Computers & Industrial Engineering 126 (2018) 99–110

Contents lists available at ScienceDirect

Computers & Industrial Engineering

journal homepage: www.elsevier.com/locate/caie

A simulation-optimization approach for scheduling in stochastic freight transportation

Safa Bhar Layeb, Amel Jaoua*, Asma Jbira, Yosra Makhlouf

UR-OASIS, National Engineering School of Tunis, University of Tunis El Manar, Tunis, Tunisia

ARTICLE INFO

Keywords:
Freight transportation
Scheduling
Stochastic behavior
Simulation-optimization

ABSTRACT

In this paper, a new Simulation-Based Optimization Model (SBO-Model) is proposed to solve scheduling problem in stochastic multimodal freight transportation systems. The model is applied to find optimal services schedule in a real-world case study. In order to handle demand and travel time inherent variability, the stochastic service network design problem is addressed. Simulation modeling is used to efficiently account for real stochastic behavior with skewed continuous distributions. Such distinctive distribution shapes were commonly reported in transportation research studies that addressed the travel time reliability modeling. Results indicate that the SBO-Model can indeed provide reliable service schedules even under realistic complex stochasticity. The main finding is that, in order to solve efficiently such stochastic optimization problem, we need to go beyond the mean and variance estimates by considering the empirical distributions of uncertain parameters. Specifically, when the data exhibit skewness and/or multimodality, which are commonly found due to the traffic congestion.

The originality of this work lies in the integration of stochastic models, commonly used in the transportation research field, for solving logistics planning problem generally addressed by Operations Research community.

1. Introduction

Global freight transportation has grown exponentially in the context of the globalized economy. Multimodal freight transportation has become consequently a key success factor for logistics service firms. Multiple transportation modes such as railway, roadway and waterway services are combined for efficient containerized freight delivery. In this context, determining the optimal service schedule which satisfies the delivery requirements for a set of commodities has become a fundamental tactical planning problem (Wang & Wallace, 2016). This NP-hard optimization problem has attracted since early 90s a considerable amount of attention from the Operations Research community (Crainic, 2000). More recent research studies (e.g. Chávez, Castillo-Villar, Herrera, & Bustos, 2017; Demir et al., 2016, Lium, Crainic, & Wallace, 2009; Meng, Wang, & Wang, 2012) have shown that solving Service Network Design Problem (SND) with the assumption that planning elements, such as demand and travel time, are deterministic might result in a complete failure of the transportation planning. Therefore, they addressed the Stochastic SND problem to handle demand and travel times variability. Commonly, the problem is modeled as a multi-stage stochastic program where demand and travel times values are assumed to follow discrete distributions. The reliability of the found freight schedules is then evaluated under different levels of disturbances.

However, it is well established in roadways, railways and waterways transportation related literature, that travel times follow skewed continuous distributions and not discrete distributions (Chalumuri & Yasuo, 2014; Harrison & Fichtinger, 2013; Krüger, Vierth, & Fakhraei Roudsari, 2013; Rakha, El-Shawarby, & Arafeh, 2010; Wen et al., 2017). In the last decade, in order to develop travel time reliability metrics, the transportation research community has intensively studied travel time distribution fitting. They found that travel time's variability has to be modeled with skewed continuous distribution such as the Lognormal and Gamma distributions. Besides, very often due to congestion, travel times data in roadway and waterway exhibit bimodality. Then, it was necessary to recourse to Mixture distributions (Guessous, Aron, Bhouri, & Cohen, 2014; Harrison & Fichtinger, 2013; Krüger et al., 2013; Lee, Lee, & Zhang, 2015; Rakha et al., 2010; Taylor & Somenahalli, 2010; Wen et al., 2017). In addition, continuous distributions is commonly used to model demand variability (Garrido & Mahmassani, 2000; Juan, Grasman, Caceres-Cruz, & Bekta ş, 2014).

Therefore, our objective is to solve the Stochastic SND problem while considering such continuous distributions of demand and travel time variabilities. Under the proposed resolution scheme, these parameters are not deterministic values but they are random stochastic variables from the fitted probability distributions to their corresponding empirical data. The accurate modeling of such stochastic aspects aims

* Corresponding author.
E-mail address: amel.jaoua@polymtl.ca (A. Jaoua).

https://doi.org/10.1016/j.cie.2018.09.021
Received 12 October 2017; Received in revised form 14 July 2018; Accepted 9 September 2018
Available online 17 September 2018
0360-8352/

to construct more reliable schedule. This issue is a major research challenge in freight transportation planning. According to SteadieSeifi, Dellaert, Nuijten, Van Woensel, and Raoufi (2014), Hrušovský, Demir, Jammernegg, and Van Woensel (2016), Lee and Song (2017), Chávez et al. (2017), taking real stochasticity, i.e. of empirical data, into account will greatly improve schedule reliability in real-world applications. For that purpose, we do not try to develop another analytical model for solving such stochastic problem. Instead, we seek to implement a different modeling paradigm, the Simulation-Based Optimization (SBO) approach to solve the freight stochastic SND problem. This SBO approach, also known as Simulation-Optimization (Fu, 2002), has proven to be a very powerful tool to solve stochastic problems. According to Oliveira, Lima, and Montevechi (2016) and Chica, Juan Pérez, Cordon, and Kelton (2017), the success of this approach in decision making is mainly due to the capability of Simulation models to incorporate uncertainties, and the real stochastic nature of the environment.

Thus, we developed a Simulation model for freight SND. Then, we coupled this Simulation model, based on the Arena language with optimizer, Optquest, in order to solve the SND problem. We first validated this SBO-model on a recently addressed real-world case study of deterministic freight transportation problem combining roadway, railway and waterway (Demir et al., 2016). It was shown that we can reach the optimal service schedule. We also investigated the capability of our SBO-model to solve the Stochastic SND problem. We went beyond simplistic uncertainties, generally modeled as levels of disturbances, by considering continuously distributed demand and travel times. It was found that our SBO-model can deal with complex stochasticity and provide a reliable solution. We were able to reach an On Time and In Full delivery (OTIF) of more than 90%. The OTIF is a key indicator for logistics systems expressed as the percentage of orders arriving on time and in the right quantity (Rushton, Croucher, & Baker, 2014). Finally, we used an extension of the Value of the Stochastic Solution (VSS) indicator to measure the sensitivity of the output performance to the distribution type of stochastic processes. The VSS is a commonly used measure, in stochastic programming, to calculate the expected loss from using the deterministic solution rather than its stochastic counterpart. This research is among the first to analyze the sensitivity to different distribution types with specific shapes as symmetry, skewness, bimodality and uniformity. We found that the performance is not only sensitive to the variance as shown in former works (Bai, Wallace, Li, & Chong, 2014; Wang, Crainic, & Wallace, 2014), but also to the shape of data distribution.

This paper is organized as follows: Literature review is discussed in Section 2. In Section 3 adequacy of SND problem to solve freight transportation tactical planning is described. The Real-world case study is also presented. Section 4 details the implementation of the SBO model for the SND problem. Section 5 presents computational results and analysis. Finally, conclusions are discussed in Section 6.

2. Literature review

Over the last decades, many works addressed the Service Network Design Problem (SND) to solve the tactical planning issue of services selection, and scheduling in multimodal freight transportation. In early studies, only deterministic SND formulations were proposed in which demand and travel times are considered as deterministic Inputs (Crainic & Rousseau, 1986; Crainic, 2000). More recently, with the need for more reliable planning, a special emphasis is made to address Stochastic Service Network Design in freight transportation. As stated in SteadieSeifi et al. (2014) and Sun, Lang, and Wang (2015), assuming the planning elements such as demand and travel time as deterministic might result in the complete failure of the computed multimodal transportation planning.

To deal with stochastic demand, Lium et al. (2009) proposed a stochastic optimization model based on scenario tree. They used a discrete version of the triangular distribution to consider three different levels of uncertainty in demand; high, low, and no uncertainty. A Mixed-integer programming model was formulated with the objective of minimizing the expected cost over all scenarios. They showed the existence of structural differences between solutions reached under deterministic assumptions versus solutions when stochastic demand is considered. Meng et al. (2012) assumed that demand follows the Normal distribution and formulated the problem as a two-stage stochastic integer programming model. A solution algorithm, integrating the sample average approximation with a dual decomposition and Lagrangian relaxation is then proposed. Demand variability was tested through different scenarios based on random sampling from the Normal distribution. They found that demand variability has a significant effect on the solution. Bai et al. (2014) assumed also that demand uncertainty follows a discrete version of the triangular distribution and extend the model of Lium et al. (2009) by introducing rerouting options to face the demand uncertainty. Wang et al. (2014) and Bai et al. (2014) recourse to the Value of the Stochastic Solution, (VSS), measure to calculate the expected loss from using the solution generated from the deterministic model rather than its stochastic counterpart. They found that when the Coefficient of Variation, CV, is high, i.e. CV $\geqslant$ 30%, the VSS could be high which indicates a bad deterministic solution. Recall that, a CV is the ratio of the standard deviation to the mean. It is a statistical measure of the dispersion of data around the mean (Krishnamoorthy, 2016). Zhang, Li, Huang, Li, and Qian (2015) used robust optimization to address the network design problem. Their objective was to find the worst-case scenario when the uncertain demand is predefined in an interval. The results show that robust optimization can reduce worst-case cost. It is also found that the variability of the uncertain demand has an effect on the solutions.

