CIRRELT

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

Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation

The Capacitated Transshipment Location Problem under Uncertainty

Roberto Tadei
Guido Perboli
Nicoletta Ricciardi

September 2009

CIRRELT-2009-38

Bureaux de Montréal :

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

Bureaux de Québec :

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

www.cirrelt.ca

The Capacitated Transshipment Location Problem under Uncertainty

Roberto Tadei[1], Guido Perboli[1,2,*], Nicoletta Ricciardi[3]

[1] Department of Control and Computer Engineering, Politecnico di Torino, Corso Duca degli Abruzzi, 24 - I-10129 Torino, Italy

[2] Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation (CIRRELT)

[3] Dipartimento di Statistica, Probabilità e Statistiche Applicate, Università di Roma “La Sapienza”, Piazzale Aldo More, 5 – 00185 Roma, Italy

Abstract. We consider the stochastic capacitated transshipment problem for freight transportation where an optimal location of the transshipment facilities, which minimizes the total cost, must be found. The total cost is given by the sum of the total fixed cost plus the expected minimum total flow cost, when the throughput costs of the facilities are random variables with unknown probability distribution. By applying the asymptotic approximation method derived from the extreme value theory, a deterministic nonlinear model, which belongs to a wide class of Entropy maximizing models, is then obtained. The computational results show a very good performance of this deterministic model when compared to the stochastic one.

Keywords.Transshipment location, random throughput costs, asymptotic approximations, Logit model.

Acknowledgements. This project has been partially supported by the Italian Ministry of University and Research, under the 2007 PRIN project “Optimization of Distribution Logistics”. Funding has also been provided by the Natural Sciences and Engineering Research Council of Canada (NSERC), through its Industrial Research Chair and Discovery Grants programs, and by the contributions of the partners of the Chair.

---

* Corresponding author: Guido.Perboli@cirrelt.ca

Dépôt légal – Bibliothèque nationale du Québec,
Bibliothèque nationale du Canada, 2009

1 Introduction

In this paper we consider the capacitated transshipment problem for freight transportation where an optimal location of the transshipment facilities must be found. In any transshipment problem the transportation process takes place in two stages: from origins to transshipment facilities and from transshipment facilities to final destinations. From an economic point of view this process implies three kinds of cost: the fixed cost of locating a transshipment facility, the transportation cost from an origin to a destination through a transshipment facility and the throughput operation cost at each transshipment facility. The throughput operation cost is due to freight treatment operations, such as loading/unloading, but also to inventories and postponed processing, such as packaging, testing etc. [15]. While the fixed costs and the transportation costs are usually well defined, deterministic and quite easy to be measured, the throughput operation costs are ill defined, stochastic and non easily measurable, so that their probability distribution remains in general unknown. Nevertheless, the costs of different throughput operation scenarios inside a transshipment facility play an important role in the choice of the transshipment facility itself by the freight suppliers, and this role becomes more and more economically relevant when moving from local to regional transshipment facilities.

The most significant contribution of this paper to the existing literature is to explicitly consider such stochastic throughput operation costs in the capacitated transshipment problem, where we want to find an optimal facility location while minimizing the total cost, given by the sum of the total fixed cost plus the expected total flow cost, subject to supply and demand satisfaction and to facility capacity constraints.

In a recent paper dedicated to the memory of Charles ReVelle [22], ReVelle et al. identify among new fields in location theory the following one:

- “models ranging from gravity types to Logit functions (which appear to be most promising in the context of location modeling)”.

This paper addresses this field. In fact, we will prove that under a very mild assumption the probability distribution of the minimum cost becomes a Gumbel (or double exponential) distribution, and the expected optimal flows turn to be multinomial Logit functions.

From the above results, a deterministic mixed-integer nonlinear model for the Capacitated Transshipment Location Problem under Uncertainty, which belongs to a wide class of Entropy maximizing models, is then derived.

The stochastic model under different probability distributions and its deterministic approximation are solved for different instances showing a mean gap between the two optima around 2%.

The remainder of the paper is organized as follows. Following a brief overview of some key papers in Section 2, we introduce the Capacitated Transshipment Location Problem under Uncertainty as a two-stage stochastic program with recourse in Section 3. Firstly, we consider only the allocation sub-problem, called Stochastic Flow Problem, and in Section 4 this problem is solved. Thanks to an equivalence between the Stochastic Flow Problem and an Entropy Maximizing Problem derived in Section 5, a deterministic approximation of the Capacitated Transshipment Location Problem under Uncertainty is then given in Section 6. In Section 7 the performance of this approximation is computationally tested. Finally, the conclusions of our work are reported in Section 8.

2 Literature review

In the huge literature on the transshipment facility location problem there are just a few papers were stochasticity is considered, but this stochasticity concerns mainly the demand, while random costs are generally ignored. Given the limited literature on costs stochasticity for the the transshipment facility location problem, in the following we give a short review over a broader set of related problems.

In [16] Klose and Drexl present a review of some contributions to the current state-of-the-art on facility location problems. In particular, models and applications for continuous location, network location, and mixed-integer programming are considered. Also probabilistic models are presented were some of the input data of the location models are subject to uncertainty. Also Ozdemir et al. give a literature review on transshipment theory, but limited to deterministic cases [21]. A more recent review covering stochastic and some non-linear facility location problems is due to Snyder [26]. In particular, this survey considers some classical problems, including the p-median problem.

Glockner and Nemhauser [10] consider a dynamic network flow problem where arc capacities are random variables and derive a multistage stochastic linear program. In [21] and [20] the different cases of transshipment capacity are modeled as a capacitated network flow problem embedded in a stochastic optimisation problem. In these papers the demand is random and with known probability distribution. In [24] the authors analyze a stochastic fractional transshipment problem with uncertain demands and prohibited routes, which is solved reformulating the stochastic transshipment problem into an equivalent deterministic transportation problem. In [29] a two-stage linear program with recourse formulation is developed to determine the optimal storage capacity to be installed on transshipment nodes by shippers in a dynamic shipper carrier network under stochastic demand. In the first stage, the shipper decides the optimal capacity to be installed on transshipment nodes. In the second stage, the shipper chooses a routing strategy based on the realized demand. Also in [32] the authors formulate the multi-location transshipment as a two-stage stochastic program with recourse, where the demand is stochastic. An interesting the paper is [31] where the authors investigate the location problem of the logistics distribution centers under the condition that setup costs, turnover costs and customer demands are fuzzy variables. As a result, a fuzzy chance-constrained programming model is developed. In [33] a capacitated location-allocation problem with stochastic demands is originally formulated as expected value model, chance-constrained programming and dependent-chance programming according to different criteria. In practice, the authors focus on a stochastic Location-Allocation problem where the demands are random.

