The multi-path Traveling Salesman Problem with stochastic travel costs

EURO J Transp Logist (2017) 6:3–23
DOI 10.1007/s13676-014-0056-2

RESEARCH PAPER

The multi-path Traveling Salesman Problem with stochastic travel costs

Roberto Tadei · Guido Perboli · Francesca Perfetti

Received: 12 February 2014 / Accepted: 4 July 2014 / Published online: 31 July 2014

Abstract Given a set of nodes, where each pair of nodes is connected by several paths and each path shows a stochastic travel cost with unknown probability distribution, the multi-path Traveling Salesman Problem with stochastic travel costs aims at finding an expected minimum Hamiltonian tour connecting all nodes. Under a mild assumption on the unknown probability distribution, a deterministic approximation of the stochastic problem is given. The comparison of such approximation with a Monte Carlo simulation shows both the accuracy and the efficiency of the deterministic approximation, with a mean percentage gap around 2% and a reduction of the computational times of two orders of magnitude.

Keywords TSP · Multiple paths · Stochastic travel costs · Deterministic approximation

1 Introduction

Recently, with an increasing worldwide concern for the environment, freight transportation has been object of new studies aiming at reducing negative

R. Tadei (✉) · G. Perboli · F. Perfetti
Politecnico di Torino, Turin, Italy
e-mail: roberto.tadei@polito.it

G. Perboli
e-mail: guido.perboli@polito.it

F. Perfetti
e-mail: francesca.perfetti@polito.it

G. Perboli
CIRRELT, Montreal, Canada

externalities due to freight distribution operations, such as pollution, accidents, noise and land use deterioration. In the past decade, City Logistics pushed researchers towards the definition of new conceptual models of transportation and supply chain integration in urban areas. These models have been extended with the introduction of the concept of Smart City (Chourabi et al. 2012), where “smart” implies a plethora of methods and disciplines within a holistic vision to mitigate the problems generated by population growth and rapid urbanization. In particular, recent efforts in the planning of freight transportation activities have focused on greenhouse gases (GHGs) emissions minimization.

In the literature concerning sustainable freight transportation planning, a huge number of papers deals with two main objectives: the quantification and forecasting of the GHGs emissions of routing activities (fuel consumption models), and the environmental concerns integration in the objective function of vehicle routing and transportation model (Demir et al. 2014). In fact, traditionally the main goal of the planning of freight transportation has been to minimize the number of vehicles or to minimize costs (usually associated to travel times or traveled distance).

Furthermore, the attention to GHGs emissions takes greater relevance within urban areas because congestion is one of the effective factors on greenhouse gases, particularly CO2. These trends (City Logistics, Smart Cities and the efforts in reducing GHGs emissions) lead to an increasing attention to the planning of road freight transportation.

At the same time, transportation technologies and fuels also have improved over the years developing electric and hybrid vehicles for freight distribution. With the introduction of new type of vehicles, new requirements arise in the planning of freight transportation.

The PIE_VERDE project, funded by the European Regional Development Fund (ERDF), aims at developing new planning tools for freight delivery in urban areas by means of environmental friendly light duty vehicles. In this project, one of the goals is to plan and manage a two-echelon delivery service. Trucks are not allowed to directly enter the city and freight is consolidated in small peripheral depots. The goods are then delivered to customers using hybrid vehicles (Perboli et al. 2011).

The planning of a hybrid vehicle tour requires the determination of the sequence of clients to visit and the selection of the powertrain during the tour. In fact, hybrid vehicles can change the powertrain during a route, impacting on their GHG emissions, energy and fuel consumption. These vehicles can be fueled by full thermal, thermal–electric or exclusively electric engine thanks to a rechargeable energy storage system able to supplement fossil fuel energy for vehicle propulsion. Additionally, some hybrid vehicles use co-generative thermal engine that exploits braking power to generate electricity while traveling. Hence, the gain in terms of GHGs emissions reduction, obtained by a hybrid vehicle instead of a traditional one, varies according to how the several powertrains have been selected during each route. An intelligent planning of the powertrain selection is a key factor to efficiently use a hybrid vehicle.

This paper has been conceived for PIE_VERDE project to create a model able to optimize both freight vehicle tour and powertrain selection for hybrid vehicles in parcel and courier deliveries. Our paper aims at meeting several current needs in the context of freight transportation planning. First, there is the need of new routing models that consider the stochasticity of generalized cost functions, which include both operational and environmental aspects.

Second, new freight distribution business models, such as parcel delivery and e-commerce freight delivery, require tiny limited time for vehicle fleet planning: about 30 minutes for planning the full fleet. Hence, the tours must be computed with a very short computational effort.

Third, in real cases, the distribution of stochastic variables are unknown and it is not easy and not always possible to derive their distribution from real data because of the small number of hybrid vehicles applied for freight distribution, the difficulty or impossibility to access to sensors data and the difficulty of getting data from the vehicles control unit.

We present the multi-path Traveling Salesman Problem with stochastic travel costs (mpTSP$_s$), a new stochastic variant of the Traveling Salesman Problem. The multi paths allow us to consider the multi powertrains of hybrid vehicles and the presence of multiple paths between two customers. In fact, with the term "path", we identify the shortest path between two nodes, performed with the selected powertrain.

In more details, given a graph characterized by a set of nodes, a set of paths, which connect pairs of nodes, and random travel costs, we want to find an Hamiltonian cycle which minimizes the expected value of the travel cost of the cycle. This is computed as the sum of the expected travel costs of the paths interconnecting the pairs of nodes, where, for each pair of node, only one path can be selected among the several ones. Furthermore, each path is characterized by a travel cost. The travel cost is a generalized cost which includes both fuel consumption and driving time. Additionally, each travel cost is composed of a deterministic term plus a random term, which represents the travel cost oscillation due to traffic congestion, driving style, etc. The several combinations between powertrain and the travel cost oscillation generate different paths between two given nodes. The model chooses the path between two nodes according to an efficiency-based decision, i.e., the path with the minimum expected travel cost is chosen. Moreover, the probability distributions of the travel costs are assumed to be unknown.

Other applications of the mpTSP$_s$ arise in the City Logistics context. Nowadays, cities offer several services, such as garbage collection, periodic delivery of goods in urban grocery distribution and bike sharing services. These services require the definition of fixed tours that will be used from one to several weeks [see, e.g., CITYLOG Consortium (2010)]. However, within urban areas, paths are affected by the uncertainty of travel time. The travel time distributions differ from one path to another and they are time dependent. Even an approximated knowledge of the travel time distribution may be made difficult due to the large size of the data involved. The usage of the travel time's mean (or other measures of the expectation) may imply relevant errors when the variance is high.

The scientific contribution of this study is threefold: (i) to introduce the mpTSP$_s$ and give a formulation, in which travel costs are assumed to be uncertain with an unknown distribution, while, at the best of our knowledge, in the literature travel costs have always been assumed to fit a specific distribution; (ii) to propose the first model able to support the planning of hybrid vehicles routing, considering the several powertrains as well as the stochasticity of the context; and (iii) to derive a deterministic approximation from the stochastic formulation, which we validate by means of extensive computational experiments.