Even though the assumption of deterministic travel time was criticized, very few works addressed travel time variability (e.g. (Bai et al., 2014; SteadieSeifi et al., 2014)). Andersen, Crainic, and Christiansen (2009) propose to add a slack variable and they formulate the problem as a mixed integer programming model with a nonlinear objective function. They add a penalty cost in their objective function to consider variability. Chávez et al. (2017) consider the inspection time variability at the border ports of entry. A triangular distribution is used to model this stochastic continuous time disruption. In their work, only road transportation mode is considered and travel time between nodes is assumed deterministic. The main finding was that ignoring the continuously distributed inspection time lead to inefficient routing plan. To the best of our knowledge, Demir et al. (2016) is the only work which addressed both demand and travel time uncertainties for solving SND problem in multimodal freight transportation. However, demand and travel times uncertainties are also assumed to follow a simple three-point discrete distribution. They formulate the problem as a stochastic linear mixed integer programming model. A real-world case study including road, rail and inland waterway services is solved using the sample average approximation method. They showed that the solution of the optimal schedule is more sensitive to travel time uncertainty than to demand uncertainty.

From this literature review, we can see that considering the variability is essential for more efficient tactical planning in multimodal freight transportation. Besides, according to Mönch, Lendermann, McGinnis, and Schirrmann (2011), SteadieSeifi et al. (2014), Chávez et al. (2017), Lee and Song (2017), for more reliable planning in freight transportation problems a challenging issue is to take into account real stochasticity, i.e. found in empirical data. Such empirical data analysis for travel times modelling have widely been addressed by researchers in the traffic and transportation field to develop reliability metrics. Rakha et al. (2010), Taylor and Somenahalli (2010), Guessous et al. (2014), Chalumuri and Yasuo (2014) have shown that roadway travel times data are skewed and possess long upper tail. Additionally, bimodality is often present. The first mode represents driving under free flow conditions, i.e. when there is no delay and congestion, the second mode when the traveler encounter long delays and queues. Through goodness of fit tests the Burr, Lognormal and Gamma distributions were proposed to model travel time in case there is a single mode but the Mixture Gamma distribution is needed to efficiently fit bimodal road travel time. They have also shown that other types, of more common distribution, such as Exponential and Normal distribution fail to capture the travel time stochastic behavior. Also, for railway freight travel times, similar data analysis were conducted on databases from different rail networks (Krüger et al., 2013; Wen et al., 2017). Studies have shown that railways travel times are skewed and exhibit fat tail. This fat tail is due to delays caused by the propagation of disturbance through the rail network. The Gamma, Lognormal and the Weibull distributions gave a good fit to model the train travel time. For vessel's travel time, works by Harrison and Fichtinger (2013) and Lee et al. (2015) indicate that the transit time of waterway freight is highly variable. They have shown that due to port congestion, the travel time presents bimodality. Though, the Mixture Lognormal distribution is found to better fit such empirical data.

Thus, in order to consider the continuously distributed demand and travel time we propose to address the problem with the simulation-based optimization modeling approach. As stated in Chica et al. (2017), unlike analytical methods, simulation modelling is a powerful tool for considering complex stochastic behavior. To the best of our knowledge, no former works have proposed to solve this class of SND problem using solely the simulation-based optimization approach. Even though some recent works (e.g. Chávez et al., 2017; Hrušovský et al., 2016) are aware of the importance to use simulation, they still model the problem as a mixed integer program and the simulation is only used to test the solution found under disturbed environment. Actually, as well explained in Juan, Faulin, Grasman, Rabe, and Figueira (2015), this is the hybrid simulation analytic approach and not the simulation-based optimization one. We elucidate this issue in Section 4.2. Besides, in former works very little, if any, attention was given to sensitivity of the solution to the shape of data's distribution: skewness, multimodality, uniformity. Here, we address this issue by computing the VSS for different types of travel time distributions.

3. Problem description and case study

The following first section presents the adequacy of SND problem with freight transportation tactical planning problem. The second section describes the addressed Real-world case study.

3.1. SND model for freight transportation

Using SND models to solve tactical freight transportation problems was firstly proposed by Crainic and Rousseau (1986). Since their pioneer work, Service Network Design is widely used by the operations research community to solve tactical issues of selecting and scheduling of services for freight transportation (e.g. Crainic, 2000; Sun et al., 2015). Usually this problem is modeled as a network composed by a set of nodes V and a set of edges E.

Each node i, $i \in V$, presents a terminal. A terminal could model, one or more at a time, an origin, a destination or a transshipment point such as a station, a port, an inland waterway port, a railway port, an airport etc. To each terminal i, $i \in V$, several characteristics could be associated such as transshipment operations' cost, the time needed for transshipment or for loading and unloading operation, etc.

Let S denote a set of services. Each service/network link s, $s \in S$, is *unique* connecting an origin terminal i_s, $i_s \in V$, to a destination terminal j_s, $j_s \in V$. Each service could be characterized by a time window availability $[W_s^{\min}, W_s^{\max}]$, a corresponding capacity u_s and a travel time t_s. Multiple transport modes could be considered (such as roads, rails, inland waterways, planes, etc.) while taking into account their corresponding characteristics. Thus, it is possible to have multiple services or

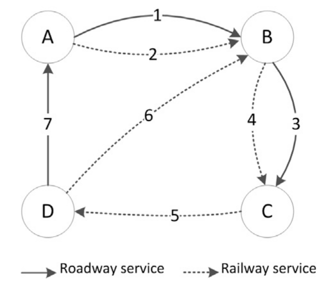

Fig. 1. Graph representation of the illustrative example.

"parallel arcs" linking the same pair of terminals while being different in term of time availability, or cost etc. In fact, each service s, $s \in S$, could be included in transportation plans.

As commonly defined, K denotes orders to be delivered through the network. Each order could be considered as a commodity with index k, $k = 1...K$, mainly characterized by a demand's amount d_k coming out from an origin/source terminal to a destination/sink terminal. Depending on the addressed problem, several other characteristics could be taken into account such as the earliest release time within the transportation might begin, and a due time. It is used to consider the release time as a hard constraint to respect because obviously, it corresponds to the time when the goods are ready for transportation. Moreover, a per unit penalty cost could be assigned for each late delivered order which exceeded the due time.

Usually the aim of the problem is to find optimal service schedule, referred to as the Path in the remainder of this paper, while considering time windows restriction, flow conservation principle and capacity constraints, etc. For sake of clearness, we consider an illustrative example detailed in Fig. 1, with 4 terminals, 7 services and 2 orders to deliver without considering any release times. Table 1 depicts the characteristics of each service: capacity, time windows and travel time. The orders characteristics are given in Table 2. No transshipment times are considered in all the terminals.

We notice that for the first order ($k = 1$), even if the path (2, 4): using services n°2 then n°4, is feasible in term of capacity, it is not chosen because the corresponding total travel time is equal to $22 = 8 + 14 > 20$ which is the due time of the order.

For the second order ($k = 2$), the service s n°2 was not selected in the optimal solution, because $t_7 \notin [W_2^{\min}, W_2^{\max}]$. Once service 7 is chosen, the transported goods will arrive at $t_7 = 20$ and fit only into the time window [0, 50] of service s n°1, which is simultaneously used to route the two orders. Even though taking the service s n°6 is a feasible path, it is not chosen because the objective is to minimize the total travel time. With service s n°6 travel time is 30 which is greater than $25 = t_7 + t_1$: the travel time of the proposed solution (7, 1).

Actually, the objective is to find the optimal Path that minimizes a transportation cost. Multi-objective optimization was also used to deal with conflicting objectives such as cost and time. Actually, minimizing the orders delivery time leads naturally to selecting the faster means of transport which are not necessarily the cheapest ones. Recent works (e.g. Baykasoğlu & Subulan, 2016; Demir et al., 2016) also consider the

Table 1
Illustrative example of SND problem: services characteristics.

Service N°s	(i_s, j_s)	Capacity u_s	$[W_s^{min}, W_s^{max}]$	Travel Time t_s
1	(A,B)	16	[0, 50]	5
2	(A,B)	13	[0, 15]	8
3	(B,C)	8	[0, 50]	10
4	(B,C)	15	[5, 10]	14
5	(C,D)	10	[10, 20]	5
6	(D,B)	10	[0, 10]	30
7	(D,A)	15	[0, 20]	20

Table 2
Illustrative example of SND problem: orders characteristics and solutions.

Order N°k	(origin→ destination)	Demand d_k	Due Time	Solution/Path (Sequence of services)
1	(A→C)	10	20	1,4
2	(D→B)	8	30	7,1

greenhouse gas emission. For that purpose, several multi-objective optimization methods were adopted such as goal programming (e.g. Yang, Low, & Tang, 2011), and weighted-sum method (e.g. Demir et al., 2016), which we also used in our work.

3.2. Real case study for freight transportation

The considered real world case study is a multimodal freight transportation problem in the Danube region as depicted in Fig. 2. This case was recently introduced by Demir et al. (2016).

