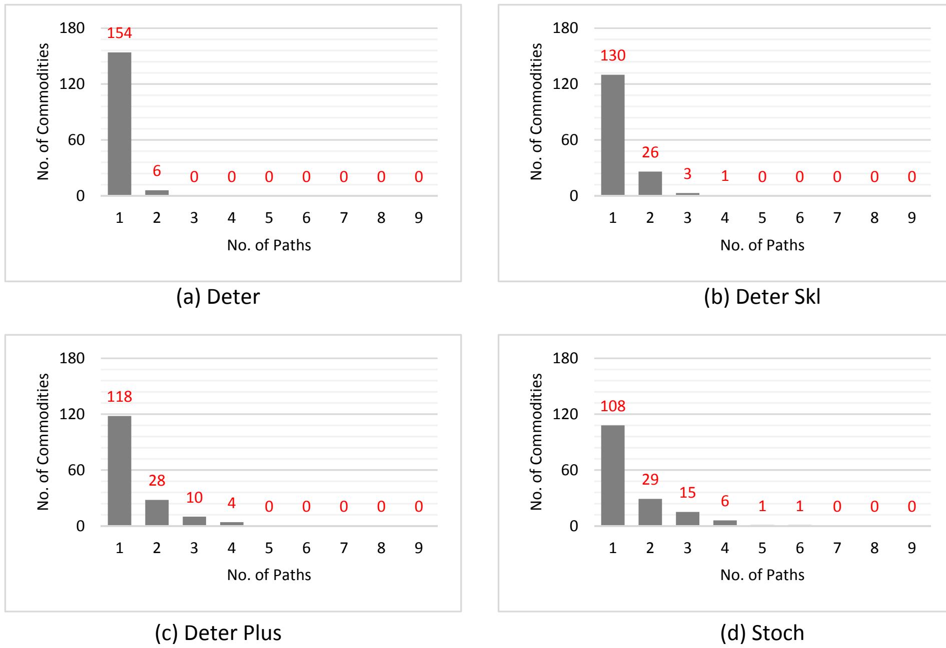

Figure 6: Measuring the level of multi-path usage for the variable capacity model.

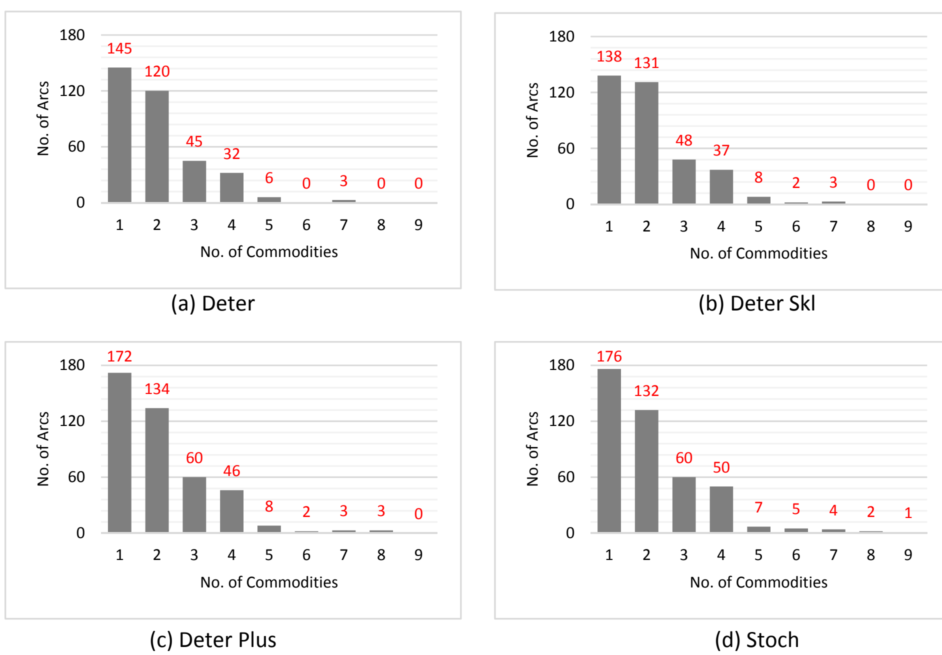

Figure 7: Measuring the level of path-sharing for the variable capacity model.

model. Rather than to allow other services to be opened, **Deter Skl** merely changes the capacities provided on the already selected services. It still brings out similar *structural improvements*, allowing higher levels of multi-path usage and path-sharing. We conclude that a design based on the deterministic skeleton is able to adapt itself structurally to uncertainty even when its options are highly limited.