In particular, the deterministic approximation becomes a TSP problem where the minimum expected total travel cost is equivalent to the maximum of the logarithm of the total accessibility of the Hamiltonian tours to the path set. We evaluate the quality of the deterministic approximation by comparing it with the Perfect Information results obtained by a Monte Carlo method. The comparison shows a good accuracy of the deterministic approximation, with a reduction of the computational times of two orders of magnitude. Besides, computational results show how the derived model can be solved with difficulty within the timing restrictions of the application with a reasonable accuracy.

The paper is organized as follows. In Section 2 a relevant literature is listed. Section 3 presents the stochastic model of the mpTSP$_s$ and Section 4 derives its deterministic approximation. In Section 5 we compare the results of the deterministic approximation with the results of a Monte Carlo simulation of the stochastic problem. Finally, in Section 6 conclusions are drawn.

2 Literature review

A routing problem is said to be static when its input data do not explicitly depend on time, while it is dynamic if some elements of information are revealed or updated during the period of time in which operations take place. (Ghiani et al. 2003; Berbeglia et al. 2010; Pillac et al. 2013) Moreover, a routing problem is deterministic if all input data are known before routes are constructed, otherwise it is stochastic.

According to these definitions, our mpTSP$_s$ is static and stochastic.

In this section, we discuss the main literature on stochastic TSP, as well as the results of similar and related problems, including Shortest Path and Vehicle Routing, showing the differences between our problem and the others introduced in the literature.

While different stochastic variants of TSP (and more in general of vehicle routing problem) are present in the literature (Gendreau et al. 1996; Golden et al. 2008; Pillac et al. 2013), the mpTSP$_s$ is absent. For this reason, we also consider some relevant literature on similar problems, highlighting the main differences with the problem faced in this paper.

In the literature several stochastic variants of the TSP problems can be found. In these problems, a known distribution affecting some problem parameters is given and the theoretical results are strongly connected with the hypotheses on such distribution. The main sources of uncertainty are related to the arc costs (Leipala 1978; Toriello et al. 2012) and the subset of cities to be visited with their location (Jaillet 1988; Goemans and Bertsimas 1991).

If we consider general routing problems, different types of uncertainty can be considered. The most studied variants are related to customer locations and demands, with the requests being both goods (Hvattum et al. 2006, 2007; Ichoua et al. 2006) and services (Bertsimas and Van Ryzin 1991). Only in recent years, the stochasticity related to travel times has been considered in the literature (Güner et al. 2012; Kenyon and Morton 2003; Taş et al. 2013).

All the papers presented in this literature review deal with uncertainty of the routing problems where the magnitude of the uncertainty is limited and the parameter values are revealed in a time interval compatible with the operations optimization. Then, even if multiple paths can be present between two given nodes, the multi-path aspects can be ignored, being possible as a priori choice of the path connecting the two nodes. In our case, the mpTSP$_s$ is thought to be used for planning a service. Thus, the enlarged time horizon as well as strong dynamic changes in travel costs due to traffic congestion and other nuisances typical of the urban transportation induce the presence of multiple paths connecting every pair of nodes, each one with its stochastic cost.

When we consider other routing problems related to the TSP, a large literature is available for the stochastic Shortest Path. One of the few papers directly dealing with multiple paths is due to Eiger et al. (1985). In their paper the authors consider an extension of the classical shorted route problem where multiple arcs interconnect the nodes and the costs are uncertain. In particular, they show how, when the preferences between the arcs are linear or exponential distributed, a Dijkstra-type algorithm using the mean of the distributions finds an optimal path. Unfortunately, the results are strictly related to the specific problem, the shortest path, and to the presence of a Dynamic Programming solution method. Moreover, differently from our case, the preferences must be exponentially distributed, while we assume that only the right tail converges to an exponential distribution. Psaraftis and Tsitsiklis (1993) introduce a variant of the stochastic Shortest Path where the arc costs are stochastic and dynamic, in the sense that the arc cost is a known function of a certain environment variable which depends on the time in which we are leaving from each node. Differently to our case, not only is there one single path associated to each node, but the environment variable associated to each arc is an independent stochastic process associated to a finite-state Markov process with a known transition probability matrix. Thus, this approach is not suitable to urban transportation, where the estimation of the Markov process could be not usable in practice. Finally, Jaillet and Melvyn Sim (2013) recently proposed criteria to design shortest paths when deadlines are imposed to the nodes and the goal of the problem is to minimize the deviation of the actual arrival time with respect to the desired one. They also show that the stochastic shortest path with deadlines under uncertainty can be solved in polynomial time when there is stochastic independence between the arc travel costs. Even in this case between any pair of node only one arc exists.

Another research direction in routing of stochastic networks is related to the usage of an objective function measuring the lateness of the dispatcher when arriving to the customer site (see Hame and Hakula (2013) and Cordeau et al. (2007) for a survey). Lecluyse et al. (2009) consider a variant of the Vehicle Routing

Problem where the travel costs are time dependent and the objective function is a linear combination of the mean and the variance of the travel costs of the arcs. The authors introduce instances based on realistic speed profiles and analyze the results in terms of the 95th percentile of the travel cost distribution, which is assumed to be lognormal. Lee et al. (2012) formulate another VRP variant where both customer demands and arc travel times are uncertain. The authors propose a Dantzig–Wolfe decomposition approach to encapsulate the uncertainty in the solution method and a Dynamic Programming algorithm to solve the column generation subproblem.

From the literature, two gaps come to light: first, problems considering multi paths have not been studied yet, particularly, problems taking into account both multi paths and stochastic costs; second, dealing with freight fleet transportation, there are no models and methods able to find a solution in short time, taking into account the effects of stochastic costs.

3 The mpTSP$_s$

To reproduce a real-world transportation network, two layers can be considered: the physical layer and the logical layer. The physical layer is built by a set of arcs and nodes, while the logical layer consists of paths and a subset of nodes of the physical layer. A path in the logical layer connects a pair of nodes via a set of arcs from the physical layer. In this paper, we consider the logical layer and each path identifies the shortest path between two nodes. In addiction, each path is characterized by a travel cost which is composed of a deterministic travel cost plus a random term, which represents the travel cost oscillation due to traffic congestion, driving style, different powertrains for hybrid vehicles, etc. In practice, such oscillations are actually very difficult to be measured. The scenarios are the possible realizations of the travel costs in different traffic situations. While at the operational level we know with a good approximation, for each path, the actual travel cost, this is not true at the planning level, where the tour must be built to cope with different traffic conditions and other parameters of the routes. Thus, at this level, even knowing the order of the nodes to visit, the travel cost of each path is a random variable with a probability distribution which is very difficult to measure in practice. This implies that such probability distribution must be assumed as unknown.

Let it be

- N: set of nodes
- U: subset of N
- L: set of scenarios
- K_{ij}: set of paths between nodes i and j
- c_{ij}^k: unit deterministic travel cost of path $k \in K_{ij}$
- $\tilde{\theta}_{ij}^{kl}$: random travel cost oscillation of path $k \in K_{ij}$ under scenario $l \in L$
- $\tilde{c}_{ij}^k(\tilde{\theta}_{ij}^{kl}) = c_{ij}^k + \tilde{\theta}_{ij}^{kl}$: unit random travel cost of path $k \in K_{ij}$ under scenario l
- x_{ij}^k: boolean variable equal to 1 if path $k \in K_{ij}$ is selected, 0 otherwise
- y_{ij}: boolean variable equal to 1 if node j is visited just after node i, 0 otherwise.