About stochastic costs, the literature is very limited. Ricciardi et al. [23] consider a p-median problem with random throughput costs at the facilities and develop a non-linear deterministic approximation model which is then solved heuristically. Daskin et al. [5] introduce a location-inventory model that minimizes the expected cost of locating facilities, transporting material, and holding inventory under stochastic daily demand. The stochastic model is then reduced to a deterministic model whose objective function is a function of the means and variances of the random parameters. Snyder et al. [27] consider a scenario-based stochastic version of the joint location-inventory model of Daskin et al. [5], allowing demand means and variances to be stochastic, as well as costs, lead times, and some other parameters. One of the most recent papers on stochastic location is due to Tadei et al. [28]. In this paper the authors consider a stochastic p-median problem where the cost for using a facility is a stochastic variable with unknown probability distribution. Under mild hypotheses, the authors are able to approximate this stochastic problem to a non-linear deterministic model.

3 Problem definition

Let be:

- I: set of origins
- J: set of destinations
- K: set of potential transshipment locations
- L_k: set of throughput operation scenarios at transshipment facility $k \in K$
- P_i: supply at origin $i \in I$
- Q_j: demand at destination $j \in J$
- U_k: throughput capacity of transshipment facility $k \in K$
- f_k: fixed cost of locating a transshipment facility $k \in K$
- y_k: binary variable which takes value 1 if transshipment facility $k \in K$ is located, 0 otherwise
- c_{ij}^k: unit transportation cost from origin $i \in I$ to destination $j \in J$ through transshipment facility $k \in K$
- θ_{kl}: unit throughput cost of transshipment facility $k \in K$ in throughput operation

scenario $l \in L_k$

- s_{ij}^k: flow from origin $i \in I$ to destination $j \in J$ through transshipment facility $k \in K$.

Let us assume

i) the system is balanced, i.e. $\sum_{i \in I} P_i = \sum_{j \in J} Q_j = T$

ii) the unit throughput costs θ_{kl} are independent and identically distributed (i.i.d.) random variables with a common and unknown probability distribution

$$Pr\{\theta_{kl} \geq x\} = F(x). \tag{1}$$

Assumption i) is a standard one and it is straightforward to balance the system if necessary. Assumption ii) is justified by the fact that the unit throughput costs usually vary among transshipment facilities and inside each of them in a random way and are quite difficult to be measured. Thus they become random variables with unknown probability distribution. Moreover, these random variables are independent each other and there is no reason to consider different shapes for their unknown probability distributions ([17], [18], [19]).

Let $r_{ij}^{kl}(\theta)$ be the stochastic generalized unit transportation cost from origin i to destination j through transshipment facility k in throughput operation scenario l given by

$$r_{ij}^{kl}(\theta) = c_{ij}^k + \theta_{kl}, \quad i \in I, j \in J, k \in K, l \in L_k \tag{2}$$

with unknown probability distribution

$$\Pr\left\{r_{ij}^{kl}(\theta) \geq x\right\} = \Pr\left\{c_{ij}^k + \theta_{kl} \geq x\right\} = \Pr\left\{\theta_{kl} \geq x - c_{ij}^k\right\} = F(x - c_{ij}^k). \tag{3}$$

Let us define

$$\overline{\theta}_k = \min_{l \in L_k} \theta_{kl}, \quad k \in K \tag{4}$$

with unknown probability distribution

$$H(x) = \Pr\left\{\overline{\theta}_k \geq x\right\} \tag{5}$$

As $\overline{\theta}_k \geq x \Longleftrightarrow \theta_{kl} \geq x,\ l \in L_k$ and θ_{kl} are independent, using (1) one gets

$$H(x) = \Pr\left\{\overline{\theta}_k \geq x\right\} = \prod_{l \in L_k} \Pr\left\{\theta_{kl} \geq x\right\} = \prod_{l \in L_k} F(x) = [F(x)]^{n_k} \tag{6}$$

where $n_k = |L_k|$ is the number of the different throughput operation scenarios at the transshipment facility k.

The stochastic generalized unit transportation cost from origin i to destination j through transshipment facility k is the minimum among the costs for the different throughput operation scenarios at facility k and becomes

$$\overline{r}_{ij}^{k}(\theta) = \min_{l \in L_k} r_{ij}^{kl}(\theta) = c_{ij}^{k} + \min_{l \in L_k} \theta_{kl} = c_{ij}^{k} + \overline{\theta}_k, \quad i \in I, j \in J, k \in K. \tag{7}$$

In order to write our stochastic problem as a two-stage program with fixed recourse [1], let us consider the variables y_k and s_{ij}^k as the first-stage decision variables and introduce the second-stage decision variables $x_{ij}^k(\theta)$, such that $x_{ij}^k(\theta) = s_{ij}^k, \ \forall \theta_{kl}, \ k \in K, l \in L_k$.

The $CTLP_u$ may be formulated as follows

$$\min_{y} \sum_{k \in K} f_k y_k + \mathbb{E}_\theta \left[\min_{s} \sum_{i \in I} \sum_{j \in J} \sum_{k \in K} \overline{r}_{ij}^{k}(\theta) s_{ij}^{k} \right] \tag{8}$$

subject to

$$\sum_{j \in J} \sum_{k \in K} s_{ij}^{k} = P_i, \quad i \in I \tag{9}$$

$$\sum_{i \in I} \sum_{k \in K} s_{ij}^{k} = Q_j, \quad j \in J \tag{10}$$

$$\sum_{i \in I} \sum_{j \in J} s_{ij}^{k} \leq U_k y_k, \quad k \in K \tag{11}$$

$$x_{ij}^{k}(\theta) = s_{ij}^{k}, \ \forall \theta_{kl}, \ i \in I, j \in J, k \in K, l \in L_k \tag{12}$$

$$x_{ij}^{k}(\theta) \geq 0, \ \forall \theta_{kl}, \ i \in I, j \in J, k \in K, l \in L_k \tag{13}$$

$$s_{ij}^{k} \geq 0, \quad i \in I, j \in J, k \in K \tag{14}$$

$$y_k \in \{0, 1\}, \quad k \in K \tag{15}$$

where $\mathbb{E}_\theta$ denotes the expected value with respect to θ; the objective function (8) expresses the minimization of the total cost given by the sum of the minimum total fixed cost plus the expected minimum total flow cost; constraints (9) and (10) ensure that supply at each origin i and demand at each destination j are satisfied; constraints (11) ensure the capacity restriction at each transshipment facility k; constraints (12) tie the first-stage decision variables s_{ij}^k (revealed flows) to the second-stage decision variables $x_{ij}^k(\theta)$ for any occurrence of the random variables θ_{kl}; (13), (14) are the non-negativity constraints, and (15) are the integrality constraints.