In **Deter Skl**, capacities are only allowed to be provided on the deterministic skeleton. So, essentially, the original *complete* network is "shrunk" to a smaller deterministic skeleton network (how much smaller depends on the problem instance and its ratio of total demand to the total capacity that may be offered on the service network). Then, the possibility of finding a new path for a given commodity is subject to whether there happens to be another combination of services (apart from the deterministic one) on the reduced network to take it from its origin to its destination. If yes, then in those scenarios where the commodity's demand is very high, it might use the new path as long as there is free capacity on this path. In Figure 6, we see the number of commodities using two paths quadrupled from 6 to 26.

A noteworthy observation is that the better the solution performs, the higher the levels of multiple-path usage and path-sharing are, but *not* vice versa. There is an obvious counter-example. If we enforce very tight capacity limits on all possible services, we can obtain a solution with an extremely high level of multiple-path usage and path-sharing, as all the commodities would have to find many paths trying to avoid expensive ad hoc handling. This might result in opening a large number of services and evidently very poor performance.

5.3 Effect of Continuous Capacities

We now turn to the continuous-capacity case, relaxing the integrality constraints of the previous model. This may correspond to an approximation of actual integer capacities (could be appropriate when capacities are large) or to applications in different fields where capacities are actually continuous.

The computational effort to perform the **Deter Skl** test (which yields well-performing designs based on deterministic solutions) can be much lower if capacity variables $X_{ij;t}$ become continuous. Given the deterministic solution, the **Deter Skl** method fixes the service selection variables $V_{ij;t}$ and determines the capacities by solving a stochastic LP. Therefore this approach can be seen as a viable heuristic.

The performances of the **Deter**, **Deter Skl** and **Deter Plus** approaches for the variable-capacity model with integer and continuous capacities are displayed in Table 1. Test results with deterministic demands set at the 50^{th} and 75^{th} percentiles of their corresponding distributions are shown. The same instances are used for every row in the

Table 1: Average loss in the stochastic environment

Model and Parameter Setting	Average Loss		
	Deter	Deter Skl	Deter Plus
Fixed Capacity, 50^{th}	19.96%		3.91%
Fixed Capacity, 75^{th}	11.23%		2.55%
Variable Integer Capacity, 50^{th}	55.67%	15.14%	5.18%
Variable Integer Capacity, 75^{th}	12.63%	8.22%	4.37%
Variable Continuous Capacity, 50^{th}	57.77%	15.90%	4.96%
Variable Continuous Capacity, 75^{th}	13.48%	9.09%	4.72%

table.

Comparing the average losses for the 50^{th} and 75^{th} percentile deterministic demand settings for all problem settings, shows that the 75^{th} percentile always produces better performances. This of course depends on problem settings, in particular, how much more expensive the ad hoc capacity is. When there is a surge in demand, extra capacity with higher price must be paid for to compensate the insufficiency of regular services. Using the 75^{th} percentile to account for higher expected demands is thus usually a better choice when regular services and capacities are cheap and reliable. This approach is in line with what is used in many industries; using demands well above the mean.

6 Asset Balance Considerations

As described and reviewed in Section 2, resource-management considerations are increasingly, accounted for within service network design models. We therefore introduce asset-balance requirements in the fixed and variable-capacity models.

The asset-balance requirements in the fixed-capacity case take the form

$$\sum_{i\in\mathcal{N}} V_{ij;t\ominus l_{ij}} = \sum_{i\in\mathcal{N}} V_{ji;t} \quad \forall j \in \mathcal{N}, \forall t \in \mathcal{T}. \tag{13}$$

Notice that this was the model used in Lium et al. (2007, 2009). Also notice that, as we use the same capacity $h_{ij;t}$ for all services, equation (13) also balances the total capacity going in and out of each node across the space-time network, which is also true in the next case.