The mpTSP$_s$ is formulated as follows

$$\min_{\{y,x\}} \mathbb{E}_{\{\tilde{\theta}_{ij}^{kl}\}} \left[\sum_{i \in N} \sum_{j \in N} y_{ij} \sum_{k \in K_{ij}} \sum_{l \in L} \tilde{c}_{ij}^{k}(\tilde{\theta}_{ij}^{kl}) x_{ij}^{k} \right] \quad (1)$$

subject to

$$\sum_{j \in N: j \neq i} y_{ij} = 1 \qquad i \in N \quad (2)$$

$$\sum_{i \in N: i \neq j} y_{ij} = 1 \qquad j \in N \quad (3)$$

$$\sum_{i \in U} \sum_{j \notin U} y_{ij} \geq 1 \qquad \forall U \subset N \quad (4)$$

$$\sum_{k} x_{ij}^{k} = y_{ij} \qquad i \in N, \quad j \in N \quad (5)$$

$$x_{ij}^{k} \in \{0, 1\} \qquad k \in K_{ij}, \quad i \in N, \quad j \in N \quad (6)$$

$$y_{ij} \in \{0, 1\} \qquad i \in N, \quad j \in N \quad (7)$$

The objective function (1) expresses the minimization of the expected total travel cost; (2) and (3) are the standard assignment constraints; (4) is the subtour elimination constraints. Constraints (5) link the variables x_{ij}^{k} to the y_{ij}. Finally, (6)-(7) are the integrality constraints.

If the probability distribution of travel costs were known, the mpTSP$_s$ would be reduced to a deterministic problem where the expected values would be substituted to the stochastic travel costs oscillations. This is due to the linearity of the objective function, as shown by Eiger et al. (1985). Unfortunately, model (1)-(7) considers stochastic costs with unknown probability distribution. A common way to represent and solve this type of problems is to discretize the stochastic sources by means of scenario generation. This implies, even in the case of the Perfect Information computation, the need to solve several times, at least one for each scenario, the deterministic counterpart that for the TSP is known to be NP-Hard. Thus, the computational effort needed to cope with the different scenarios increases rapidly both with the size of the instance in terms of nodes and multiple paths, and with the number of scenarios. As shown in the computational experiences, in fact, a reasonable trade off between accuracy and efficiency is 100 scenarios. Due to the need of solving the mpTSP$_s$ in a very limited computational time, more efficient ways for solving it are needed. Thus, in the following we derive a deterministic approximation of the stochastic problem able to approximate, under a mild hypothesis on the probability distribution shape, even the case where the probability distribution in unknown or it is varying from path to path.

Given any pair of nodes $i, j \in N$ and any path $k \in K_{ij}$ between them, we assume that, across the alternative scenarios $l \in L$, the travel cost oscillations $\tilde{\theta}_{ij}^{kl}$ are independent and identically distributed random variables with unknown probability distribution, given by the following cumulative right distribution function

$$F_{ij}^{k}(x) = Pr\{\tilde{\theta}_{ij}^{kl} \geq x\} \tag{8}$$

Although the independence assumption could seem unrealistic, it is frequently done in the urban path-finding literature (Lecluyse et al. 2009; Jaillet and Melvyn Sim 2013). Moreover, we will see from our computational results, obtained using empirical data of a middle-sized city, that even with travel cost oscillations which are dependent our deterministic approximation gives very good results.

The assumption of identical distributions for the travel cost oscillations is quite more rare and stronger. Nevertheless, it is mitigated, as it is shown in 4, by the fact that the only common property required to these distributions is to be asymptotically exponential in their left tail. This is a very mild assumption as we observe that many probability distributions show such behavior, among them are the widely used distributions Exponential, Normal, Lognormal, Gamma, Gumbel, Laplace, and Logistic. Also in this case, our computational results show that even with different probability distributions for the scenarios the deterministic approximation is very accurate.

Following Tadei et al. (2012), we define $\tilde{\theta}_{ij}^{k}$ as the minimum of the random travel cost oscillations $\tilde{\theta}_{ij}^{kl}$ of path $k \in K_{ij}$ across the alternative scenarios $l \in L$

$$\tilde{\theta}_{ij}^{k} = \min_{l \in L} \tilde{\theta}_{ij}^{kl} \qquad k \in K_{ij}, \quad i \in N, \quad j \in N \tag{9}$$

Let B_{ij}^{k} be the cumulative right distribution function of $\tilde{\theta}_{ij}^{k}$

$$B_{ij}^{k}(x) = \Pr\left\{\tilde{\theta}_{ij}^{k} \geq x\right\} \tag{10}$$

As, for any path $k \in K_{ij}$, $\tilde{\theta}_{ij}^{k} \geq x \Longleftrightarrow \tilde{\theta}_{ij}{}^{kl} \geq x$, $l \in L$ and $\tilde{\theta}_{ij}{}^{kl}$ are independent and identically distributed across the alternative scenarios $l \in L$, using (8) one gets

$$B_{ij}^{k}(x) = \prod_{l \in L} \Pr\left\{\tilde{\theta}_{ij}{}^{kl} \geq x\right\} = \prod_{l \in L} F_{ij}^{k}(x) = \left[F_{ij}^{k}(x)\right]^{|L|} \tag{11}$$

We relax the problem by assuming that we can choose across all scenarios $l \in L$. Being the routing efficiency-based, the scenario $l \in L$ that minimizes the random travel cost $\tilde{c}_{ij}^{k}(\tilde{\theta}_{ij}^{kl})$ will be selected.

Then, the random travel cost of path $k \in K_{ij}$ becomes

$$\tilde{c}_{ij}^k(\tilde{\theta}_{ij}^k) = \min_{l \in L} \tilde{c}_{ij}^k(\tilde{\theta}_{ij}^{kl}) = c_{ij}^k + \min_{l \in L} \tilde{\theta}_{ij}^{kl} = c_{ij}^k + \tilde{\theta}_{ij}^k \qquad k \in K_{ij}, \quad i \in N, \quad j \in N \tag{12}$$

The minimum travel cost oscillation $\tilde{\theta}_{ij}^k$ can be either positive or negative, but, in practice, its support is such that no negative travel costs c_{ij}^k exist, so that $\tilde{c}_{ij}^k(\tilde{\theta}_{ij}^k)$ is always non-negative.

For each pair of node (i, j), let us consider the path k^* (for the sake of simplicity, we assume it is unique) which gives the minimum random travel cost.