We model this case study as a network having $|V| = 10$ nodes or terminals and $|S| = 32$ directed links or services as shown in Fig. 3. There are 3 service types: roadway, railway and water services. Roadway services are represented with solid arrows. Waterway services are represented with small dashed arrows and railway services are represented with large dashed arrows. Each service is labeled by its number and is characterized by a unit transportation set up cost as detailed in Appendix Table A.2.

In this real-world case, CO_2 emissions, via emissions costs, are taken into account. To each service s, $s \in S$, and terminal i, $i \in V$, is associated a prefixed emission ε_s, ε_i, respectively. Table A.1 details terminal characterization per each container transshipped. In such cases, $W_s^{min} = 0$ and W_s^{max}= total frame time fixed to a week (168 h) Demir et al., 2016. There are $K = 5$ commodities (d_1, …, d_5) to be transported from 3 different departure locations: Budapest Port, Prague and Vienna Port represented respectively by nodes A, E and J in Fig. 3. Demand settings could be found in Appendix Table A.3.

Let $\mathscr{S}$ be a feasible solution for the problem. Thus, the objective is to find the selection and the scheduling of services in order to form a Path with a minimum global cost $\Psi(\mathscr{S})$. The multiple objectives of this problem is to simultaneously minimize: (i) Transport costs: direct services costs, (ii) Time, via penalty costs for delay (iii) CO_2 emissions, via emissions costs: services emission. A weighted-sum method is used to derive a single objective function Ψ, by associating to each objective (i)–(iii) a non-negative weight w_τ, ω_γ and ω_ε respectively, as follows:

$$\Psi(\mathscr{S}): \min_{\mathscr{S}\in\Pi} \omega_\tau \Psi_\tau(\mathscr{S}) + \omega_\gamma \Psi_\gamma(\mathscr{S}) + \omega_\varepsilon \Psi_\varepsilon(\mathscr{S}), \tag{1}$$

where Π is the set of feasible solutions.

To solve this freight transportation planning problem, a Mixed Integer Programming formulation was proposed by Demir et al. (2016). They solved the deterministic version of this problem using the general MIP solver CPLEX. In our work, we will compare our results to theirs, in order to validate the deterministic version of our Simulation-Based Optimization model.

4. The simulation based optimization model

The purpose of the following first section is to elucidate some simulation modelling issues related to the developed freight SND simulation model. The second section deals with the implementation of the simulation based optimization approach.

4.1. The freight SND simulation model

In this work, the SND simulation model is built with the popular Arena Software. This choice is motivated by the fact that, according to the recent review (Oliveira et al., 2016), Arena is the most used software in logistics and supply chain simulation. For more insight, Arena interested reader can consult the reference (Kelton, Sadowski, & Sturrock, 2007). Also, the Arena model (i.e., .doe file), is available upon request.

The objective of the developed simulation model is to reproduce the processing of multiple commodities, i.e. orders, through the freight Network. Two main components were implemented with the Arena module: the Network's physical structure, the logic of freight orders flowing through the Network.

For the Network's structure, we use the STATION module to define each node i, $i \in V$, representing a terminal. The VARIABLE module is used to implement characteristics associated with each terminal i, such as transshipment operations cost. For that purpose, we define the variable *vTerminals* as a two-dimensional array with the number of rows

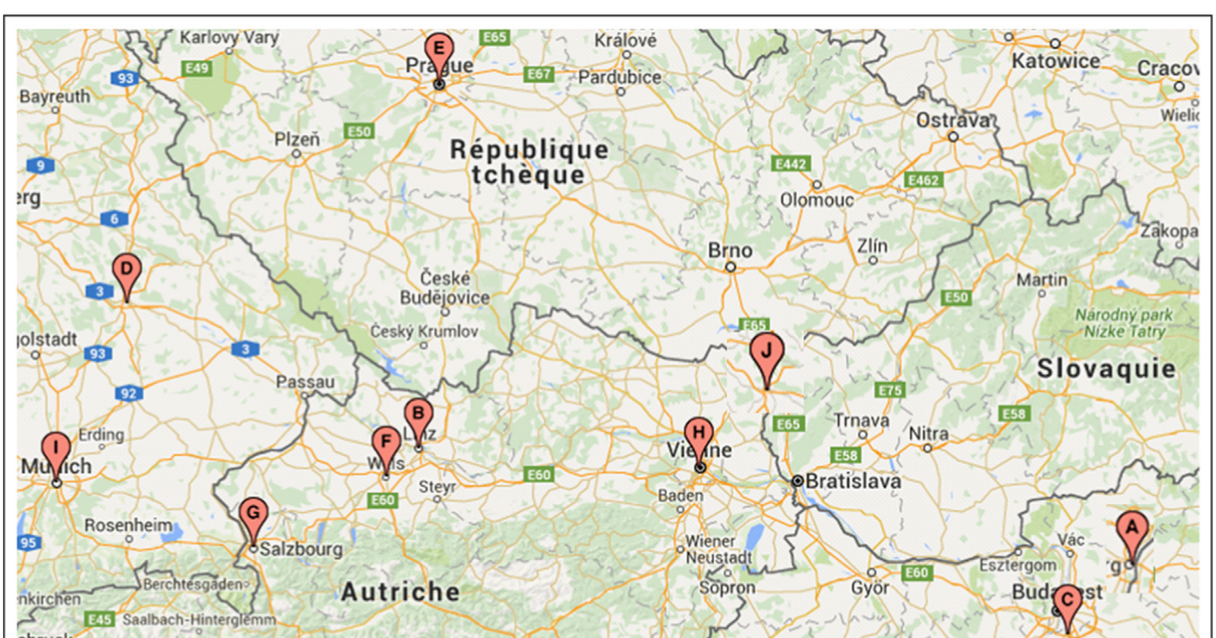

Fig. 2. Overview of Terminals location in the real-world case study.

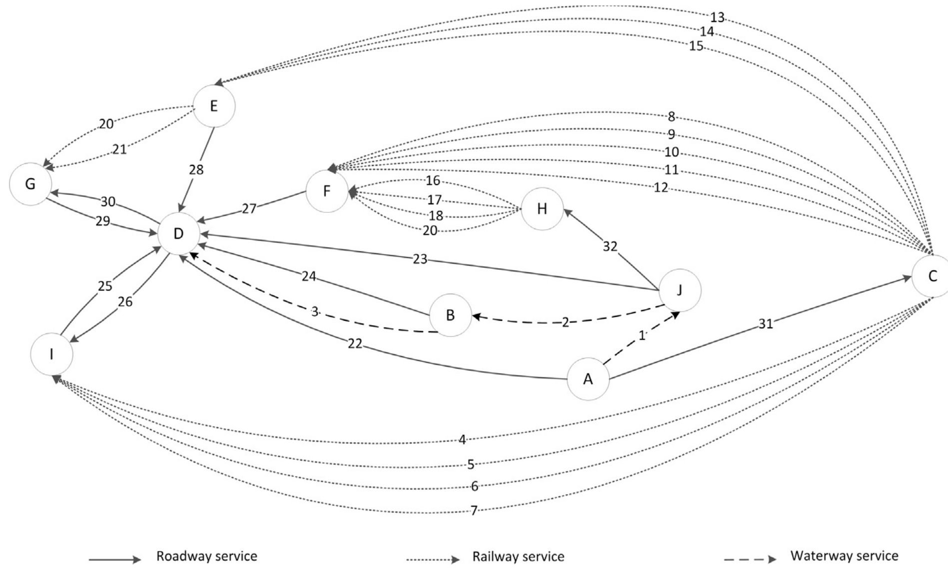

Fig. 3. Graph representation of the case study.

equal to $|V| = 10$ and columns correspond to characteristics from Appendix Table A.1. After terminal definition, the VARIABLE module is also used to define services/network links s. The variable *vServices* is a two-dimensional array with the number of rows equal to $|S| = 32$ and columns correspond to characteristics given in Appendix Table A.2. Finally, a variable *vLink*, a two-dimensional array (32×2), is used to define the connection (i_s, j_s) between an origin terminal i_s to a destination terminal j_s ; according to data given in Appendix Table A.2.

Once the Network's structure is defined the next step was to implement the freight orders flowing through this Network. At first, the CREATE module generates 1 entity to initiate the model. Then, this entity goes through a SEPARATE module to be duplicated and create the $|K| = 5$ different orders. When an entity leaves this SEPARATE module, it represents an order in the model. An ASSIGN module gives to each order the corresponding characteristics taken from variable *vCommodities*. The variable *vCommodities* is a two-dimensional array with the number of rows equal to $|K| = 5$ and columns correspond to characteristics given in Appendix Table A.3.

An order is routed through the different nodes of the network according to a path sequence given in the ROUTE module when leaving a STATION. This path is the freight service schedule, i.e. the decision variable of our SND model. For this purpose, we introduced a variable *vPath* which is a two-dimensional array with number of rows equal to the number of orders $|K| = 5$ and columns correspond to the sequence of Services to be used. When an entity goes through a service; emission, transportation cost and the remaining capacity are computed to avoid overflow on the given service. The travel time in the ROUTE module is specified by the variable *vServices*. Once the K orders reached their destination; output performances such as transportation costs, carbon dioxide emissions and the penalty for late order deliveries are computed. It is worth to mention that cardinality of V, K and S are only given to illustrate the implementation of the case study. However, the model can reproduce any different/larger Network instances. For example in Section 5.3, new instances are defined with bigger V, K and S values.