4 Solving the Stochastic Flow Problem

Let us consider first the Stochastic Flow Problem obtained from problem (8)-(15) by disregarding the demand satisfaction constraints (10) and the capacity constraints (11), and assuming that a location of the transshipment facilities $\{y_k\}$ is already known. Flows s_{ij}^k remain the only unknowns and the Stochastic Flow Problem becomes

$$\mathbb{E}_\theta \left[\min_s \sum_{i \in I} \sum_{j \in J} \sum_{k \in K/y_k=1} \overline{r}_{ij}^k(\theta) s_{ij}^k \right] \tag{16}$$

subject to

$$\sum_{j \in J} \sum_{k \in K/y_k=1} s_{ij}^k = P_i, \quad i \in I \tag{17}$$

$$x_{ij}^k(\theta) = s_{ij}^k, \ \forall \theta_{kl}, \ i \in I, j \in J, k \in K/y_k = 1, l \in L_k \tag{18}$$

$$x_{ij}^k(\theta) \geq 0, \ \forall \theta_{kl}, \ i \in I, j \in J, k \in K/y_k = 1, l \in L_k \tag{19}$$

$$s_{ij}^k \geq 0, \quad i \in I, j \in J, k \in K/y_k = 1. \tag{20}$$

Let define $\overline{r}_i(\theta)$ as the minimum stochastic generalized unit transportation cost among the different alternatives (j,k) for a supplier i. Due to (7), this cost is given by

$$\overline{r}_i(\theta) = \min_{j,k/y_k=1} \left\{ \overline{r}_{ij}^k(\theta) \right\} \quad = \quad \min_{j,k/y_k=1} \left\{ c_{ij}^k + \overline{\theta}_k \right\}, \quad i \in I \tag{21}$$

with probability distribution

$$G(x) = \Pr\left\{ \overline{r}_i(\theta) \geq x \right\} = \Pr\left\{ \min_{j,k/y_k=1} \left\{ c_{ij}^k + \overline{\theta}_k \right\} \geq x \right\}. \tag{22}$$

As

$$\min_{j,k/y_k=1} \left\{ c_{ij}^k + \overline{\theta}_k \right\} \geq x \Longleftrightarrow \left\{ c_{ij}^k + \overline{\theta}_k \right\} \geq x, \quad j \in J, k \in K/y_k = 1 \tag{23}$$

and the random variables $\overline{\theta}_k$ are independent (because θ_{kl} are independent), using (5) and (6) the probability distribution (22) becomes

$$\begin{aligned} G(x) \quad &= \prod_{j \in J} \prod_{k \in K/y_k=1} \Pr\left\{ c_{ij}^k + \overline{\theta}_k \geq x \right\} = \prod_{j \in J} \prod_{k \in K/y_k=1} \Pr\left\{ \overline{\theta}_k \geq x - c_{ij}^k \right\} = \\ &\prod_{j \in J} \prod_{k \in K/y_k=1} H(x - c_{ij}^k) = \prod_{j \in J} \prod_{k \in K/y_k=1} \left[F(x - c_{ij}^k) \right]^{n_k}. \end{aligned} \tag{24}$$

The fact that the probability distribution $F(x)$ is unknown prevents the use of the probability distribution given by (24). A possible way to overcome this problem and get an explicit form for $G(x)$ is to consider its asymptotic approximation. A similar approach has been adopted by the authors for a different problem in [28].

4.1 The asymptotic approximation for the probability distribution $G(x)$ of the minimum stochastic generalized unit transportation cost

The asymptotic approximation method to derive $G(x)$ is based on the following observation. If under mild conditions on the probability distribution $F(x)$ of the random unit throughput costs θ_{kl}, the distribution of the stochastic variables $r_{ij}^{kl}(\theta)$ (and then of their minimum $\bar{r}_i(\theta)$) tends to a specific functional form as the number n_k of the different throughput operation scenarios at any transshipment facility k becomes large, we do not need further specific knowledge of the probability distribution $F(x)$.

Galambos [9] gives a sufficient condition on $F(x)$ to guarantee the existence of sequences $a_n, b_n > 0$ of constants such that the following limit

$$\lim_{n\to\infty} G(x \mid n) = \lim_{n\to\infty} G(b_n x + a_n) = G(x) \tag{25}$$

does exist for all continuity points of $G(x)$, where $G(x)$ is a nondegenerate distribution function. In other terms, the distribution function $G(x \mid n)$ *weakly converges* (we remind that $G(x \mid n)$ is said to weakly converge if, as $n \to \infty$, $\lim G(x \mid n) = G(x)$ exists for all continuity points x of the limit $G(x)$).

The Galambos' sufficient condition requires that the probability distribution $F(x)$ is asymptotically exponential in its left tail, i.e. there is a constant $\beta > 0$ such that

$$\lim_{y\to-\infty} \frac{1 - F(x+y)}{1 - F(y)} = e^{\beta x}. \tag{26}$$

We then assume condition (26) and prove that $G(x)$ assumes a specific form as the number n_k of the different throughput operation scenarios at any transshipment facility k becomes large (condition that can be easily verified in our case).

Consider first the following aspect: the solution of problem (16)-(20) does not change if an arbitrary constant is added to the random variables θ_{kl}.

Let us choose this constant as the root a_{n_k} of the equation

$$1 - F(a_{n_k}) = 1/n_k. \tag{27}$$

Replacing $\overline{\theta}_k$ with $\overline{\theta}_k - a_{n_k}$ in (24)

$$G(x \mid n_k) = \prod_{j \in J} \prod_{k \in K/y_k=1} \left[F(x - c_{ij}^k + a_{n_k})\right]^{n_k}. \tag{28}$$

Let us assume that $n_k, k \in K/y_k = 1$ are large enough to use $\lim_{n_k \to \infty} G(x \mid n_k)$ as an approximation of $G(x)$.

The following property holds

Property 1 *Under condition (26), the unknown probability distribution $G(x)$ becomes*

$$G(x) = \lim_{n_k \to \infty} G(x \mid n_k) = \exp\left(-A_i e^{\beta x}\right) \tag{29}$$

where

$$A_i = \sum_{j \in J} \sum_{k \in K/y_k=1} e^{-\beta c_{ij}^k}, \quad i \in I \tag{30}$$

is the accessibility, in the sense of Hansen [12], of a supplier in i to the overall system of located transshipment facilities and destinations.