The minimum random travel cost between i and j is then

$$\tilde{c}_{ij}(\tilde{\theta}_{ij}^{k^*}) = \min_{k \in K_{ij}} \tilde{c}_{ij}^k(\tilde{\theta}_{ij}^k) \qquad i \in N, \quad j \in N \tag{13}$$

and the optimal variables $\left\{x_{ij}^k\right\}$ of problem (1)-(7) become

$$x_{ij}^k = \begin{cases} 1, & \text{if} \quad \mathrm{k} = \mathrm{k}^* \\ 0, & \text{otherwise} \end{cases} \tag{14}$$

Using (13), (14), and the linearity of the expected value operator $\mathbb{E}$, the objective function (1) becomes

$$\min_{\{y\}} \mathbb{E}_{\{\tilde{\theta}_{ij}^{k^*}\}} \left[\sum_{i \in N} \sum_{j \in N} y_{ij} \tilde{c}_{ij}(\tilde{\theta}_{ij}^{k^*}) \right] = \min_{\{y\}} \sum_{i \in N} \sum_{j \in N} y_{ij} \mathbb{E}_{\{\tilde{\theta}_{ij}^{k^*}\}} \left[\tilde{c}_{ij}(\tilde{\theta}_{ij}^{k^*}) \right] = \min_{\{y\}} \sum_{i \in N} \sum_{j \in N} y_{ij} \hat{c}_{ij} \tag{15}$$

where

$$\hat{c}_{ij} = \mathbb{E}_{\{\tilde{\theta}_{ij}^{k^*}\}} \left[\tilde{c}_{ij}(\tilde{\theta}_{ij}^{k^*}) \right] \qquad i \in N, \quad j \in N \tag{16}$$

The mpTSP$_s$ then becomes

$$\min_{\{y\}} \sum_{i \in N} \sum_{j \in N} y_{ij} \hat{c}_{ij} \tag{17}$$

subject to (2)–(7).

However, the calculation of $\hat{c}_{ij}$ in (17) requires knowing the probability distribution of the minimum random travel cost between i and j, i.e. $\tilde{c}_{ij}(\tilde{\theta}_{ij}^{k^*})$, which will be derived in the next section.

4 The deterministic approximation of the mpTSP$_s$

By (12) and (13), let

$$G_{ij}(x) = Pr\left\{\tilde{c}_{ij}(\tilde{\theta}_{ij}^{k^*}) \geq x\right\} = Pr\left\{\min_{k \in K_{ij}} \tilde{c}_{ij}^k(\tilde{\theta}_{ij}^k) \geq x\right\} \qquad i \in N, \quad j \in N \tag{18}$$

be the cumulative right distribution function of the minimum random travel cost between i and j.

As, for any pair of nodes $i, j \in N$ and any path $k \in K_{ij}$, $\min_{k \in K_{ij}} \tilde{c}_{ij}^k(\tilde{\theta}_{ij}^k) \geq x \Longleftrightarrow \tilde{c}_{ij}^k(\tilde{\theta}_{ij}^k) \geq x$, due to (10) and (11), $G_{ij}\{x\}$ in (18) becomes a function of the total number $|L|$ of scenarios as follows

$$\begin{aligned} G_{ij}(x, |L|) &= Pr\left\{\min_{k \in K_{ij}} \tilde{c}_{ij}^k(\tilde{\theta}_{ij}^k) \geq x\right\} = \prod_{k \in K_{ij}} Pr\left\{\tilde{c}_{ij}^k(\tilde{\theta}_{ij}^k) \geq x\right\} \\ &= \prod_{k \in K_{ij}} Pr\left\{\tilde{\theta}_{ij}^k \geq x - c_{ij}^k\right\} = \prod_{k \in K_{ij}} B_{ij}^k\left(x - c_{ij}^k\right) \\ &= \prod_{k \in K_{ij}} \left[F_{ij}^k\left(x - c_{ij}^k\right)\right]^{|L|} \qquad i \in N, \quad j \in N \end{aligned} \tag{19}$$

Let us assume that $|L|$ is large enough to use the asymptotic approximation $\lim_{|L| \to +\infty} G_{ij}(x, |L|)$ as a good approximation of $G_{ij}(x)$, i.e.

$$G_{ij}(x) = \lim_{|L| \to +\infty} G_{ij}(x, |L|)) \qquad i \in N, \quad j \in N \tag{20}$$

The calculation of the limit in (20) would require knowing the probability distribution $F_{ij}^k(.)$ in (8), which is unknown. From Tadei et al. (2012), we know that under a mild assumption on the shape of the unknown probability distribution $F_{ij}^k(.)$ (i.e. it is asymptotically exponential in its left tail), the limit in (20) tends towards the following Gumbel probability distribution (Gumbel 1958), which is known as the extreme values distribution

$$G_{ij}(x) = \lim_{|L| \to +\infty} G_{ij}(x, |L|)) = \exp\left(-A_{ij}e^{\beta x}\right) \qquad i \in N, \quad j \in N \tag{21}$$

where $\beta > 0$ is a parameter to be calibrated and

$$A_{ij} = \sum_{k \in K_{ij}} e^{-\beta c_{ij}^k} \qquad i \in N, \quad j \in N \tag{22}$$

is the accessibility, in the sense of Hansen (1959), of the pair of nodes i, j to the set of paths between i and j.

Using the probability distribution $G_{ij}(x)$ given by (21), after some manipulations, $\hat{c}_{ij}$ in (16) becomes

$$\hat{c}_{ij} = -\int_{-\infty}^{+\infty} x dG_{ij}(x) = \int_{-\infty}^{+\infty} x \exp\left(-A_{ij}e^{\beta x}\right) A_{ij} e^{\beta x} \beta dx = -\frac{1}{\beta}(\ln A_{ij} + \gamma) \tag{23}$$
$$i \in N, \quad j \in N$$

where $\gamma \simeq 0.5772$ is the Euler constant.

By (23) and up to the constant $-\frac{\gamma}{\beta}\sum_{i\in N}\sum_{j\in N} y_{ij} = -\frac{\gamma}{\beta}|N|$, (17) becomes

$$\begin{aligned} &\min_{\{y\}} \sum_{i\in N}\sum_{j\in N} -\frac{1}{\beta} y_{ij} \ln A_{ij} = \\ &= \frac{1}{\beta}\max_{\{y\}} \sum_{i\in N}\sum_{j\in N} \ln A_{ij}^{y_{ij}} = \\ &= \frac{1}{\beta}\max_{\{y\}} \ln \prod_{i\in N}\prod_{j\in N} A_{ij}^{y_{ij}} = \\ &= \frac{1}{\beta}\max_{\{y\}} \ln \Phi \end{aligned} \tag{24}$$

subject to (2)-(7), where $\Phi = \prod_{i\in N}\prod_{j\in N} A_{ij}^{y_{ij}}$ is the total accessibility of the set of arcs of an optimal Hamiltonian tour to the global set of paths.

From (24), it is interesting to observe that the expected minimum total travel cost is equivalent, but the constant $\frac{1}{\beta}$, to the maximum of the logarithm of the total accessibility.

5 Computational results

In this section, we present and analyze the results of the computational experiments. The goal is to evaluate the effectiveness of the deterministic approximation of the mpTSP$_s$ we derived. In our computational experiments, travel costs are associated to travel times.