It is important to notice that this SND simulation model does not explicitly include any resources. In other words, transport means such as trucks train and boats are not defined with the RESOURCE or the TRANSPORTER modules. This modelling logic is used in order to reproduce the same level of abstraction used on the freight SND problem formulation.

In order to implement the stochastic SND model, we had to consider demand and travel time, respectively D_k where $k \in \{1, 2, ..., 5\}$ and T_s where $s \in \{1, 2, ..., 32\}$, as random variables following specific probability distribution functions. Notice that, the deterministic SND model is a particular case of the Stochastic SND model with $Var[D_k] = 0, \forall\, k \in \{1, 2, ..., |K|\}$, and $Var[T_s] = 0, \forall\, s \in \{1, 2, ..., |S|\}$.

For the Stochastic SND model, we assume that demand D_k; for each commodity k where $k \in \{1, 2, ..., 5\}$; follow a Lognormal distribution, as proposed in Juan et al. (2014). The deterministic demand d_k, Appendix Table A.3, is then converted into probabilistic demand D_k with $E[D_k] = d_k$. To introduce a low level of variance, we can choose a Coefficient of Variation (CV) of 10% or 20%, with $Var[D_k] = (CV \times E[D_k])^2$. Then, we apply the method of moments to estimate the location parameter, μ_k, and the scale parameter σ_k, of each lognormally distributed demand D_k, according to Eqs. (2)–(4).

$$D_k \sim Lognormal(\mu_k, \sigma_k), \tag{2}$$

$$\mu_k = \ln(E[D_k]) - \frac{1}{2}\ln\left(1 + \frac{Var[D_k]}{(E[D_k])^2}\right), \tag{3}$$

$$\sigma_k = \left|\sqrt{\left(1 + \ln\left(\frac{Var[D_k]}{(E[D_k])^2}\right)\right)}\right|. \tag{4}$$

For example, commodity 1 with the deterministic demand $d_1 = 20$, is converted to probabilistic demand $D_1 \sim Lognormal(2.99, 0.09)$; the corresponding plot is given in Fig. 4(a) below.

For each service s, we used the same method of moments to determine the parameters of the corresponding travel time distribution, T_s for $s \in \{1, 2, ..., 32\}$. For ship/Waterway services, $s \in \{1, 2, 3\}$, we assume that transportation time follow Mixture/bimodal distribution with two modes, as proposed in Harrison and Fichtinger (2013) and Lee et al. (2015). The first mode represents travel under uncongested conditions, and the second mode represents travel under congested conditions. For this mixture distribution, we fixed that the first mode is

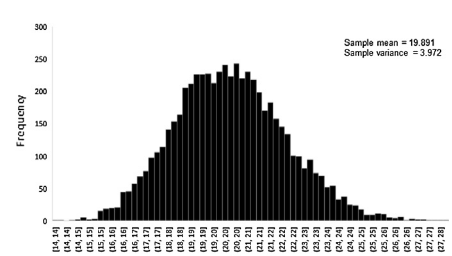

Fig. 4. (a). Histogram of Commodity 1 demand. (b). Histogram of Service 22 Travel time.

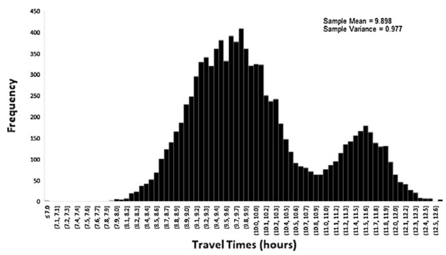

Fig. 4. (*continued*)

encountered with probability $p_1 = 80\%$ versus a probability of $p_2 = 20\%$ for the second mode. The deterministic travel time t_s, Appendix Table A.2, is then converted into the probabilistic travel time T_s, with

$$T_s \sim p_1 Lognormal(\mu_{1,s}, \sigma_{1,s}) + p_2 Lognormal(\mu_{2,s}, \sigma_{2,s}). \tag{5}$$

For Train/Railway services, $s \in \{4, 5, ..., 21\}$, we assume that transportation time T_s follow Lognormal distribution, as proposed in Krüger et al. (2013). We convert the fixed travel time t_s for $s \in \{4, 5, ..., 21\}$, from Appendix Table A.2, into probabilistic travel time T_s, with $T_s \sim Lognormal(\mu_s, \sigma_s)$. As for the Lognormally distributed demand, we fixed $E[T_s] = t_s$, and $Var[T_k] = (CV \times E[T_k])^2$; then we use the method of moment to determine μ_s and σ_s for $s \in \{4, 5, ..., 21\}$.

Finally for Truck/Roadway services, $s \in \{22, 23, ..., 32\}$, we assume that travel time follow Mixture/bimodal Gamma distribution, as proposed in Rakha et al. (2010), Taylor and Somenahalli (2010), and Guessous et al. (2014). The first mode represents travel under free flow; the second mode represents travel under congested flow. Under this second mode, the average travel time is increased by 20%. For this Mixture Gamma distribution, it is assumed that the first mode is encountered with probability $p_1 = 80\%$ versus a probability of $p_2 = 20\%$ for the second mode. Deterministic travel times t_s, given in Appendix Table A.2, are then converted into the probabilistic travel time T_s, with

$$T_s \sim p_1 Gamma(\alpha_{1,s}, \beta_{1,s}) + p_2 Gamma(\alpha_{2,s}, \beta_{2,s}), \ \ \forall\, s \in \{22, 23, ..., 32\}, \tag{6}$$

where the shape parameter ($\alpha_{i,s}$) and the scale parameter ($\beta_{i,s}$), $i = 1, 2$ are estimated using the method of moments. These parameters are expressed by Eqs. (7) and (8) stated as follows.

$$\alpha_{i,s} = \frac{E[T_{i,s}]^2}{Var[T_{i,s}]}, \ \ i = 1, 2, \ \ s \in \left\{22, 23, ..., 32\right\}, \tag{7}$$

$$\beta_{i,s} = \frac{E[T_{i,s}]}{Var[T_{i,s}]}, \ \ i = 1, 2, \ \ s \in \left\{22, 23, ..., 32\right\}. \tag{8}$$

For example, service 22 with the deterministic travel time $t_{22} = 10$ *hours*, is converted to probabilistic travel time T_{22}; such as

$$T_{22} \sim 0.8 Gamma(204, 0.047) + 0.2 Gamma(1111, 0.01), \tag{9}$$

with $E[T_{22}] = t_{22} = 10$. The corresponding plot is given in Fig. 4(b).

4.2. Coupling simulation with optimization

Once the simulation model is developed, the next step consists of coupling this model with an optimization technique in order to find an optimal freight Path. The technique of combining simulation and optimization methods, also called simulation-based optimization has been successfully applied to solve complex decision making problems (Fu, 2002). The basic concept of this method (Fig. 5) is to use an optimization module (generally based on a metaheuristic algorithm) to find a set of values for the input parameters (i.e., decision variables), and uses the responses generated by the simulation model output (i.e., performance indicators) to search for the next trial solution. This iterative search aims at finding a set of input values with the highest contribution to the performance indicators (i.e., best optimal/suboptimal Solution).

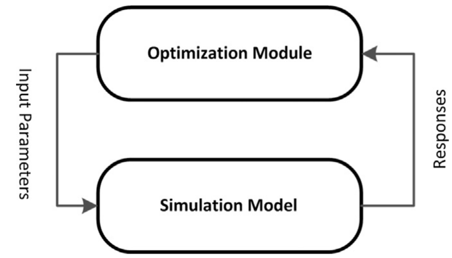

Fig. 5. Simulation-based Optimization approach (Fu, 2002).

In the last two decades, research papers have reported benefits of this simulation-optimization process to solve problems with stochastic components (Jaoua, Gamache, & Riopel, 2012; Juan et al., 2015; Oliveira et al., 2016; Yegul, Erenay, Striepe, & Yavuz, 2017). Also, commercial software proposed powerful optimizer modules, such as Optquest (provided by OptTek System Inc.), to deal with complex stochastic optimization. The Search Algorithm used in Optquest is based on a combination of the metaheuristics of tabu search, neural networks, and scatter search. Furthermore, Optquest handle multi-objective optimization using the weighted sum method (Glover, Laguna, & Marti, 2003).

Even though research papers have reported the success of the Simulation based Optimization, its use in practice remains scarce (Halim & Seck, 2011). Among the reasons, Halim and Seck (2011) and Jaoua et al. (2012) criticize the simple conceptualization like depicted in Fig. 5, or even the one based on flowcharts (Chávez et al., 2017), which do not provide a formal structure of the looping and conditional behavior characterizing the simulation based optimization process. They state that there is a need to use specification language, such as the UML standard, for rigorous computer implementation, in practice.

Therefore, we developed an UML sequence diagram, Fig. 6, to formally specify the modules conditional behavior and the temporal sequence involved in the simulation-optimization interaction process. For the behavior modeling, the branching improved concept provided in UML2.0® is used. These concepts are well documented in the work of Rumbaugh, Jacobson, and Booch (2004).