Proof. By (28) one has

$$\begin{aligned} G(x) &= \lim_{n_k \to \infty} G(x \mid n_k) = \lim_{n_k \to \infty} \prod_{j \in J} \prod_{k \in K/y_k=1} \left[F(x - c_{ij}^k + a_{n_k})\right]^{n_k} = \\ &= \prod_{j \in J} \prod_{k \in K/y_k=1} \lim_{n_k \to \infty} \left[F(x - c_{ij}^k + a_{n_k})\right]^{n_k}. \end{aligned} \tag{31}$$

As $\lim_{n_k \to \infty} a_{n_k} = -\infty$, from (26) one obtains

$$\lim_{n_k \to \infty} \frac{1 - F(x - c_{ij}^k + a_{n_k})}{1 - F(a_{n_k})} = e^{\beta(x - c_{ij}^k)}. \tag{32}$$

By (32) and (27) one has

$$\lim_{n_k \to \infty} F(x - c_{ij}^k + a_{n_k}) = \lim_{n_k \to \infty} \left(1 - [1 - F(a_{n_k})] e^{\beta(x - c_{ij}^k)}\right) = \lim_{n_k \to \infty} \left(1 - \frac{e^{\beta(x - c_{ij}^k)}}{n_k}\right) \tag{33}$$

and, by reminding that $\lim_{n \to \infty}(1 + \frac{x}{n})^n = e^x$

$$\lim_{n_k \to \infty} \left[F(x - c_{ij}^k + a_{n_k})\right]^{n_k} = \lim_{n_k \to \infty} \left(1 - \frac{e^{\beta(x - c_{ij}^k)}}{n_k}\right)^{n_k} = \exp\left(-e^{\beta(x - c_{ij}^k)}\right). \tag{34}$$

Substituting (34) into (31) and using (30) one finally gets

$$G(x) = \prod_{j \in J} \prod_{k \in K / y_k = 1} \exp\left(-e^{\beta(x - c_{ij}^k)}\right) = \exp\left(-A_i e^{\beta x}\right). \quad \square \tag{35}$$

■

The only hypothesis made by Property 1 is that $F(x)$ is an exponential function in its left tail. This is a very mild hypothesis, in fact we observe that many probability distributions show such a behavior, among them the following widely used distributions: Gamma, Gumbel, Laplace, and Logistic.

It is interesting to observe that $G(x)$ in (29) becomes a Gumbel (or double exponential) distribution [11].

4.2 Finding the optimum of the Stochastic Flow Problem

Using the probability distribution $G(x)$ given by (29) we are now able to calculate the expected value of the minimum stochastic generalized unit transportation cost for a supplier in i as follows

$$\hat{r}_i = \mathbb{E}_\theta\left[\bar{r}_i(\theta)\right] = -\int_{-\infty}^{+\infty} x dG(x) = \int_{-\infty}^{+\infty} x \exp\left(-A_i e^{\beta x}\right) A_i e^{\beta x} \beta dx, \quad i \in I. \tag{36}$$

Substituting for $t = A_i e^{\beta x}$ one gets

$$\begin{aligned} \hat{r}_i &= 1/\beta \int_0^{+\infty} \ln(t/A_i) e^{-t} dt = \\ &= 1/\beta \int_0^{+\infty} e^{-t} \ln t dt - 1/\beta \ln A_i \int_0^{+\infty} e^{-t} dt = \\ &= -\gamma/\beta - 1/\beta \ln A_i = \\ &= -1/\beta(\ln A_i + \gamma) \end{aligned} \tag{37}$$

where $\gamma = -\int_0^{+\infty} e^{-t} \ln t \; dt \simeq 0.5772$ is the Euler constant.

It is clear that the optimum of the Stochastic Flow Problem (16)-(20), due to (37), is given by

$$\sum_{i \in I} P_i \hat{r}_i = -\frac{1}{\beta} \sum_{i \in I} P_i (\ln A_i + \gamma) \tag{38}$$

which can also be interpreted as an additive expected utility model [19], where the term $\frac{1}{\beta}(\ln A_i + \gamma)$ is the expected utility for a supplier in i.

If one defines the *total accessibility* of the overall transshipment system as

$$\Phi = \prod_{i \in I} A_i^{P_i} \tag{39}$$

one easily derives from (38)

$$\begin{aligned} \sum_{i \in I} P_i \hat{r}_i = -\frac{1}{\beta} \sum_{i \in I} P_i \ln A_i - \frac{\gamma}{\beta} \sum_{i \in I} P_i = -\frac{1}{\beta} \sum_{i \in I} \ln A_i^{P_i} - \frac{\gamma}{\beta} T = \\ = -\frac{1}{\beta} \ln \prod_{i \in I} A_i^{P_i} - \frac{\gamma}{\beta} T = -\frac{1}{\beta} \ln \Phi - \frac{\gamma}{\beta} T. \end{aligned} \tag{40}$$

Then we get a very interesting result: the optimum of the Stochastic Flow Problem (16)-(20) (i.e. the expected minimum total cost) is proportional (but the constant $-\frac{\gamma}{\beta}T$) to the opposite of the logarithm of the total accessibility Φ.

We are now interested to calculate the optimal flows s_{ij}^k of the Stochastic Flow Problem, which is done in the next Section.

4.3 Finding the optimal flows of the Stochastic Flow Problem

Let define p_{ij}^k as the probability for a supplier i to choose the alternative (j, k) for freight shipping.

The following property holds

Property 2 *At optimality, the probability p_{ij}^k is given by*

$$p_{ij}^k = \frac{e^{-\beta c_{ij}^k}}{\sum_{j \in J} \sum_{k \in K / y_k = 1} e^{-\beta c_{ij}^k}}, \quad i \in I, j \in J, k \in K / y_k = 1. \tag{41}$$

Proof. At optimality, the probability that a supplier i chooses the alternative (j, k) is equal to the probability that such alternative is that of minimum cost. Then, from the

Total Probability Theorem [6], condition (26) and eq. (30), one obtains

$$\begin{aligned} p_{ij}^k &= \int_{-\infty}^{+\infty} Pr\left\{x < c_{ij}^k \leq x + dx\right\} Pr\left\{c_{ir}^s > x,\ \forall (r,s) \neq (j,k)\right\} = \\ &= \int_{-\infty}^{+\infty} \beta e^{\beta(x - c_{ij}^k)} exp(-A_i e^{\beta x}) dx = \\ &= e^{-\beta c_{ij}^k} \int_0^{+\infty} e^{-A_i t} dt = \frac{e^{-\beta c_{ij}^k}}{A_i} = \\ &= \frac{e^{-\beta c_{ij}^k}}{\sum_{j \in J} \sum_{k \in K / y_k = 1} e^{-\beta c_{ij}^k}} \quad i \in I, j \in J, k \in K / y_k = 1 \end{aligned} \tag{42}$$

where $t = e^{\beta x}$. $\square$

■

The optimal flows s_{ij}^k then become

$$s_{ij}^k = P_i p_{ij}^k = P_i \frac{e^{-\beta c_{ij}^k}}{\sum_{j \in J} \sum_{k \in K / y_k = 1} e^{-\beta c_{ij}^k}}, \quad i \in I, j \in J, k \in K / y_k = 1 \tag{43}$$

and it is trivial to check the satisfaction of constraints (17).

The formulation of s_{ij}^k in (43) represents a multinomial Logit model [7], which is widely used in choice theory. In our case it describes how the freight delivered by a supplier i is split among the different alternatives (j,k), due to the stochasticity of the throughput costs of the transshipment facilities which the freight passes through.

We need now one last step to eventually obtain a deterministic model for the Capacitated Transshipment Location Problem under Uncertainty. This step consists in proving that the Stochastic Flow Problem (16)-(20) is equivalent to a well known Entropy Maximizing Problem.