We do this by comparing our deterministic approximation with the Perfect Information case, computed by means of a Monte Carlo simulation performed on the stochastic problem. The Perfect Information is one of the most used methods in stochastic programming to evaluate whether an approximated approach is nearly optimal or inaccurate. The expected value of perfect information (EVPI) measures the maximum amount a decision maker would be ready to pay in return for complete information about the future (Birge and Louveaux 1997). An alternative method to EVPI is the one obtained by replacing all random variables with their expected values. This is called the expected value problem or mean value problem and it is used to calculate the value of the stochastic solution (VSS). EVPI measures the value of knowing the future with certainty, while VSS assesses the value of knowing and using distributions on future outcomes. Thus, the former is used for deciding whether to undertake additional efforts becomes more practically relevant, while the latter is used where no further information about the future is available. After these considerations, we decided to implement the EVPI approach.

The Monte Carlo simulation is implemented in C++, with the underlying TSP instances solved by means of the Concorde TSP solver (Applegate et al. 2007; Cook 2012). Experiments were performed on an Intel I7 2 GHz workstation with 8 GB of RAM.

Section 5.1 introduces the instance sets. The details of the Monte Carlo simulation are presented in Section 5.2. The calibration of the parameter involved in the deterministic approximation of the mpTSP$_s$ is described in Section 5.3, whilst the comparison between the Monte Carlo simulation and the approximated results is given in Section 5.4.

5.1 Instance sets

No real-life instances are present in the literature for this stochastic version of the TSP problem. Then, we generate two instance sets. In the first instance set, Set1, travel times are generated according to realistic rules, i.e. they are related to the length of the associated arcs of the considered TSP instances and there is no correlation between different scenarios and the stochastic variables are independent.

In urban areas, travel times are linked to the vehicle speed profile distributions. Moreover, to assume an actual independence of the vehicle speed of different paths could be wrong, both from the geographic (portions of paths in common) and from the time (evolution of the traffic flows in contiguous interval times) point of view. To take into account these two aspects, the second instance set, Set2, is heavily based on the real traffic sensor network of the medium sized city, Turin in Italy, which allows to better reflect real cases of City Logistics applications. Moreover, time correlation is considered using data taken from a large and continuous interval time (a full week) and a proper scenario generation algorithm.

5.1.1 Set1

In Set1, we generated instances, partially based on those available in the TSPLIB (Reinelt 1991) for the deterministic TSP problem. According to the literature, we generated the stochastic travel times according to the guidelines presented in Kenyon and Morton (2003):

- Instances. To limit the computational time, which is mainly due to the Monte Carlo simulation, we considered all instances with a number of nodes up to 200 in the TSP Library set. In particular, we split those instances into two sets: 11 instances with up to 100 nodes (N100) and 15 instances with number of nodes between 101 and 200 (N200).
- Nodes. The nodes and their position on the plane are the same as the original TSP instances.
- Multiple paths. The number of paths between any pair of nodes is set to 1, 3, and 5. Although the mpTSP$_s$ hypothesizes that several paths are present between any pair of nodes, we decided to also test the case where only one path is available. In fact, it is interesting to observe the behavior of the approximation in an extreme situation where the aspect characterizing the problem is just the stochasticity of the travel times on a single path.
- Path travel times. The travel time c_{ij}^k associated to each path k between nodes i and j is considered as a function of the Euclidean distance between i and j. In

detail, this travel time has been drawn from $U(EC_{ij}, 3EC_{ij})$, where EC_{ij} is the Euclidean distance between i and j and U is the uniform distribution. The random travel time oscillations θ_{ij}^k have been drawn as $\mathcal{D}(-c_{ij}^k/2, 2c_{ij}^k)$, where $\mathcal{D}$ is a probability distribution with its support limited to 50 and 100% of the corresponding deterministic cost, such that $c_{ij}^k + \theta_{ij}^k \geq EC_{ij}$. For $\mathcal{D}$ we have considered both the Uniform and the Gumbel distribution.

5.1.2 Set2

As instances of Set1 do not fully reflect real cases of City Logistics applications, in the following we discuss how we have generated new instances based on the real traffic sensor network of the city of Turin, in Italy. In fact, the assumptions about the independence and the equal distribution of the stochastic travel times do not hold in real urban settings. The aim of Set2 is to introduce spatial and time correlation between the variables. The spatial correlation is obtained using data taken from real speed sensors, while the time correlation is imposed by generating scenarios where the speed profiles of the paths are grouped in subsequent time intervals. In details, we consider that travel times are directly correlated to speed profile distributions. Thus, we apply two different speed profile distributions: an empirical one, whose values are obtained by data from a real sensor network in the city of Turin, and a theoretical one where the speed values are distributed accordingly to a given distribution. Being our deterministic approximation based on the extreme values theory, we choose the Gumbel distribution for the second speed profile. In this way, the theoretical distribution allows us to measure the error due to the bias introduced by our approximation itself. Hence, the comparison between the empirical and theoretical speed distribution results shows the error due to the bias introduced by our approximation and the error due to the real data distribution.

In the literature, a consistent number of papers investigates the correlation between speed distribution and travel times, particularly when dealing with road congestion (Weisbrod et al. 2001; Figliozzi 2010a, b). For example, Figliozzi (2010a) studies the correlation between congestion, travel times and depot-customer travel distance.

We follow the same schema presented in Figliozzi (2011) for the empirical speed profile distribution, because also this paper refers to data provided by a real sensor network. Furthermore, dealing with travel times, we use a simplified distance computation, applied in many other papers in the literature, such as Kenyon and Morton (2003) and Franceschetti et al. (2013):

- Instances. As in Set1, we split these instances into two sets: 11 instances with up to 100 nodes (N100) and 15 instances with number of nodes between 101 and 200 (N200).
- Nodes. Given the portion of plane containing the nodes of the original TSP instances and their position, they are mapped over a square of 14 km edge,

which is equivalent to a medium sized city like Turin. The set of nodes is partitioned into two subsets:

- Central nodes: the nodes belonging to the city center, which are the nodes in the circle where the center coincides with the geometric center of the 14 km square and a radius equal to 7 km;
- Suburban nodes: the nodes which are not central.

– Pair of nodes types: the pairs of nodes can be homogeneous or heterogeneous.

- Homogeneous: they are pairs of nodes where the starting node i and the destination node j are both central. In this case all the multiple paths between the nodes present the empirical speed profile of a central speed sensor.
- Heterogeneous: these are pairs of nodes where at least i or j belongs to the suburban set. In this case the multiple paths between the nodes present the empirical speed profile of a central speed sensor for $1/3$ of the paths and a suburban one for the $2/3$ of them if the paths are more than 1. If there is only one path between i and j, it has a suburban speed profile.

– Multiple paths. The number of paths is set to 3, and 5.

– Speed profile. We use two speed profiles, one empirical based on real data taken from speed sensors placed in the town of Turin and a theoretical one based on a Gumbel distribution of the speed profiles. In details, for each path k connecting nodes i and j and each scenario l, the speed velocity v_{ij}^{kl} is computed as follows:

- Empirical speed profile distributions, ve_{ij}^{kl}: we generate central and suburban speed profile distributions from real data on the traffic of Turin available at the website http://opendata.5t.torino.it/get_fdt. The data of the mean vehicle speed, expressed in kilometers per hour (km/h), are accessible with an accuracy of 5 min. We aggregated them in blocks of 30 minutes, for a total of 48 observations per day. The instances refer to 9 central speed sensors locations and 18 suburban ones in the period from 13 to 17 February 2013 (see the two circles in Fig. 1, giving the distribution of the actual sensors). Thus, given a path k associated to a pair of nodes (i,j) in the scenario l, an empirical speed ve_{ij}^{kl} is randomly taken from the database of the real data.
- Theoretical speed profile distribution: $vt_{ij}^{kl} = \mathcal{G}(-\gamma_{ij}^{k}/2, 2\gamma_{ij}^{k})$, where $\mathcal{G}$ is a Gumbel distribution truncated between $-\gamma_{ij}$ and $2\gamma_{ij}^{k}$ and γ_{ij}^{k} is the mean over all speed velocities generated by the empirical speed profile distribution of a path k between the nodes i and j in all the generated scenarios.