When the assets controlled correspond to the number of services (e.g., power units, ships, etc.), equation (13) may be also used within the variable-capacity formulation. When, on the other hand, the controlled assets are the units of capacity, the constraints have to be written in the appropriate units as in equation (14) where, for simplicity of presentation we assume all units are the same for all services.

$$\sum_{i \in \mathcal{N}} X_{ij;t \ominus l_{ij}} = \sum_{i \in \mathcal{N}} X_{ji;t} \qquad \forall j \in \mathcal{N}, \forall t \in \mathcal{T} \tag{14}$$

Table 2 displays the results of the experimentation performed with the modified formulations using the same instances as before. Again, using the 75^{th} percentile of the demand distribution is a better choice when obtaining the deterministic solution. For the fixed and the variable capacity models, with both integer and continuous capacity settings, **Deter** (75^{th}) produces average losses that are all less than 10% in this case. The **Deter Skl** method can further improve the performance of the solution with not much computational efforts: a much smaller MIP for the integer capacity case and an LP for the continuous capacity case, both on a reduced deterministic skeleton network.

Table 2: Average loss in the stochastic environment with asset balance

Model and Parameter Setting	Average Loss		
(with asset balance)	Deter	Deter Skl	Deter Plus
Fixed Capacity, 50^{th}	18.61%		4.62%
Fixed Capacity, 75^{th}	9.01%		3.77%
Variable Integer Capacity, 50^{th}	27.19%	11.91%	5.03%
Variable Integer Capacity, 75^{th}	8.47%	6.68%	3.85%
Variable Continuous Capacity, 50^{th}	26.36%	12.80%	4.75%
Variable Continuous Capacity, 75^{th}	9.55 %	6.20%	4.13%

7 Conclusion

In this paper we discussed the value and the upgradeability of the deterministic solution in the scheduled stochastic service network design problem, for the fixed and the variable capacity models with both integer and continuous capacity settings. In those situations where deterministic solutions can be found, optimally or heuristically, we may upgrade these solutions into much better performing ones to the stochastic problem.

For the fixed capacity model, by adding a limited number of extra arcs, the deterministic design can become structurally different, and much better suited for the stochastic environment. For the variable capacity model, this can also be achieved by using part of the deterministic design information (the deterministic skeleton) and also with not much computational efforts. In particular, when the capacities are continuous, the **Deter Skl** method becomes an LP on a reduced skeleton network. We also showed that it is a better practice to use the 75^{th} percentile of the random demands when obtaining the deterministic solutions.

To quantitatively investigate the structural improvements from the deterministic design to better performing solutions in the stochastic environment, a measurement scheme has been proposed to evaluate the level of the potentially beneficial structural features: multi-path usage and path-sharing. It was concluded that, in general, the better the solution performs in the stochastic environment, the higher the level of multiple-path usage and path-sharing it displays. The reverse is not true, but still, this might lead to possible ways to develop heuristic approaches for the stochastic problem.

Therefore, an interesting direction of future research may be identified: can we find the "correct" extra services based on the deterministic solution (or even a feasible solution), using the beneficial structural features confirmed in this paper? For example, if certain services increase the level of multi-path usage and path-sharing in the network, then these might be the potentially "correct" extra services for the stochastic problem.

Another research avenue is to investigate the existence of similar upgradeability of deterministic solutions in general network design problems. As mentioned earlier, such upgradeability is not obvious at all in some other stochastic problems (Maggioni and Wallace, 2012). We may be able to determine in what circumstances the deterministic solution is useful in the stochastic environment, if a certain modeling factor is found to have great impact on the upgradeability of the deterministic solution.

Acknowledgments

Partial funding for this project has been provided by the Natural Sciences and Engineering Council of Canada (NSERC), through its Industrial Research Chair and Discovery Grant programs.

References

Andersen, J., Christiansen, M., Crainic, T. G., and Gronhaug, R. (2011). Branch and price for service network design with asset management constraints. *Transportation Science*, 45(1):33–49.

Andersen, J., Crainic, T. G., and Christiansen, M. (2009a). Service network design with asset management: Formulations and comparative analyses. *Transportation Research Part C: Emerging Technologies*, 17(2):197 – 207.

Andersen, J., Crainic, T. G., and Christiansen, M. (2009b). Service network design with management and coordination of multiple fleets. *European Journal of Operational Research*, 193(2):377–389.

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

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

Barnhart, C. and Schneur, R. (1996). Network design for express freight service. *Operations Research*, 44(6):852–863.

Birge, J. R. (1982). The value of the stochastic solution in stochastic linear programming with fixed recourse. *Mathematical Programming*, 24:314–325.

Buedenbender, K., Grünert, T., and Sebastian, H.-J. (2000). A hybrid tabu search/branch and bound algorithm for the direct flight network design problem. *Transportation Science*, 34(4):364–380.