5 Equivalence between the Stochastic Flow Problem and a Entropy Maximizing Problem

In general, given a location of the transshipment facilities $\{y_k\}$ and a corresponding flows $\{s_{ij}^k\}$, the *Entropy* [30] of the flow system is defined as

$$E = -\sum_{i \in I} \sum_{j \in J} \sum_{k \in K / y_k = 1} s_{ij}^k \ln s_{ij}^k \tag{44}$$

and it gives a measure of the uncertainty represented by the discrete probability distribution of the flows [25]. It is clear that, to make inferences which avoid bias on the basis of partial information, one must use that flow probability distribution which has maximum Entropy [13]. So that, the flows are derived as solution of a problem which maximizes the Entropy of the flow system.

We show now that the Stochastic Flow Problem (16)-(20) is equivalent to a Entropy Maximizing Problem, whose objective function expresses the maximization of the Entropy of the flow system plus a term, which is a combination of the opposite of the total deterministic transportation cost multiplied by β plus a constant, and subject to the supply satisfaction constraints at the origins.

The following property holds

Property 3 *The Stochastic Flow Problem (16)-(20) is equivalent to the following Entropy Maximizing Problem*

$$\max_s \left[-\sum_{i \in I} \sum_{j \in J} \sum_{k \in K / y_k = 1} s_{ij}^k \ln s_{ij}^k - \beta \sum_{i \in I} \sum_{j \in J} \sum_{k \in K / y_k = 1} s_{ij}^k (c_{ij}^k - \frac{1}{\beta}) \right] \tag{45}$$

subject to

$$\sum_{j \in J} \sum_{k \in K / y_k = 1} s_{ij}^k = P_i, \quad i \in I \tag{46}$$

$$s_{ij}^k \geq 0, \quad i \in I, j \in J, k \in K / y_k = 1. \tag{47}$$

Proof. Clearly, see (40), the optimum of the Stochastic Flow Problem is equivalent to $-ln\Phi$.

Firstly, we show that $-ln\Phi$ is equivalent to

$$\min_{\nu_i} \left[-\sum_{i \in I} A_i e^{-\nu_i} - \sum_{i \in I} P_i \nu_i \right] \tag{48}$$

where $\nu_i, i \in I$ are real variables.

Then, we prove that (48) is equivalent to problem (45)-(47), and this ends the proof.

1) Equivalence between $-ln\Phi$ and (48).

By imposing the necessary first order conditions for the variables ν_i in (48) one gets

$$\partial \left[-\sum_{i\in I} A_i e^{-\nu_i} - \sum_{i\in I} P_i \nu_i \right] / \partial \nu_i = e^{-\nu_i} A_i - P_i = 0 \tag{49}$$

then

$$e^{-\nu_i} = P_i / A_i \tag{50}$$

and

$$\nu_i = -\ln P_i + \ln A_i. \tag{51}$$

By substituting (50) and (51) into (48) one obtains

$$\begin{aligned} &-\sum_{i\in I} P_i + \sum_{i\in I} P_i \ln P_i - \sum_{i\in I} P_i \ln A_i = \\ &= -\sum_{i\in I} P_i \ln A_i - \sum_{i\in I} P_i (1 - \ln P_i) = \\ &= -\ln \Phi - \sum_{i\in I} P_i (1 - \ln P_i) \end{aligned}$$

which coincides with $-ln\Phi$, but the additive constant $-\sum_{i\in I} P_i(1-\ln P_i)$.

2) Equivalence between (48) and problem (45)-(47).

Let us consider (45) and the constraints (46). By making a Lagrangian relaxation of these constraints by means of the Lagrangian multipliers $\nu_i, \ i \in I$ given by (51) one gets

$$\min_s \left[\sum_{i\in I} \sum_{j\in J} \sum_{k\in K/y_k=1} s_{ij}^k \ln s_{ij}^k + \beta \sum_{i\in I} \sum_{j\in J} \sum_{k\in K/y_k=1} s_{ij}^k (c_{ij}^k - \frac{1}{\beta}) + \sum_{i\in I} \nu_i (\sum_{j\in J} \sum_{k\in K/y_k=1} s_{ij}^k - P_i) \right]. \tag{52}$$

Then, by imposing the necessary first order conditions for the optimal flows s_{ij}^k one gets

$$\begin{aligned}
&\partial \left[\sum_{i\in I} \sum_{j\in J} \sum_{k\in K/y_k=1} s_{ij}^k \ln s_{ij}^k + \beta \sum_{i\in I} \sum_{j\in J} \sum_{k\in K/y_k=1} s_{ij}^k (c_{ij}^k - \frac{1}{\beta}) \right. \\
&\left. + \sum_{i\in I} \nu_i (\sum_{j\in J} \sum_{k\in K/y_k=1} s_{ij}^k - P_i) \right] / \partial s_{ij}^k = \\
&= \ln s_{ij}^k + \beta c_{ij}^k + \nu_i = 0
\end{aligned}$$

and the optimal flows become

$$s_{ij}^k = e^{-\nu_i} e^{-\beta c_{ij}^k}, \quad i \in I, j \in J, k \in K/y_k = 1. \tag{53}$$

Using (50) and (30), it is easy to see that the optimal flows in (53) are equal to the optimal flows of the Stochastic Flow Problem given by (43).

By substituting (53) into (52) one gets

$$\begin{aligned}
&-\sum_{i\in I} e^{-\nu_i} \sum_{j\in J} \sum_{k\in K/y_k=1} e^{-\beta c_{ij}^k} - \sum_{i\in I} Pi\nu_i = \\
&= -\sum_{i\in I} A_i e^{-\nu_i} - \sum_{i\in I} P_i \nu_i
\end{aligned}$$

which is exactly (48). $\square$ $\blacksquare$

We are now able to derive a deterministic approximation for the Capacitated Transshipment Location Problem under Uncertainty, which is presented in the next Section.

6 The deterministic approximation of $CTLP_u$

We remind that problem (45)-(47) is a deterministic approximation of the Stochastic Flow Problem (16)-(20), obtained from the original $CTLP_u$ by disregarding the demand satisfaction constraints at the destinations (10) and the capacity constraints of the transshipment facilities (11), and assuming a given location for the transshipment facilities. Now, by reintroducing those constraints, dropping the assumption of a given optimal location and multiplying the objective function (45) by $1/\beta$, we get the final Deterministic Approximation of the Capacitated Transshipment Location Problem under Uncertainty, named $CTLP_d$, as follows

$$\min_y \sum_{k\in K} f_k y_k + \max_s \left[-\frac{1}{\beta} \sum_{i\in I} \sum_{j\in J} \sum_{k\in K} s_{ij}^k \ln s_{ij}^k - \sum_{i\in I} \sum_{j\in J} \sum_{k\in K} s_{ij}^k (c_{ij}^k - \frac{1}{\beta}) \right] \tag{54}$$

subject to

$$\sum_{j\in J} \sum_{k\in K} s_{ij}^k = P_i, \quad i \in I \tag{55}$$

$$\sum_{i\in I} \sum_{k\in K} s_{ij}^k = Q_j, \quad j \in J \tag{56}$$

$$\sum_{i\in I} \sum_{j\in J} s_{ij}^k \leq U_k y_k, \quad k \in K \tag{57}$$

$$s_{ij}^k \geq 0, \quad i \in I, j \in J, k \in K \tag{58}$$

$$y_k \in \{0,1\}, \quad k \in K \tag{59}$$

which is a mixed-integer deterministic nonlinear model in the unknowns y_k and s_{ij}^k.