Four objects are involved in the interaction: *User*, *Optimizer*, *Simulator* and *RNG* (for *R*andom *N*umber *G*enerator). To these Objects, we associate the following instances: *Optquest*, *SNDModel. p* the executable '.p' file from Arena software and the *CMRG* (for *C*ombined *M* ultiple *R*ecursive *G*enerator) (L'ecuyer, Simard, Chen, & Kelton, 2002). Actually, the proposed UML sequence diagram is not restricted to the use of these instances, i.e. other metaheuristics algorithms and Random Number Generator as well as other simulation language rather than Arena, could be used.

Fundamentally, the role of the *Optimizer* is to send a path to the *Simulator*. The *Simulator* tests this *path* and gives the corresponding *Cost* to the *Optimizer*. Then, the *Optimizer* applies the search Algorithm to find a better path to be tested by the *Simulator*. The diagram in Fig. 6, presents the interactions between objects in the sequential order.

At first, *User* sends a request to start the search for an optimal path, i.e. services schedule; specifies whether the deterministic or the stochastic version of the freight SND problem is to be resolved. An optimization loop fragment is then created to iterate through an unknown number of configurations, i.e. path. The guard condition indicates that this interaction for path generation and simulation is repeated until *TimeUp* or An Automatic Stop condition is met. These two stopping rules are given in *Optquest*. One can choose to run the process for a fixed time period, for example one hour, or to stop it when there is no improvement in the value of the objective function for a fixed number of consecutive tested configurations.

At this point, an alternative, *Alt*, fragment is created to consider the two different optimization executions: for the deterministic and for the stochastic model. Recall that, a null Variance corresponds to the deterministic SND problem. This condition is expressed on the guard of

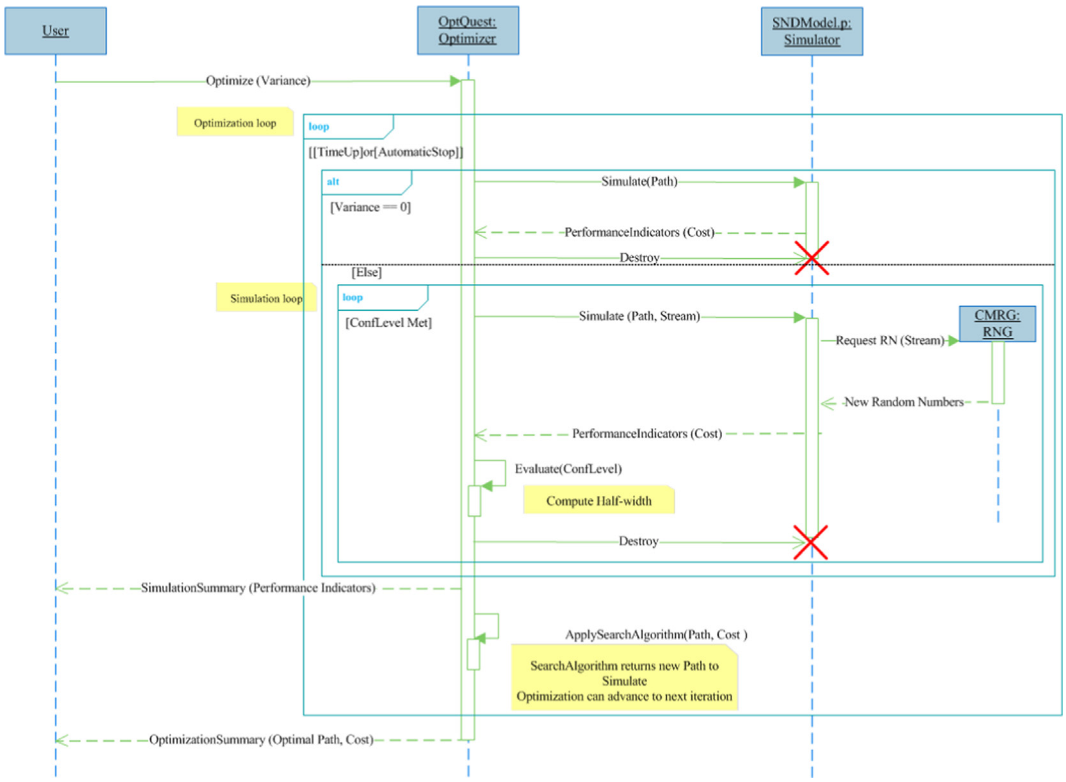

Fig. 6. UML Sequence diagram for the Simulation-based Optimization interaction.

the *Alt* operator. If the deterministic model is concerned, the simulation is executed only one time through the invocation of *Simulate* (*Path*).When the execution is terminated, the *Simulator* sends the output performance measures to the *Optimizer* through the invocation of *Performance Indicators* (*Cost*).Then, the *Optimizer* destroys the current instance of the simulator; because this instance was created to calculate the Output of the specific Path.

The second section of the *Alt* fragment shows the different interactions related to the stochastic model. Fundamentally, another iteration loop is started. In this loop fragment, the simulation is executed many times, i.e. replications, through the invocation of *Simulate* (*Path*, *Stream*) until the confidence level is met. The number of replications depends on the resulted variance from each configuration. Through invocation of *Evaluate* (*ConfLevel*), the required number of replications is computed based on the Half-width formula given in Kelton et al. (2007). For each replication, new random numbers are needed. These numbers are generated using different streams from *CMRG* which is instantiated during the construction of *Simulator*. An important issue is to specify the parameter *StreamtoSimulate*(). Using the same random number stream for each stochastic variable is known as the Common Random Number (CRN) technique of variance reduction. It allows reducing the variance of the difference between sample means and increasing the accuracy of the comparison. According to Fu, Hu, Chen, and Xiong (2007), it is important to apply such method to accelerate the simulation-based optimization process.

Finally, when the *Alt* fragment is completed, the *Optimizer* has the path and the corresponding *Cost*. These values are used when *ApplySearchAlgorithm* (*Path*, *Cost*) is invoked. A new Path is found then a new iteration can begin.

5. Computational results and analysis

For the computational analysis, we first apply our proposed SBO-Model to the deterministic version of the freight SND problem. Once the model is validated, we use it on a stochastic version of the problem. The last part aims to investigate the sensitivity of the performance indicator to different types of stochasticity. Specifically, we analyze the effect of skewness and multimodality that are frequently found in travel time empirical data.

5.1. Results for the deterministic freight SND problem

The purpose of this section is to validate the simulation model. Actually, in simulation modelling validation is a fundamental step before experimentation. We choose to realize this phase by comparing our results to the one found analytically in Demir et al. (2016), for the deterministic version of the multimodal freight transportation problem presented in Section 3.2.

Let M_0 denote the deterministic SND simulation model. Also, we denote $SBO\text{–}M_0$ as the application of the Simulation-Based Optimization approach to find the optimal path for M_0.We apply $SBO\text{–}M_0$ to solve the 9 cases proposed by Demir et al. (2016). For each case, different weights are assigned to the transportation costs (w_t), the time penalty costs (ω_γ) and to the emission-related costs (ω_ε). Obtained results are reported in Table 3.

The optimal Paths presented in Table 3 are similar to the results obtained by Demir et al. (2016) when they used analytical formulation of the freight SND problem. That means that our $SBO\text{–}M_0$ can efficiently tackle this NP-hard optimization problem. Even if the computational time seems high compared to the analytical formulation (some seconds), our approach remains competitive as such planning decision is not taken in real time. Actually, when using Simulation, we have to be aware that it is not a competitive faster method.

We observe in Table 3 that optimal solutions/paths are strongly correlated to the weights of each objective. For example, in cases 1 and 6 where the weight of transportation service costs is sensibly high, railway and waterway services are selected, because they are less costly. Whereas in case 2 and 5, when the weight of time penalty cost in increased, roadways services are mostly selected since they are faster. Another example of reducing emission related cost is achieved in case

Table 3
Computational results of $SBO{-}M_0$.

Case	Weights			Optimal Paths (sequence of services)					Total cost	*CPU Time*
	ω_τ	ω_γ	ω_ε	1	2	3	4	5	(€)	
1	1	0	0	1,2,3	1,2,3	31,5	2,3	21	24 681	1 h10 mn
2	0	1	0	22	22	22,26	23	28,3	33 919	2 h12 mn
3	0	0	1	31,5,25	31,6,25	31,7	2,3	21	35 229	10 mn
4	0.4	0.4	0.2	1,2,3	1,2,3	31,5	2,3	28,3	23 285	2 h13 mn
5	0.2	0.6	0.2	1,2,3	1,2,3	22,26	2,3	28,3	26 010	1 h25 mn
6	0.6	0.3	0.1	1,2,3	1,2,3	3 1,5	2,3	2 1	24 681	1 h50 mn
7	0.1	0.8	0.1	1,2,3	1,2,3	22,26	2,3	28,3	26 010	1 h45 mn
8	1	1	1	1,2,3	1,2,3	31,5	2,3	28,3	23 285	2 h30 mn
9	1	10	10	1,2,3	1,2,3	31,8,27,26	2,3	28,3	26 179	1 h20 mn

3.In this case the first commodity corresponds to a transport request of 20 orders from Budapest Port (A) to Regensburg (D). As shown in Fig. 3 many Paths could ensure this request. The optimal Path proposed in case 3 is the sequence of services (31, 5, 25): 31 and 25 are roadway services and 5 is a railway service. This solution is preferred to the Path (1, 2, 3) proposed in case 1 because it reduces the CO2 emission.