Christiansen, M., Fagerholt, K., Nygreen, B., and Ronen, D. (2007). Maritime transportation. In Barnhart, C. and Laporte, G., editors, *Transportation*, volume 14 of *Handbooks in Operations Research and Management Science*, chapter 4, pages 189–284. North-Holland, Amsterdam.

Christiansen, M., Fagerholt, K., and Ronen, D. (2004). Ship routing and scheduling: Status and perspectives. *Transportation Science*, 38(1):1–18.

Cordeau, J.-F., Toth, P., and Vigo, D. (1998). A survey of optimization models for train routing and scheduling. *Transportation Science*, 32(4):380–404.

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

Crainic, T. G. (2003). Long haul freight transportation. In Hall, R., editor, *Handbook of Transportation Science*, pages 451–516. Kluwer Academic Publishers, Norwell, MA, second edition.

Crainic, T. G., Dejax, P., and Delorme, L. (1989). Models for multimode multicommodity location problems with interdepot balancing requirements. *Annals of Operations Research*, 18:279–302.

Crainic, T. G., Ferland, J.-A., and Rousseau, J.-M. (1984). A tactical planning model for rail freight transportation. *Transportation Science*, 18(2):165–184.

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. and Kim, K. H. (2007). Intermodal transportation. In Barnhart, C. and Laporte, G., editors, *Transportation*, volume 14 of *Handbooks in Operations Research and Management Science*, chapter 8, pages 467–537. North-Holland, Amsterdam.

Crainic, T. G. and Laporte, G. (1997). *Design and Operation of Civil and Environmental Engineering S ystems*, chapter Planning Models for Freight Transportation, pages 343–394. Wiley-Interscience, New York.

Crainic, T. G. and Rousseau, J.-M. (1986). Multicommodity, multimode freight transportation: A general modeling and algorithmic framework for the service network design problem. *Transportation Research B: Methodology*, 20B:225–242.

Crainic, T. G. and Roy, J. (1988). O.R. Tools for Tactical Freight Transportation Planning. *European Journal of Operational Research*, 33(3):290–297.

Dejax, P. and Crainic, T. G. (1987). A review of empty flows and fleet management models in freight transportation. *Transportation Science*, 21(4):227–247.

Delorme, L., Roy, J., and Rousseau, J.-M. (1988). Motor-carrier operation planning models: A state of the art. In Bianco, L. and Bella, A. L., editors, *Freight Transport Planning and Logistics*, pages 510–545. Springer-Verlag, Berlin.

Farvolden, J. and Powell, W. (1991). A dynamic network model for less-than-truckload motor carrier operations. Working Paper 90-05, Department of Industrial Engineering, University of Toronto, Toronto, ON, Canada.

Farvolden, J. and Powell, W. (1994). Subgradient methods for the service network design problem. *Transportation Science*, 28(3):256–272.

Farvolden, J., Powell, W., and Lustig, I. (1992). A Primal Partitioning Solution for the Arc–Chain Formulation of a Multicommodity Network Flow Problem. *Operations Research*, 41(4):669–694.

Gorman, M. (1998a). An application of genetic and tabu searches to the freight railroad operating plan problem. *Annals of Operations Research*, 78:51–69.

Gorman, M. (1998b). Santa fe railway uses an operating-plan model to improve its service design. *Interfaces*, 28(4):1–12.

Grünert, T. and Sebastian, H.-J. (2000). Planning models for long-haul operations of postal and express shipment companies. *European Journal of Operational Research*, 122:289–309.

Grünert, T., Sebastian, H.-J., and Thäerigen, M. (1999). The design of a letter-mail transportation network by intelligent techniques. In Sprague, R., editor, *Proceedings Hawaii International Conference on System Sciences 32*.

Haghani, A. (1989). Formulation and solution of combined train routing and makeup, and empty car distribution model. *Transportation Research B: Methodological*, 23B(6):433–452.

Higle, J. L. and Wallace, S. W. (2003). Sensitivity analysis and uncertainty in linear programming. *Interfaces*, 33:53–60.

Hoff, A., Lium, A.-G., Løkketangen, A., and Crainic, T. G. (2010). A metaheuristic for stochastic service network design. *Journal of Heuristics*, 16(5):653–679.

Høyland, K., Kaut, M., and Wallace, S. W. (2003). A heuristic for moment-matching scenario generation. *Computational Optimization and Applications*, 24(2–3):169–185.