We observe that the nonlinearity affects only the objective function through the Entropy term
$-\frac{1}{\beta} \sum_{i\in I} \sum_{j\in J} \sum_{k\in K} s_{ij}^k \ln s_{ij}^k$, while the constraints are all linear.

It is interesting to note that when $\beta \to \infty$ problem (54)-(59) turns into the classical Capacitated Transshipment Location Problem. In fact, the Entropy term in the objective function disappears and only the linear classical total transportation cost does remain. This is also coherent with the well-known property of the multinomial Logit model which states that this model collapses into a classical minimum transportation cost choice model as $\beta \to \infty$ [7].

We also observe that, provided $\max_s -\frac{1}{\beta} \sum_{i\in I} \sum_{j\in J} \sum_{k\in K} s_{ij}^k \ln s_{ij}^k \geq 0$, the optimum of the classical Capacitated Transshipment Location Problem is a Lower Bound for the Capacitated Transshipment Location Problem under Uncertainty.

7 Computational results

In this section we compare the Capacitated Transshipment Location Problem under Uncertainty $CTLP_u$, given by (8)-(15), with its Deterministic Approximation $CTLP_d$, given by (54)-(59).
This section is organized as follows. As no instance for $CTLP_u$ is available in literature, new instances are generated and introduced in Subsection 7.1. The setting of a commercial stochastic solver for solving $CTLP_u$ as well as the identification of an appropriate nonlinear solver for solving $CTLP_d$ are given in Subsection 7.2. A detailed comparison between $CTLP_u$ under different probability distributions and its deterministic approximation $CTLP_d$ is discussed in Subsection 7.3. Finally, in order to show the effect of the Entropy term in $CTLP_d$, in Subsection 7.4 we compare $CTLP_d$ with the classical $CTLP$, and we evaluate the speed of convergence of $CTLP_d$ to the classical $CTLP$, when the value of parameter β does increase.

7.1 Instance generation

We consider a subset of the test classes given in [15], where the authors generate 4 classes with 20 instances in each class. Here, due to the much higher computational effort required to solve the stochastic and the nonlinear problems, we consider only the first class from [15] and generate 10 instances instead of 20, using uniform distribution with corresponding ranges according to the following criteria:

- number of depots $|I|$ is drawn from $U[2, 3]$;
- number of customers $|J|$ is drawn from $U[30, 40]$;
- number of possible locations for the transshipments $|K|$ is drawn from $U[10, 20]$;
- supply P_i is drawn from $U[900, 1000]$;
- demand Q_j is drawn from $U[1, \sum_{i \in I} P_i / \mid J \mid]$. If necessary, the demand of the last customer is adjusted so that the total demand is equal to the total supply;
- capacity U_k is drawn from $U[0.5avU, 3avU]$, where $avU = \sum_{i \in I} P_i / \mid K \mid$;
- unit transportation cost c_{ij}^k is drawn from $U[1, 10]$;
- fixed cost $f_k = TC\ U_k / (\mid I \mid\mid J \mid)$, where TC is the total unit transportation cost over all the possible arcs.
- random cost $\overline{\theta}_k$ is generated using three different probability distributions, Gumbel, Laplace, and Uniform, as follows (the cumulative distribution functions are considered):
  - Gumbel: $exp{-}e^{-\beta x}$ (with mode equal to 0).
    The parameter β is set to 0.1, which ensures to have a mean of the Gumbel distribution ($\simeq 5.7$) quite close to the mean of the distribution used to obtain the deterministic unit costs c_{ij}^k. In this way, the random costs $\overline{\theta}_k$ have the same order of magnitude of the deterministic unit costs c_{ij}^k;

– Laplace:

$$\begin{cases} 0.5\, exp(\frac{x-\mu}{b}) & \text{if } x < \mu \\ 1 - 0.5\, exp(-\frac{x-\mu}{b}) & \text{if } x \geq \mu \end{cases}$$

with mean equal to μ. The parameters of the distribution are set such that the mean of the Laplace distribution is the same of the Gumbel one;

– Uniform

$$\begin{cases} 0 & \text{if } x < a \\ \frac{x-a}{b-a} & \text{if } a \leq x < b \\ 1 & \text{if } x \geq b \end{cases}$$

The costs are generated in the range $[a, b] = [1, 10]$, such that the mean of the Uniform distribution is quite close to the Gumbel one.

The random unit generalized transportation costs $\bar{r}_{ij}^k$ in (8) are computed by (7). If some of them become negative, they are set to 1.

7.2 Stochastic solver setting and nonlinear solver identification

As stated above, we compare $CTLP_u$ with its deterministic approximation $CTLP_d$. The solution of $CTLP_u$ is generated by means of a two-stage implementation of the stochastic model in XPress-SP, the stochastic programming module provided by XPress [8]. The tests are performed by generating an appropriate number of scenarios for each instance. In order to tune this number, we start with 50 scenarios and increase them by step 50. Then we solve each instance 10 times, reinitializing every time the pseudo-random generator of the stochastic components with a different seed, and compute the standard deviation and the mean of the optima over the 10 runs. The appropriate number of scenarios is then fixed to the smallest value ensuring for each instance a maximum ratio between the standard deviation and the mean less than 0.5% [14]. According to our tests, this value is fixed to 100 scenarios, which show a maximum ratio between standard deviation and mean equal to 0.17%.

In order to solve the deterministic approximation $CTLP_d$ we consider the most efficient and effective state-of-the-art nonlinear solvers: BonMIN, MinLP, KNITRO, LINGO and FilMINT. In order to have uniformity in input, output and computational results, we use the NEOS infrastructure [4] to make the tests, giving to the solvers a time limit of 1000 seconds per instance. According to our results, BonMIN and KNITRO outperform the other solvers, obtaining the best solutions on the overall set of instances. By comparing each other these two solvers, BonMIN is 10 times faster than KNITRO, which also shows some memory problems when running large instances (with more than 50000 arcs). For these reasons, we select BonMIN (release 1.1) [2, 3] for solving $CTLP_d$ within a time limit of 1000 seconds. The parameters are set to their default values, which show a satisfactory behavior both in accuracy and computational effort.