– Path travel times. The travel time $\tilde{c}_{ij}^{kl}$ is a function of the Euclidean distance between i and j, EC_{ij}, the type of pair of nodes, k, and the speed profile v_{ij}^{kl} associated to the path k between i and j under scenario l. v_{ij}^{kl} is equal to ve_{ij}^{kl} or vt_{ij}^{kl}

accordingly to the used speed distribution, empirical or theoretical respectively. In detail, this travel time has been computed as

$$\tilde{c}_{ij}^{kl} = c_{ij}^{k} + \tilde{\theta}_{ij}^{kl} = \frac{EC_{ij}}{v_{ij}^{kl}} \tag{25}$$

and

$$\tilde{c}_{ij}^{k} = \mathbb{E}_{l \in L} \frac{EC_{ij}}{v_{ij}^{kl}} \tag{26}$$

is the average travel time over all scenarios $l \in L$, associated to the path k between nodes i and j. The random travel time oscillations are then computed as

$$\tilde{\theta}_{ij}^{kl} = \frac{EC_{ij}}{v_{ij}^{kl}} - \mathbb{E}_{l \in L} \frac{EC_{ij}}{v_{ij}^{kl}} \tag{27}$$

– Time correlation. The time correlation is defined as follows. Given a scenario l^* and the speed sensor associated to the path k between i and j , we randomly choose among the 48 available a given time block $t*$. From the empirical speed profile distribution we obtain $v_{ij}^{kl^*}$. Thus, for the next σ scenarios the speed sensor associated to the path k between i and j is fixed and $v_{ij}^{k(l^*+\sigma)}$ is given by the empirical speed profile distribution at time block $t^* + \sigma$. The values of time correlation used in our experiments are 0, 2 and 4 ($\sigma = 0$ means that the association between a path, a real speed sensor and the time block is randomly chosen in each scenario).

5.2 Monte Carlo simulation

To evaluate the stochastic objective function of our problem for Set1, we used a Monte Carlo simulation. Our Monte Carlo simulation repeats the following overall process I times:

– Create S scenarios with the random costs θ_{ij}^{k} generated as described in 5.1.
– Solve each scenario as follows. Build a TSP with the node set equal to the node set of the stochastic problem. Set the cost c_{ij} between nodes i and j as $c_{ij} = \min_k (c_{ij}^k + \theta_{ij}^k)$. Indeed, when a cost scenario becomes known, its optimal solution is obtained using, as a path between the two nodes, the path with the minimum random travel time. The scenarios are solved to optimality by means of the Concorde TSP solver.
– Given the scenario optima, compute the expected value of the total cost.
– Compute the distribution of the expected value of the total cost for the scenario-based simulations.

To obtain the most reliable results of the Monte Carlo simulation, we performed a set of tuning testbeds using a subset of instances (5 from N100 and 5 from N200). The values for the parameters I (number of repetitions) and N (number of scenarios)

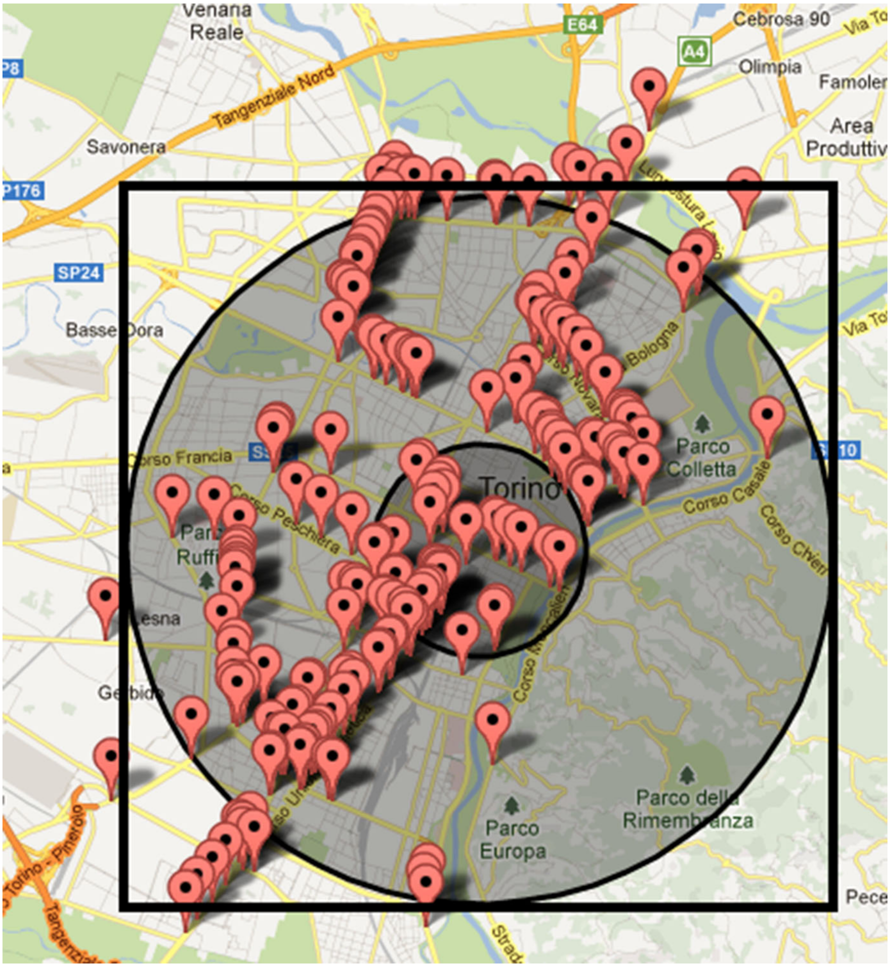

Fig. 1 Distribution of central (*gray circle*) and suburban speed sensors in the city of Turin in Italy

have been set such that the standard deviation of the distribution of the expected value was less than 1% of its mean. These values were $I = 10$ and $N = 100$.

Concerning Set2, to evaluate the quality of the deterministic approximation compared to the stochastic objective function of our problem, we used the same approach previously described for Set1. Thus, the stochastic problem is solved by means of a Monte Carlo simulation, while when needed, the TSP instances are solved by means of the Concorde TSP solver (Applegate et al. 2007; Cook 2012). Each instance is solved using the empirical speed profile and the Gumbel distribution defined as in Set1. This is done to give a comparison between the ideal situation for the deterministic approximation (the speed distribution is a Gumbel) and the empirical one.