From this section, we can conclude that our SBO-model is valid. Thus, we can use it to solve complex stochastic versions of the freight transportation problem.

5.2. Results for the stochastic freight SND problem

The first objective herein is to investigate the capability of the Simulation-Based Optimization approach to solve the Stochastic SND problem while considering realistic continuous distributions. Let's M_1 denotes the stochastic freight SND simulation model where demands and travel times are random variables generated according to the corresponding distributions from the literature. Namely, the lognormally distributed demand (Garrido & Mahmassani, 2000), also lognormally distributed railway travel times (Krüger et al., 2013), the Mixture (bimodal) Lognormal distribution for waterway travel time variability, as proposed in Harrison and Fichtinger (2013), Lee et al. (2015); and the Mixture (bimodal) Gamma distribution of travel time variability on roads, as found in Rakha et al. (2010), Taylor and Somenahalli (2010), Guessous et al. (2014).

We denote by $SBO{-}M_1$ the application of the Simulation-Based Optimization approach to find the optimal path for the model M_1. We apply $SBO{-}M_1$ for the same addressed nine cases; we also compute the On Time and In Full delivery (OTIF) indicator. We opt for the OTIF as it is a performance indicator widely used in freight logistics (Rushton et al., 2014). Table 4 shows the computational results. For the Optimization Loop, we fixed the Time Up to 60 min. Since we are using stochastic simulation model, the OTIF output performance is estimated, $\widehat{OTIF}$, as the sample average over n independent replications. The corresponding half-widths of the 95% confidence intervals are computed.

As the results in Table 4 indicate, the $SBO{-}M_1$ finds paths that lead to a good $\widehat{OTIF}$ of more than 90%, in all tested cases. In fact, despite the realistic and complex stochasticity implemented in this freight SND model, this resolution approach can find a good solution in only 60 min. Recall that, as we use the CRN technique, this process of simulation-optimization is efficiently accelerated.

In addition, we observe that, in all cases, the $SBO{-}M_1$ proposes different optimal/suboptimal Paths from the ones given by the $SBO{-}M_0$ in Table 3. That means that the deterministic solution given by $SBO{-}M_0$ is not reliable enough to deal with the real stochasticity introduced in M_1. For example for the 1*st* case, the new proposed Path for commodity 4 is (2, 24) rather than (2, 3). In fact, congestion encountered with service 2 induces a high risk of performance deterioration due to the missing of service 3 time window restriction. Recall that travel time through service 2, T_2, in M_1 follows the bimodal Lognormal distribution. Then to avoid the risk of performance loss, $SBO{-}M_1$ uses service 24 because it did not have a tight time window.

However, we found for some commodities, that the path proposed by the deterministic solution is maintained even under stochastic environment, such as the 1*st* case paths for commodity 1 and 2. That means that, for some commodities, using the Path proposed by $SBO{-}M_0$ does not induce a loss of performance in stochastic environment. Then, we recourse to the Value of Stochastic Solution indicator (VSS) to measure to which extent the deterministic solution could be used in stochastic environment.

The VSS is a well-known standard measure in stochastic programming, already used in Bai et al. (2014) and Wang et al. (2014). A high VSS indicates bad deterministic solution; i.e unreliable services schedule to variations. In this work, we compute the VSS as the $\widehat{OTIF}$ difference given by Eq. (10). The VSS is equal to the $\widehat{OTIF}$ reached with the Path found by $SBO{-}M_1$ minus the $\widehat{OTIF}$ reached with the Path found by $SBO{-}M_0$ tested on the stochastic SND model M_1:

$$VSS = \widehat{OTIF}(Path_{SBO-M_1}) - \widehat{OTIF}(Path_{SBO-M_0}). \tag{10}$$

As we want to compute the difference between two stochastic simulation outputs, $\widehat{OTIF}$, we need advanced statistical comparison

Table 4
Computational results of $SBO{-}M_1$.

Case	Optimal Paths (sequence of services)					OTIF	Half Width	Number of
	1	2	3	4	5	(%)	(%)	Replications (n)
1	1,2,3	1,2,3	22,26	2,24	28,30	97.04%	1.30	310
2	22	22	22,26	2,24	28,30	98.03%	1.60	408
3	31,5,25	31,7,29	22,26	2,24	28,30	92.05%	0.50	358
4	22	31,5,25	22,26	2,3	28,30	95.43%	1.70	352
5	1,2,3	31,5,25	22,26	2,3	28,30	93.83%	110	420
6	1,2,3	1,2,3	22,26	2,24	28,30	97.04%	1.30	310
7	1,2,3	31,5,25	22,26	2,3	28,30	93.83%	1.70	420
8	1,2,3	1,2,3	22,26	23	28,30	94.13%	0.21	350
9	31,5,25	22	22,26	23	28,30	91.20%	1.23	456

Table 5
Average loss in the stochastic environment.

	Simulation of deterministic solution			SBO-model solution			VSS	
Case	OTIF (%)	n	std	OTIF (%)	n	std	Mean (%)	C.I (%)
1	42.05	303	0.212	97.04	310	0.116	54.99	2.72
2	89.32	423	0.021	98.03	408	0.164	8.71	1.61
3	20.45	340	0.141	92.05	358	0.048	71.60	1.58
4	65.98	486	0.035	95.43	352	0.162	29.45	1.72
5	78.45	316	0.126	93.83	420	0.115	15.38	1.77
6	42.05	303	0.212	97.04	310	0.116	54.99	2.72
7	78.45	316	0.126	93.83	420	0.115	15.38	1.77
8	65.23	586	0.049	94.13	350	0.162	28.90	1.74
9	71.90	989	0.256	91.20	456	0.023	19.30	1.61

techniques. Thus, we calculate the VSS and the corresponding confidence interval using the Welch test at a confidence level of 95% (Krishnamoorthy, 2016). Results are presented in Table 5. The number n of replications is determined such that, at a 95% confidence level, the half-width of the confidence interval is less than 0.05.

As shown in Table 5, for all tested cases the confidence interval on the VSS difference does not cover 0. Therefore, we can conclude that this difference is statistically significant at an α level of 0.05, i.e. the OTIF output values are statistically different. We also see that the average loss from using the deterministic Path proposed by $SBO{-}M_0$ in stochastic environment M_1 can variate between 8.71% and 71.6% These average losses vary from a case to another dependently on the Path structure. This is consistent with the results of Wang et al. (2014) and Bai et al. (2014), where they show that freight demand uncertainty induces VSS variation between 5% and 55%. For example we see in Table 5 that for the 4*th* case using the Path found in Table 3, rather than the Path given in Table 4, in the real stochastic freight transportation network will deteriorate the OTIF by an average of 29.45% ± 1.72.

However, Bai et al. (2014) found that with low level on demand uncertainty there is no significant difference between the stochastic and the deterministic solution, i.e. the output is insensitive to the enrolled stochasticity. Even though in our experiments we assumed a low level of variance, we did found a considerably high VSS average. That means that the OTIF output performance is indeed sensitive to stochasticity even under low levels of variance. In fact, in former works stochasticity is reproduced with simplistic discrete distributions whereas in our work we use realistic continuous distributions. Furthermore, in the present work, we address both demand and travel time variabilities. In order to elucidate this issue, we conduct, in the next section, a sensitivity analysis to assess the effect of using different continuous distribution types on the OTIF performance measure.

5.3. Sensitivity of the OTIF to distribution types

In our next set of experiments, we focus solely on changing the roadway travel time distributions. We want to analyze the sensitivity of the OTIF to the following types of travel time distributions: Mixed Gamma, Uniform, Triangular, Normal and Lognormal distributions. The fundamental difference between these distributions exists in their shapes. In order to analyze the effect of the different shapes, we fixed the same mean and variance for all these travel time distributions. Our baseline model is M_1 where roadway travel time is distributed according to the Mixture/bimodal. Let's M_2, M_3, M_4 and M_5 denote Stochastic SND simulation model for freight transportation where the roadway travel times are respectively assumed to follow Uniform, Triangular, Normal and Lognormal distributions.

To analyze the sensitivity of the OTIF to the different distributions types, we introduce the VSS_{M_i} as an extension of the VSS. Generally, the VSS is used to measure the expected loss when using the deterministic solution. We introduce the VSS_{M_i} to measure the expected loss when

Table 6
Summary Statistics of VSS_{M_i}.

Statistics	VSS_{M_0}	VSS_{M_2}	VSS_{M_3}	VSS_{M_4}	VSS_{M_5}
Mean	27.45	23.92	19.60	19.22	11.13
95% confidence level	[26.45, 28.45]	[23.13, 24.70]	[18.75, 20.45]	[18.20, 20.24]	[10.20, 12.07]
SD	3.52	2.77	2.99	3.58	3.28
Minimum	20.48	18.74	13.68	13.09	5.64
Q1 (0.25 percentile)	24.99	22.24	17.64	16.35	8.76
Q2 (Median)	26.73	23.55	19.57	18.66	10.88
Q3 (0.75 percentile)	30.03	25.57	21.85	20.89	12.47
Maximum	35.99	31.41	25.51	27.57	19.92

considering another type of stochasticity rather than the real stochasticity, i.e. herein implemented in the baseline model M_1. VSS_{M_i} is calculated following Expression (11) below:

$$VSS_{M_i} = \widehat{OTIF}(Path_{SBO-M_1}) - \widehat{OTIF}(Path_{SBO-M_i}), \quad \forall\ i = 0, 2, \ldots, 5. \qquad (11)$$

In order to evaluate the statistical significance of the VSS_{M_i}, we created a set of 50 new instances. The basis for these additional instances is a network having $|V| = 20$ nodes, $|S| = 45$ services and a number of $|K| = 15$ commodity. The instances vary according to demand's amount and the corresponding origin/destination terminals, herein randomly generated. We ran experiments and computed the VSS_{M_i} as defined in (11).