Kaut, M. and Wallace, S. W. (2007). Evaluation of scenario-generation methods for stochastic programming. *Pacific Journal of Optimization*, 3(2):257–271.

Keaton, M. (1989). Designing optimal railroad operating plans: Lagrangian relaxation and heuristic approaches. *Transportation Research B: Methodological*, 23B(6):415–431.

Keaton, M. (1991). Service–cost tradeoffs for carload freight traffic in the u.s. rail industry. *Transportation Research A: Policy and Practice*, 25A(6):363–374.

Keaton, M. (1992). Designing optimal railroad operating plans: A dual adjustment method for implementing lagrangian relaxation. *Transportation Science*, 26:263–279.

Kim, D., Barnhart, C., Ware, K., and Reinhardt, G. (1999). Multimodal express package delivery: A service network design application. *Transportation Science*, 33(4):391–407.

Lamar, B., Sheffi, Y., and Powell, W. (1990). A capacity improvement lower bound for fixed charge network design problems. *Operations Research*, 38(4):704–710.

Lium, A.-G., Crainic, T. G., and Wallace, S. W. (2007). Correlations in stochastic programming: A case from stochastic service network design. *Asia-Pacific Journal of Operational Research*, 24(2):161–179.

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

Maggioni, F. and Wallace, S. W. (2012). Analyzing the quality of the expected value solution in stochastic programming. *Annals of Operations Research*, 200(1).

Newton, H. (1996). *Network Design Under Budget Constraints with Application to the Railroad Blocking Problem*. PhD thesis, Industrial and Systems Engineering, Auburn University, Auburn, Alabama, U.S.A.

Newton, H., Barnhart, C., and Vance, P. (1998). Constructing railroad blocking plans to minimize handling costs. *Transportation Science*, 32(4):330–345.

Pedersen, M., Crainic, T. G., and Madsen, O. (2009). Models and tabu search metaheuristics for service network design with asset-balance requirements. *Transportation Science*, 43(2):158–177.

Powell, W. (1986). A local improvement heuristic for the design of less-than-truckload motor carrier networks. *Transportation Science*, 20(4):246–357.

Powell, W. and Sheffi, Y. (1983). The load-planning problem of motor carriers: Problem description and a proposed solution approach. *Transportation Research A: Policy and Practice*, 17(6):471–480.

Powell, W. and Sheffi, Y. (1986). Interactive optimization for motor carrier load planning. *Journal of Business Logistics*, 7(2):64–90.

Powell, W. and Sheffi, Y. (1989). Design and implementation of an interactive optimization system for the network design in the motor carrier industry. *Operations Research*, 37(1):12–29.

Roy, J. and Delorme, L. (1989). Netplan: A network optimization model for tactical planning in the less-than-truckload motor-carrier industry. *INFOR*, 27(1):22–35.

Sharypova, M., Crainic, T. G., van Woensel, T., and Fransoo, J. (2012). Scheduled Service Network Design with Synchronization and Transshipment Constraints for Intermodal Container Transportation Networks. Publication CIRRELT-2012-77, Centre interuniversitaire de recherche sur les réseaux d'entreprise, la logistique et le transport, Montréal, QC, Canada.

Smilowitz, K., Atamtürk, A., and Daganzo, C. (2003). Deferred item and vehicle routing within integrated networks. *Transportation Research Part E: Logistics and Transportation*, 39:305–323.

Thapalia, B. K., Crainic, T. G., Kaut, M., and Wallace, S. W. (2011). Single-commodity stochastic network design with multiple sources and sinks. *INFOR: Information Systems and Operational Research*, 49(3):193–211.

Thapalia, B. K., Crainic, T. G., Kaut, M., and Wallace, S. W. (2012a). Single-commodity network design with random edge capacities. *European Journal of Operational Research*, 220(2):394–403.

Thapalia, B. K., Crainic, T. G., Kaut, M., and Wallace, S. W. (2012b). Single source single-commodity stochastic network design. *Computational Management Science*, 9(1):139–160. Special issue on 'Optimal decision making under uncertainty'.

Wallace, S. (2000). Decision making under uncertainty: Is sensitivity analysis of any use? *Operations Research*, 48:20–25.

Wallace, S. W. (2010). Stochastic programming and the option to do it differently. *Annals of Operations Research*, 177(1):3–8.

Zhu, E., Crainic, T. G., and Gendreau, M. (2013). Scheduled service network design for freight rail transportation. *Operations Research*, 62(2):383–400.