5.3 Calibration of the β parameter

The deterministic approximation of the mpTSP$_s$ requires, see (24), an appropriate value of the parameter β. This parameter describes the propensity of the model to choose among the set of the paths characterized by different random travel times.

β is obtained by calibration as follows. Let us consider the standard Gumbel distribution $G(x) = \exp(e^{-x})$. If an approximation error of 2‰ is accepted, then $G(x) = 1 \Leftrightarrow x = 6.08$ and $G(x) = 0 \Leftrightarrow x = -1.76$. Let us consider the distribution range $[m, M]$. The following equations hold

$$\beta(m - \zeta) = -1.76 \tag{28}$$

$$\beta(M - \zeta) = 6.08 \tag{29}$$

where ζ is the mode of the Gumbel distribution $G(x) = \exp\left(e^{-\beta(x-\zeta)}\right)$.

By subtracting (28) from (29) one gets for β the value

$$\beta = \frac{6.08 - (-1.76)}{M - m} = \frac{7.84}{M - m} \tag{30}$$

According to our random oscillations rule, m is set equal to $\min_{i,j} EC_{ij}$. To calculate M we need to know the order of magnitude of the travel time oscillations in the final solution. This is needed to avoid considering those paths with travel times very far from the travel times in the solution, which could lead us to overestimate M. In fact, the presence of paths with a travel time much greater of the mean travel time is a quite common situation in the TSP and VRP problems.

M has been calculated as follows

- Solve a TSP instance with the same node set of the stochastic problem and the cost of each arc determined as $c_{ij} = \min_k c_{ij}^k$. Let us call C_D the optimum of this deterministic instance.
- Set $m = \min_{i,j} EC_{ij}$ and $M = \frac{2KC_D}{|N|}$, where $|N|$ is the number of nodes and K is the number of paths. The rationale of the formula for calculating M is that $C_D/|N|$ gives us the order of magnitude of the mean deterministic cost, which, given the rules we used to generate the instances, can have a maximum oscillation of 100%. The number of paths K is used for normalizing the accessibility effect when the path cardinality increases.

More sophisticated methods to calibrate β can be found in Galambos et al. (1994).

5.4 Comparison of deterministic approximation results and Monte Carlo simulation

Here we summarize the results for all instances with different combinations of the parameters. The performance, in terms of percentage gap, is defined as the relative percentage error of the approximated optimum when compared to the mean of the expected value given by the Monte Carlo simulation (Maggioni and Wallace 2012).

Table 1 reports the percentage gap for all combinations of the parameters, while varying the probability distribution (either Uniform or Gumbel) in Set1. The first two columns display the instance set and the number of paths between any pair of nodes, while Columns 3-4 report the mean of the percentage gaps. The best mean values are obtained for the Gumbel distribution. For both distributions, the best results are obtained with one path between the nodes, with a gap of less than 1% for the Gumbel distribution. This gap increases with the number of paths. The quality of the approximation seems to be inversely correlated with the number of nodes. However, the percentage gap is, in all cases, quite limited, with a worst case of 7.77% for the Uniform and 4.46% for the Gumbel distribution.

Computational times of Set1, expressed in seconds, are reported in Table 2. Notice that, as the computational time in both cases (Monte Carlo and deterministic approximation) are mainly given by the TSP instances computational time and the number of the TSP instances are independent from the number of multiple paths, the computational times are independent from the number of multiple paths. Thus, the results are summarized by considering the aggregation of the paths. The Monte Carlo simulation needs a computational time of about 2 orders of magnitude greater than the deterministic approximation. This makes the deterministic approximation increasingly appealing when applied to large instances, where the Monte Carlo simulation becomes impracticable.

Table 1 Set1: Percentage gap between the deterministic approximation and the Monte Carlo simulation

Nodes	Path	Uniform	Gumbel
N100	1	1.32	0.62
	3	3.41	1.86
	5	4.01	2.22
Avg		2.91	1.57
N200	1	0.71	0.35
	3	7.46	3.13
	5	7.77	4.46
Avg		5.31	2.64
Global avg		4.30	2.19

Table 2 Set1: Computational times in seconds of the deterministic approximation and the Monte Carlo simulation

Nodes	Approx	Monte Carlo
N100	5.52	523.40
N200	14.54	1507.71
Global avg	10.72	1091.27

Table 3 Set2: Percentage gaps between the deterministic approximation and the Monte Carlo simulation for the empirical and the theoretical speed profile distributions

Nodes	Path	Empirical			Theoretical
		$\sigma = 0$	$\sigma = 2$	$\sigma = 4$	
N100	1	2.72	2.87	3.41	0.87
	3	4.05	4.32	4.44	2.85
	5	4.93	5.87	6.26	2.73
Avg		3.90	4.35	4.70	2.15
N200	1	2.57	3.19	3.24	1.14
	3	6.93	8.47	8.69	3.84
	5	7.52	9.13	9.84	5.16
Avg		5.68	6.93	7.26	3.38
Global avg		4.79	5.64	5.98	2.76

As stated before, our assumptions about the independence and the equal distribution of the stochastic costs are not holding in real urban case studies. To show how the accuracy of the deterministic approximation is deteriorated by time and spatial correlation we present the results of Set2 in Table 3. The table gives the percentage gaps between the deterministic approximation and the Monte Carlo simulation for the empirical speed profile distribution and the theoretical one. In particular, for each value of the Time correlation parameter, a different column of the empirical speed profile distribution is given.

The first thing that can be noticed is how, by introducing real data which imply a spatial correlation between the different real roads of the city considered in this study ($\sigma = 0$), we have a deterioration of the results quite limited (about 2 % points). Also the time correlation implies worse results, which are, in any case, limited in mean to 1%. From the point of view of the number of path, the empirical speed profile distribution presents a sort of asymptotic behavior, with the main increase due to the presence of multiple paths, while the gap when considering 3 and 5 multiple paths remains almost stable.

The average percentage gap of the theoretical speed profile distribution shows that our deterministic approximation introduces a percentage error of about 2 in the case of 100 nodes and about 3 in the case of 200 nodes. Furthermore, the percentage error due to the anomalous distribution of real data is about 2 in both cases. Hence, we can deduce that the approximation correctly behaves even with real data distribution.

From the computational time point of view, we do not report the details, being almost equal to the ones of Set1, with the deterministic approximation showing better performances that can be measured in about 2 orders of magnitude.

In conclusion, the results seem very promising. The deterministic approximation performs quite well for all types of instances and distributions and guarantees a good accuracy. The best performance is obtained when the random travel times have a Gumbel distribution, that is usually the case for real travel time random oscillations.

6 Conclusions

In this paper we have addressed the multi-path Traveling Salesman Problem with stochastic travel costs, which consists in finding an expected minimum Hamiltonian tour connecting all nodes, where between each pair of nodes several paths do exist and each path shows a stochastic travel cost with unknown distribution.

From a theoretical perspective, the paper shows that, under a mild assumption, the probability distribution of the minimum random travel cost between any pair of nodes becomes a Gumbel distribution. Moreover, the expected minimum total travel cost is proportional to the maximum of the logarithm of the total accessibility of the Hamiltonian tours to the path set.

The deterministic approximation of the stochastic model provides very promising results on a large set of instances in negligible computational times.