Table 6 reports the summary statistics of the experiments. We made box-and-whisker plots as shown in Fig. 7 for the VSS_{M_i} result values. This plot shows numerical results beyond the means, such as the minimum, maximum of the data and a box that encapsulates fifty percent of the data, i.e. the interquartile range equal to (Q3-Q1), from Table 6. We used these box-whisker plots, Fig. 7, to compare the distribution of the responses across the Model M_i, $i = 0, 2, \ldots, 5$.

As shown in Fig. 7, all VSS_{M_i} values, including plotted outliers, are strictly positive. That means that the calculated difference is statistically significant at an alpha level of 0.05, hence the OTIF is indeed sensitive to the type of distribution. In order to determine which model M_i leads to the worst solution, i.e. less reliable Path, we computed in Table 7 the 95% Welch confidence intervals for all pairwise comparison. We used this advanced technique for results comparison because the central boxes overlap, as shown in Fig. 7.

From Table 7, we conclude that the difference is not significant only for the comparison between the Normal and the Triangular models. The corresponding 95% Welch confidence interval plot ($VSS_{M3}{-}VSS_{M_4}$) contains 0, as shown in Fig. 8. Then, we can deduce that the deterministic model M_0 leads to the worst solution which induces an average loss of $VSS_{M_0} = 27.45\%$. In fact, even the 25th percentile is equal to 24.99 which is a substantial performance loss. That means that the deterministic assumption may lead to such highly unreliable solution.

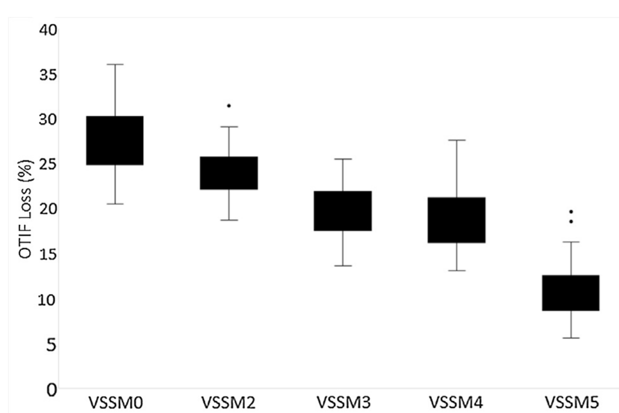

Fig. 7. Box-and-whisker Plots of VSS_{M_i}.

Table 7
95% Welch confidence intervals for the VSS_{M_i} comparison.

	VSS_{M_0}	VSS_{M_2}	VSS_{M_3}	VSS_{M_4}	VSS_{M_5}
VSS_{M_0}	–	[−4.78, −2.29]	[−9.13, −6.56]	[−9.63, −6.83]	[−17.66, −14.98]
VSS_{M_2}	–	–	[−5.45, −3.18]	[−5.95, −3.43]	[−13.98, −11.59]
VSS_{M_3}	–	–	–	[−1.68, 0.92]	[−9.71, −7.23]
VSS_{M_4}	–	–	–	–	[−9.44, −6.73]
VSS_{M_5}	–	–	–	–	–

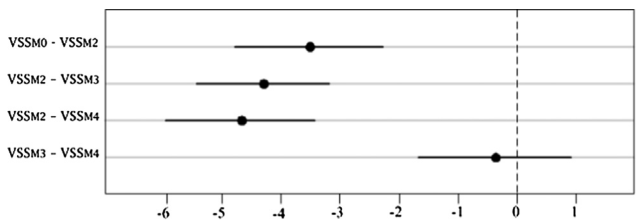

Fig. 8. Plot for the pairwise comparaison of the confidence intervals.

We also observe that the Lognormal model M_5, gave the less OTIF loss, $VSS_{M_5} = 11.13\%$. In other words, if we assume that the roadway travel time is Lognormally distributed and we solve the corresponding stochastic SND problem, the found Path deteriorates the OTIF by an average of 11.13% comparatively to the OTIF we can reach in case we solved the problem using the real Mixture Gamma distributed travel times. This is not surprising since the skewed Lognormal distribution better fit at least the first mode of the Mixture Gamma. We also notice that the Uniform, Triangular and Normal assumptions on travel time distributions gave substantial loss of respectively $VSS_{M_2} = 23.92\%$, $VSS_{M_3} = 19.60\%$ and $VSS_{M_4} = 19.22\%$. It is worth to mention that in former work (Demir et al., 2016), the travel time is assumed to follow such Triangular distribution. We shown that this model may lead to an average performance loss of 19.60%. In fact, the Triangular distribution is unable to model travel time with strong positive skew and long upper tail. Also, since the Uniform distribution has no mode, this model can lead to substantial performance loss. Actually, the more accurate we fit the baseline model M_1, where travel times are distributed according to the Mixture Gamma, the less OTIF loss we have, i.e. the more reliable Path.

We can conclude that, even when we fixed the same mean and variance for all Models M_i, the difference in the data distribution shapes have considerable effect on results. That means that, in order to find reliable service Path, we have to go beyond the empirical mean and variance calculation from historical data by finding the best fitting distribution type. Hence, to solve efficiently such stochastic optimization problem, we need to consider true empirical distributions of uncertain parameters. This need is enhanced when the corresponding empirical data exhibit multimodality.

6. Conclusion

In this paper, we have proposed a new simulation-based optimization model for scheduling problem in stochastic multimodal freight transportation systems. At first, this SBO-model was applied to solve a deterministic optimization problem in a real-world case study. We showed that, although computationally more expensive, our model provides the optimal freight service schedule found with the analytical formulation. Then, this SBO-model was applied to solve complex stochastic SND problem. As we are using the simulation modelling, we integrated demand and travel times stochasticity with the corresponding realistic continuous distributions. We considered typical distribution shapes, commonly used in the transportation research field, such as skewness and multimodality. The results show that our SBO-model can handle complex stochasticity and find good solutions, with an OTIF of more than 90%, in reasonable computational time of 60 min. Finally, with an extension of the VSS, we assessed the schedule reliability by calculating the average loss of using the deterministic assumption or the wrong distribution of stochastic processes.

The main finding is that, to reach reliable freight service schedule, even with low variance in stochastic input data, we should go beyond mean and variance estimation. Specifically, when such empirical data exhibit skewness and multimodality. For that purpose, data mining techniques have to be used for efficient statistical model fitting. Although we conducted our experiments with a set of known distribution types, this is not a limitation of the proposed simulation based optimization modelling approach. Indeed, we can introduce in the simulator any other distribution type or even an empirical distribution as long as there is sufficient historical data. Also, a significant achievement presented in this paper is the rigorous implementation, with the UML standard, of the sequential interaction involved in the simulation based optimization process. The lack of such formalism, for effective computer implementation, has widely been criticized in former works.

Future research is intended to extend the results based on the proposed SBO-model. The research will focus on the implementation of the Simulation-based optimization approach to solve other type of application such as for vehicle fleet sharing problems in city logistics. Using this approach to solve such class of NP-hard optimization problems will allow taking into account different stochastic aspects of the traffic system. Also, future research into simulation based multi-objective optimization is intended to accelerate the searching process.

Appendix A

See Tables A1–A3.

Table A.1
Terminal characterization per each container transshipped Demir et al. (2016).

Terminal Designation	Terminal number	Time in hour	Emission in kg	Cost in €
Budapest Port	A	0.130	2.80	19.5
Linz	B	0.015	2.80	20.5
Budapest Bilk	c	0.050	0.92	23
Regensburg	D	0.050	0.92	20
Prague	E	0.070	0.92	23
Wels	F	0.050	2.53	20
Salzburg	G	0.025	2.53	20
Vienna Rail	H	0.020	0.92	20
Munich	I	0.050	3.00	20
Vienna Port	J	0.050	0.92	15

Table A.2
Services characteristics per each container transshipped Demir et al. (2016).