7.3 Comparison between $CTLP_u$ under different probability distributions and $CTLP_d$

In the following we analyze the performance of the deterministic approximation $CTLP_d$, by comparing its results with those obtained from $CTLP_u$ under the three different probability distributions. All the tests are done on a Pentium Q6600 2.4GHz Machine with 2 Gb of RAM. The parameter β in the $CTLP_d$ objective function (54) is set to 0.1 as the corresponding β value in the Gumbel distribution for $CTLP_u$.

A comparison between the optimum of the deterministic approximation $CTLP_d$ and that of the stochastic model $CTLP_u$ under the three different probability distributions is presented in Table 1.

The table columns have the following meaning:

- Column 1: instance number;
- Column 2: optimum of $CTLP_d$;
- Columns 3-5: optimum of $CTLP_u$ using Gumbel, Laplace, and Uniform distributions, respectively;
- Columns 6-8: percentage gap between the stochastic optimum using Gumbel, Laplace, and Uniform distributions, respectively, and the deterministic approximation one. The gap is computed as $100(S - D)/S$, where S is the optimum of $CTLP_u$ and D the optimum of $CTLP_d$.

Table 1 reports the results for each instance, as well as their mean in the last row.

The optima of the stochastic problem and its deterministic approximation are quite close together, with a mean gap around 2%. The gap is lower for the Gumbel and Laplace distributions than the Uniform one, which is coherent with the assumption done for the probability distribution used to derive $CTLP_d$ (i.e. it acts as an exponential function in its left tail). About the negative value for some gap values present in Table 1, we remind that when the random cost $\overline{\theta}_k$ is generated and added to the deterministic unit cost in (7), if the resulting $\overline{r}^k_{ij}$ is negative it is set to 1. This implies a slight change in the distribution functions, producing such negative gap values.

The results of Table 1, even if satisfactory, are not sufficient to qualify the performance of $CTLP_d$. In fact, besides the optimum, another important comparison concerns the optimal solution of the two models. This is considered in Tables 2 and 3.

Table 1: Comparison between the optimum of the deterministic approximation $CTLP_d$ and that of the stochastic model $CTLP_u$ under the three different probability distributions

Instances	**Objective function**				**Gap**		
	Det	Stoch					
		Gumbel	Laplace	Uniform	Gumbel	Laplace	Uniform
1	142713	137460	139664	134570	3.82%	2.18%	6.05%
2	209429	207013	209238	202495	1.17%	0.09%	3.42%
3	150860	144510	145031	147152	4.39%	4.02%	2.52%
4	167359	164393	165939	161654	1.80%	0.86%	3.53%
5	157160	151061	152683	148561	4.04%	2.93%	5.79%
6	211108	210291	210567	213969	0.39%	0.26%	-1.34%
7	244105	243214	245251	239280	0.37%	-0.47%	2.02%
8	248086	243645	245213	249019	1.82%	1.17%	-0.37%
9	247005	243887	246621	239930	1.28%	0.16%	2.95%
10	188291	181987	184353	185853	3.46%	2.14%	1.31%
Mean	**196612**	**192746**	**194456**	**192248**	**2.25%**	**1.43%**	**2.93%**

Table 2 compares the optimal solution of $CTLP_d$ with that of $CTLP_u$, in terms of open facilities. The table columns have the following meaning:

- Column 1: instance number;
- Columns 2-5: number of open facilities in $CTLP_d$ and $CTLP_u$ under Gumbel, Laplace, and Uniform distributions, respectively;
- Columns 6, 8, 10: number of open facilities which are equal in the optimal solution of $CTLP_d$ and $CTLP_u$ under the three different distributions;
- Columns 7, 9, 11: percentage of open facilities which are equal in the optimal solution of $CTLP_d$ and $CTLP_u$ under the three different distributions.

Table 2 reports the results for each instance, as well as their mean in the last row.

The main conclusion we can draw from these results is that the optimal solution of the stochastic model under the three different distributions is quite similar to that of its deterministic approximation, since on average 75% of the open facilities are the same. Let us consider instance 1 with Laplace and Uniform distributions, for which the open facilities in the optimal solution are exactly the same of those of the deterministic approximation. Nevertheless, the gap between the two optima is 2.18% and 6.05% (see Table 1), respectively. We should then conclude that this gap is due to a different flow distribution in the two optimal solutions.

In order to verify this conclusion we consider the optimum of $CTLP_u$ when the open facilities are those of the $CTLP_d$ optimal solution and we compare this optimum with the original $CTLP_u$ optimum in Table 3.

Table 2: Comparison of the open facilities in the optimal solutions of the deterministic approximation $CTLP_d$ and the stochastic model $CTLP_u$ under different probability distributions

	Number of open facilities				**Common open facilities**					
Instances	Det	Stoch								
		Gumbel	Laplace	Uniform	Gumbel		Laplace		Uniform	
					Number	%	Number	%	Number	%
1	8	6	8	8	4	67%	8	100%	8	100%
2	9	7	7	8	5	71%	5	71%	6	75%
3	12	10	10	10	8	80%	8	80%	7	70%
4	9	6	7	8	4	67%	3	43%	6	75%
5	9	8	8	8	5	63%	6	75%	6	75%
6	13	12	12	12	9	75%	10	83%	10	83%
7	9	7	7	7	5	71%	3	43%	5	71%
8	9	8	9	8	7	88%	9	100%	6	75%
9	9	8	9	8	6	75%	6	67%	5	63%
10	13	12	11	11	11	92%	9	82%	9	82%
Mean	**10**	**8**	**9**	**9**	**6**	**75%**	**7**	**74%**	**7**	**77%**

Table 3: Performance of the optimal solution of the deterministic approximation $CTLP_d$ when used as optimal solution of the stochastic model $CTLP_u$

Instances	**Gumbel**	**Laplace**	**Uniform**
1	0.30%	0.00%	0.00%
2	0.22%	0.16%	0.36%
3	0.19%	0.47%	0.34%
4	0.73%	0.08%	0.29%
5	1.31%	0.40%	0.77%
6	0.57%	0.81%	0.03%
7	0.36%	0.45%	0.06%
8	0.27%	0.40%	0.79%
9	0.53%	0.26%	0.15%
10	0.38%	0.45%	0.84%
Mean	**0.49%**	**0.35%**	**0.38%**

The table columns have the following meaning:

- Column 1: instance number;
- Column 2-4: percentage gap between the stochastic optimum when the facilities are compulsory opened as in the $CTLP_d$ optimal solution and the stochastic optimum when $CTLP_u$ can decide which facilities must be opened.

According to these results, the gap between the optimum of $CTLP_u$ obtained with given open facilities and the original one is on average less than 0.5 for all the three distributions. This implies that the optimal decisions taken by $CTLP_d$ and $CTLP_u$ in terms of open facilities are equivalent (i.e. they generate almost the same optimum) and that the gap between the $CTLP_d$ optimum and the $CTLP_u$ one when the open facilities are different is mainly due to a different flow distribution in the two optimal solutions.