In conclusion, the performance of the methodology proposed is particularly good when the probability distribution of the random travel costs of the stochastic model is a Gumbel distribution, even if good results are also provided with the Uniform distribution. This feature makes our deterministic approximation a good predictive tool for addressing stochastic travel costs in multi-path networks.

Acknowledgments Partial funding was provided by the Regional Council of Piedmont (Italy) under the ”PIE VERDE” project, P.O.R.— FESR 2007-2013— ”AUTOMOTIVE”, and by the Italian University and Research Ministry under the UrbeLOG project-Smart Cities and Communities.

References

Applegate DL, Bixby RE, Chvátal V, Cook WJ (2007) The Traveling Salesman Problem: A Computational Study. Princeton University Press, Princeton

Baldi MM, Ghirardi M, Perboli G, Tadei R (2012) The capacitated transshipment location problem under uncertainty: A computational study. Procedia - Social and Behavioral Sciences 39:424–436

Berbeglia G, Cordeau J-F, Laporte G (2010) Dynamic pickup and delivery problems. Eur J Operat Res 202:8–15

Bertsimas D, Van Ryzin G (1991) A stochastic and dynamic vehicle-routing problem in the Euclidean plane. Operat Res 39:601–615

Birge JR and Louveaux FV (1997) Introduction to stochastic programming. Springer

Chourabi H, Nam T, Walker S, Gil-Garcia JR, Mellouli S, Nahon K, Pardo TA, Scholl HJ (2012) Understanding smart cities: An integrative framework. Proceedings of the 2012 45th Hawaii International Conference on System Sciences HICSS ’12. IEEE Computer Society, Washington, pp 2289–2297

CITYLOG Consortium (2010). CITYLOG European project

Cook WJ (2012) In Pursuit of the Traveling Salesman: Mathematics at the Limits of Computation. Princeton University Press, Princeton

Cordeau J-F, Laporte G, Savelsbergh M, Vigo D (2007) Vehicle Routing. In: Barnhart C, Laporte G (eds) Transportation Handbooks on Operations Research and Management Science. North Holland, Amsterdam, pp 367–428

Demir E, Bektas T, Laporte G (2014) A Review of Recent Research on Green Road Freight Transportation. Eur J Operat Res 237:775–793

Eiger A, Mirchandani PB, Soroush H (1985) Path preferences and optimal paths in probabilistic networks. Transportation Science 19:75–84

Figliozzi MA (2010) The impacts of congestion on commercial vehicle tour characteristics and costs. Transp res part E: logist transp rev 46:496–506

Figliozzi MA (2010) An iterative route construction and improvement algorithm for the vehicle routing problem with soft time windows. Transp Res Part C: Emerg Technol 18:668–679

Figliozzi MA (2011) The impacts of congestion on time-definitive urban freight distribution networks CO2 emission levels: Results from a case study in Portland, Oregon. Transp Res Part C: Emerg Technol 19:766–778

Franceschetti A, Honhon D, Van Woensel T, Bektaş T, Laporte G (2013) The time-dependent pollution-routing problem. Transp Res Part B: Methodol 56:265–293

Galambos J, Lechner J, Simiu E (1994) Extreme Value Theory and Applications. Kluwer

Gendreau M, Laporte G, Séguin R (1996) Stochastic vehicle routing. Eur J Operat Res 88:3–12

Ghiani G, Guerriero F, Laporte G, Musmanno R (2003) Real-time vehicle routing: Solution concepts, algorithms and parallel computing strategies. Eur J Operat Res 151:1–11

Goemans MX, Bertsimas D (1991) Probabilistic analysis of the Held and Karp lower bound for the Euclidean traveling salesman problem. Math Operat Res 16:72–89

Golden BL, Raghavan S, Wasil EA (2008) The Vehicle Routing Problem: Latest Advances and New Challenges. Springer, New York

Gumbel EJ (1958) Statistics of Extremes. Columbia University Press, Columbia

Güner AR, Murat A, Chinnam RB (2012) Dynamic routing under recurrent and non-recurrent congestion using real-time ITS information. Comput Operat Res 39:358–373

Hame L, Hakula H (2013) Dynamic journeying under uncertainty. Eur J Operat Res 225:455–471

Hansen W (1959) How accessibility shapes land use. J Am Inst Planners 25:73–76

Hvattum LM, Lokketangen A, Laporte G (2006) Solving a dynamic and stochastic vehicle routing problem with a sample scenario hedging heuristic. Transp Sci 40:421–438

Hvattum LM, Lokketangen A, Laporte G (2007) A branch-and-regret heuristic for stochastic and dynamic vehicle routing problems. Networks 49:330–340

Ichoua S, Gendreau M, Potvin J-Y (2006) Exploiting knowledge about future demands for real-time vehicle dispatching. Transp Sci 40:211–225

Jaillet P (1988) A priori solution of a traveling salesman problem in which a random subset of the customers are visited. Operat Res 36:929–936

Jaillet P, Melvyn Sim JQ (2013) Routing Optimization with Deadlines under Uncertainty. "Tech. Rep". MIT

Kenyon AS, Morton DP (2003) Stochastic vehicle routing with random travel times. Transp Sci 37:69–82

Lecluyse C, van Woensel T, Peremans H (2009) Vehicle routing with stochastic time-dependent travel times. 4OR 7:363–377

Lee C, Lee K, Park S (2012) Robust vehicle routing problem with deadlines and travel time/demand uncertainty. J Operat Res Soc 63:1294–1306

Leipala T (1978) On the solutions of stochastic traveling salesman problems. EurJOperat Res 2:291–297

Maggioni F, Kaut M, Bertazzi L (2009) Stochastic optimization models for a single-sink transportation problem.Comput Manag Sci 6:251–267

Maggioni F, Wallace S (2012) Analyzing the quality of the expected value solution in stochastic programming. Annals Operat Res 200:37–54

Perboli G, Tadei R, Baldi MM (2012) The stochastic generalized bin packing problem. Discrete Appl Math 160:1291–1297

Perboli G, Tadei R, Vigo D (2011) The two-echelon capacitated vehicle routing problem: Models and math-based heuristics.Transp Sci 45:364–380

Pillac V, Gendreau M, Guéret C, Medaglia AL (2013) A review of dynamic vehicle routing problems. Eur J Operat Res 225:1–11

Psaraftis HN, Tsitsiklis JN (1993) Dynamic shortest paths in acyclic networks with markovian arc costs. Operat Res 41:91–101

Reinelt G (1991) TSPLIB- a traveling salesman problem library. ORSA J Comput 3:376–384

Tadei R, Perboli G, Ricciardi N, Baldi MM (2012) The capacitated transshipment location problem with stochastic handling utilities at the facilities. Int Trans Operat Res 19:789–807

Taş D, Dellaert N, Van Woensel T, De Kok T (2013) Vehicle routing problem with stochastic travel times including soft time windows and service costs. Comput Biomed Res Operat Res 40:214–224

Toriello A, Haskell WB,Poremba M (2012) A Dynamic Traveling Salesman Problem with Stochastic Arc Costs. Technical Report Optimization Online

Weisbrod G, Vary D, Treyz G (2001) Economic implications of congestion. Project A2–21 FY'97