Service N°s	(i_s, j_s)	Capacity	$[W_s^{\min}, W_s^{\max}]$	Time in hours	Emission in kg	Cost in €
1	(A,J)	60	[32, 32]	42	127	115
2	(J,B)	60	[76, 97]	29	53	63
3	(B,D)	42	[107, 141]	49	86	102
4	(C,I)	20	[18, 18]	38	70	182
5	(C,I)	20	[42, 42]	84	69	181
6	(C,I)	20	[114, 114]	38	71	183
7	(C,I)	20	[162, 162]	38	68	184
8	(C,F)	16	[21, 21]	35	50	201
9	(C,F)	16	[45, 45]	58	49	202
10	(C,F)	16	[117, 117]	35	51	200
11	(C,F)	16	[141, 141]	35	48	203
12	(C,F)	16	[165, 165]	35	52	199
13	(C,E)	16	[8, 8]	40	163	250
14	(C,E)	16	[80, 80]	40	162	251
15	(C,E)	16	[128, 128]	40	164	249
16	(H,F)	16	[19, 19]	11	9	113
17	(H,F)	16	[115, 115]	11	8	114
18	(H,F)	16	[139, 139]	11	10	112
19	(H,F)	16	[163, 163]	11	11	115
20	(E,G)	16	[17, 17]	81	53	109
21	(E,G)	16	[137, 137]	35	52	110
22	(A,D)	60	[0, 168]	10	390	484
23	(J,D)	60	[0, 168]	6	252	327
24	(B,D)	60	[0, 168]	3	140	196
25	(I,D)	60	[0, 168]	2	83	129
26	(D,I)	60	[0, 168]	2	83	129
27	(F,D)	60	[0, 168]	3	123	165
28	(E,D)	60	[0, 168]	4	158	209
29	(G,D)	60	[0, 168]	4	157	193
30	(D,G)	60	[0, 168]	4	157	193
31	(A,C)	60	[0, 168]	0	14	62
32	(J,H)	60	[0, 168]	0	12	63

Table A.3
Commodities/Orders characterization Demir et al. (2016).

Order N°k	(origin→destination)	Demand d_k	Release time (h)	Due time (h)	Cost (€/late hour)
1	(A→D)	20	10	160	30
2	(A→D)	10	25	170	100
3	(A→I)	15	20	80	70
4	(J→D)	9	70	159	80
5	(E→G)	6	30	102	50

References

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

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

Baykasoğlu, A., & Subulan, K. (2016). A multi-objective sustainable load planning model for intermodal transportation networks with a real-life application. *Transportation Research Part E: Logistics and Transportation Review, 95*, 207–247.

Chalumuri, R. S., & Yasuo, A. (2014). Modelling travel time distribution under various uncertainties on Hanshin expressway of Japan. *European Transport Research Review, 6*(1), 85–92.

Chávez, H., Castillo-Villar, K. K., Herrera, L., & Bustos, A. (2017). Simulation-based multi-objective model for supply chains with disruptions in transportation. *Robotics and Computer-Integrated Manufacturing, 43*, 39–49.

Chica, M., Juan Pérez, A. A., Cordon, O., & Kelton, D. (2017). Why Simheuristics? Benefits, Limitations, and Best Practices When Combining Metaheuristics with Simulation. Available at SSRN: https://ssrn.com/abstract=2919208.

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

Crainic, T. G., & Rousseau, J. M. (1986). Multicommodity, multimode freight transportation: A general modeling and algorithmic framework for the service network design problem. *Transportation Research Part B: Methodological, 20*(3), 225–242.

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

Fu, M. C. (2002). Optimization for simulation: Theory vs. practice. *INFORMS Journal on Computing, 14*(3), 192–215.

Fu, M. C., Hu, J. Q., Chen, C. H., & Xiong, X. (2007). Simulation allocation for determining the best design in the presence of correlated sampling. *INFORMS Journal on Computing, 19*(1), 101–111.

Garrido, R. A., & Mahmassani, H. S. (2000). Forecasting freight transportation demand with the space–time multinomial probit model. *Transportation Research Part B: Methodological, 34*(5), 403–418.

Glover, F., Laguna, M., & Marti, R. (2003). Scatter search and path relinking: Advances and applications. *Handbook of metaheuristics* (pp. 1–35). US: Springer.

Guessous, Y., Aron, M., Bhouri, N., & Cohen, S. (2014). Estimating travel time distribution under different traffic conditions. *Transportation Research Procedia, 3*, 339–348.

Halim, R. A., & Seck, M. D. (2011). The simulation-based multi-objective evolutionary optimization (SIMEON) framework. In *Proceedings of the winter simulation conference* (pp. 2839–2851).

Harrison, A., & Fichtinger, J. (2013). Managing variability in ocean shipping. *The International Journal of Logistics Management, 24*(1), 7–21.

Hrušovský, M., Demir, E., Jammernegg, W., & Van Woensel, T. (2016). Hybrid simulation and optimization approach for green intermodal transportation problem with travel time uncertainty. *Flexible Services and Manufacturing Journal*, 1–31.

Jaoua, A., Gamache, M., & Riopel, D. (2012). Specification of an intelligent simulation-based real time control architecture: Application to truck control system. *Computers in Industry, 63*(9), 882–894.

Juan, A. A., Faulin, J., Grasman, S. E., Rabe, M., & Figueira, G. (2015). A review of simheuristics: Extending metaheuristics to deal with stochastic combinatorial optimization problems. *Operations Research Perspectives, 2*, 62–72.

Juan, A. A., Grasman, S. E., Caceres-Cruz, J., & Bekta ş, T. (2014). A simheuristic algorithm for the single-period stochastic inventory-routing problem with stock-outs. *Simulation Modelling Practice and Theory, 46*, 40–52.

Kelton, W. D., Sadowski, R. P., & Sturrock, D. T. (2007). *Simulation with Arena* (4th ed.). Boston: Mc Graw-Hill Higher Education.

Krishnamoorthy, K. (2016). *Handbook of statistical distributions with applications* (2nd ed.). CRC Press.

Krüger, N. A., Vierth, I., & Fakhraei Roudsari, F. (2013). Spatial, temporal and size distribution of freight train delays: evidence from Sweden. Working papers in Transport Economics 2013:8, CTS - Centre for Transport Studies Stockholm (KTH and VTI).

L'ecuyer, P., Simard, R., Chen, E. J., & Kelton, W. D. (2002). An object-oriented random-number package with many long streams and substreams. *Operations Research, 50*(6), 1073–1075.

Lee, C. Y., Lee, H. L., & Zhang, J. (2015). The impact of slow ocean steaming on delivery reliability and fuel consumption. *Transportation Research Part E: Logistics and Transportation Review, 76*, 176–190.

Lee, C. Y., & Song, D. P. (2017). Ocean container transport in global supply chains: Overview and research opportunities. *Transportation Research Part B: Methodological, 95*, 442–474.

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

Meng, Q., Wang, T., & Wang, S. (2012). Short-term liner ship fleet planning with container transshipment and uncertain container shipment demand. *European Journal of Operational Research, 223*(1), 96–105.

Mönch, L., Lendermann, P., McGinnis, L. F., & Schirrmann, A. (2011). A survey of challenges in modelling and decision-making for discrete event logistics systems. *Computers in Industry, 62*(6), 557–567.

Oliveira, J. B., Lima, R. S., & Montevechi, J. A. B. (2016). Perspectives and relationships in Supply Chain Simulation: A systematic literature review. *Simulation Modelling Practice and Theory, 62*, 166–191.

Rakha, H., El-Shawarby, I., & Arafeh, M. (2010). Trip travel-time reliability: issues and proposed solutions. *Journal of Intelligent Transportation Systems, 14*(4), 232–250.

Rumbaugh, J., Jacobson, I., & Booch, G. (2004). *The unified modeling language reference manual* (second ed.). Addison-Wesley, Pearson Higher Education.

Rushton, A., Croucher, P., & Baker, P. (2014). *The handbook of logistics and distribution management: Understanding the supply chain.* Kogan Page Publishers.

SteadieSeifi, M., Dellaert, N. P., Nuijten, W., Van Woensel, T., & Raoufi, R. (2014). Multimodal freight transportation planning: A literature review. *European Journal of Operational Research, 233*(1), 1–15.

Sun, Y., Lang, M., & Wang, D. (2015). Optimization models and solution algorithms for freight routing planning problem in the multi-modal transportation networks: A review of the state-of-the-art. *The Open Civil Engineering Journal, 9*, 714–723.

Taylor, M. A., & Somenahalli, S. (2010). Travel time reliability and the bimodal travel time distribution for an arterial road. *Road & Transport Research: A Journal of Australian and New Zealand Research and Practice, 19*(4), 37.

Wang, X., Crainic, T. G., & Wallace, S. W. (2014). Stochastic scheduled service network design: the value of deterministic solutions. Technical Report CIRRELT-2014-46, Centre interuniversitaire de recherche sur les réseaux d'entreprise, la logistique et le transport, Université de Montréal.

Wang, X., & Wallace, S. W. (2016). Stochastic scheduled service network design in the presence of a spot market for excess capacity. *EURO Journal on Transportation and Logistics, 5*(4), 393–413.

Wen, C., Li, Z., Lessan, J., Fu, L., Huang, P., & Jiang, C. (2017). Statistical investigation on train primary delay based on real records: Evidence from Wuhan–Guangzhou HSR. *International Journal of Rail Transportation,* 1–20.

Yang, X., Low, J. M., & Tang, L. C. (2011). Analysis of intermodal freight from China to Indian Ocean: A goal programming approach. *Journal of Transport Geography, 19*(4), 515–527.

Yegul, M. F., Erenay, F. S., Striepe, S., & Yavuz, M. (2017). Improving configuration of complex production lines via simulation-based optimization. *Computers & Industrial Engineering, 109*, 295–312.

Zhang, D., Li, X., Huang, Y., Li, S., & Qian, Q. (2015). A robust optimization model for green regional logistics network design with uncertainty in future logistics demand. *Advances in Mechanical Engineering, 7*(12) 1687814015620518.