7.4 Comparison between $CTLP_d$ and the classical $CTLP$

The discussion of the computational results ends by showing the behavior of the Entropy term of $CTLP_d$ in (54).

In Table 4 the contribution given by the Entropy term to the optimum of $CTLP_d$ is presented. The table compares the optimum of $CTLP_d$ with that of the classical $CTLP$, which differs from the former by the Entropy term. The table columns have the following meaning:

- Column 1: instance number;
- Column 2: percentage gap between the optimum of $CTLP_d$ and that of the classical $CTLP$;
- Column 3: number of open facilities which are equal in the two optimal solutions;
- Column 4: percentage of open facilities which are equal in the two optimal solutions.

Table 4 reports the results for each instance, as well as their mean in the last row.

According to the results, even if a large part of the open facilities are common to the two optimal solutions, the gap between the two optima is relevant, showing an important role played by the Entropy term in $CTLP_d$.

We remind (see Section 6) that when $\beta \to \infty$ the coefficient of the Entropy term tends to 0 and $CTLP_d$ turns into the classical $CTLP$.

The last test we perform is devoted to show the speed of convergence of $CTLP_d$ to the classical $CTLP$, while the value of parameter β increases. Figure 1 reports the mean

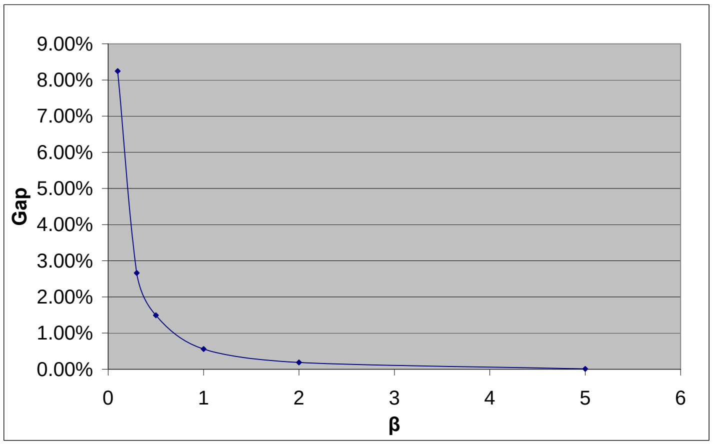

Figure 1: Convergence of the deterministic approximation $CTLP_d$ to the classical $CTLP$ as $\beta \rightarrow \infty$

gap between the optimum of $CTLP_d$ and that of the classical $CTLP$ while β varies. According to Figure 1, the gap is almost zero for β equal to 5, so a very fast convergence is guaranteed.

Table 4: Contribution of the Entropy term in $CTLP_d$

Instances	**Comparison** Gap	Common open facilities	Common open facilities (%)
1	14.68%	5	71%
2	11.89%	6	75%
3	6.48%	8	73%
4	10.59%	5	71%
5	13.66%	6	75%
6	3.85%	10	77%
7	8.93%	9	100%
8	8.87%	6	75%
9	10.24%	5	71%
10	3.89%	11	92%
Mean	**9.31%**	**7**	**78%**

8 Conclusions

In this paper the Capacitated Transshipment Location Problem under Uncertainty $CTLP_u$, which is a two-stage stochastic program with recourse, has been approximated to an equivalent non-linear deterministic Capacitated Transshipment Location Problem $CTLP_d$, which belongs to a wide class of Entropy maximizing models. The performance of $CTLP_d$ is quite good. In fact, the mean gap betwwen the two optima is around 2%.

In particular, when the probability distribution of the random costs in $CTLP_u$ acts as an exponential function in its left tail the performance of $CTLP_d$ is particularly good. This is coherent with the Galambos' sufficient condition of the extreme values theory on which the deterministic approximated problem has been derived.

It is interesting to observe that the optimal transshipment facility location obtained by $CTLP_d$ is equivalent (in terms of optimum) to that of $CTLP_u$ and the remaining small gap is due to a slightly different distribution of the flows from origins to facilities and from those to destinations.

The role of the Entropy term in $CTLP_d$, weighted by the non-negative parameter β set to 0.1, is particularly relevant. Nevertheless, this role is highly affected by the value of the parameter β. In fact, as this parameter does increase the contribution of the Entropy term to the optimum rapidly decreases and for $\beta = 5$ $CTLP_d$ collapses into the classical linear Capacitated Transshipment Location Problem.

Both $CTLP_u$ and $CTLP_d$ have been exactly solved for small size instances (up to 3 origins, 20 potential transshipment locations and 40 destinations) in 1000 seconds by means of existing solvers. Larger instances will probably deserve some heuristic approaches.

9 Acknowledgements

This project has been partially supported by the Italian Ministry of University and Research, under the 2007 PRIN Project "Optimization of Distribution Logistics". The authors thank Dr. Laure Simon for her valuable help in running the software for solving the stochastic model.

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

References

[1] Birge, J.R., Louveaux, F.V.: Introduction to stochastic programming. Springer-Verlag (1997)

[2] Bonami, P., Biegler, L.T., Conn, A.R., Cornuéjols, G., Grossmann, I.E., Laird, C.D.,

Lee, J., Lodi, A., Margot, F., Sawaya, N., Wächter, A.: An algorithmic framework for convex mixed integer nonlinear programs. Discrete Optimization **5**, 186–204 (2008)

[3] Bonami, P., Lee, J.: BONMIN 1.1 Users' Manual. COIN-OR (2009)

[4] Czyzyk, J., Mesnier, M., Moré, J.: The NEOS server. IEEE Journal on Computational Science and Engineering **5**, 68–75 (1998)

[5] Daskin, M.S., Coullard, C.R., Shen, Z.M.: An inventory-location model: formulation, solution algorithm and computational results. Annals of Operations Research **5**, 227–246 (2002)

[6] DeGroot, M., Schervish, M.: Probability and Statistics (Third Edition). Addison-Wesley (2002)

[7] Domencich, T., McFadden, D.: Urban travel dynamics: a behavioral analysis. North Holland (1975)

[8] Dormer, A., Vazacopoulos, A., Verma, N., Tipi, H.: Supply Chain Optimization, chap. Modeling & Solving Stochastic Programming Problems in Supply Chain Management Using Xpress-SP, pp. 307–354. Springer-Verlag (2005)

[9] Galambos, J.: The asymptotic theory of extreme order statistics. John Wiley (1978)

[10] Glockner, G.D., Nemhauser, G.L.: A dynamic network flow problem with uncertain arc capacities: Formulation and problem structure. Operations Research **48**, 233–242 (2000)

[11] Gumbel, E.J.: Statistics of Extremes. Columbia University Press (1958)

[12] Hansen, W.: How accessibility shapes land use. Journal of the American Institute of Planners **25**, 73–76 (1959)

[13] Jaynes, E.: Information theory and statistical mechanics. Phys. Rev. **106**, 620–630 (1957)

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

[15] Keskin, B.B., Uster, H.: A scatter search-based heuristic to locate capacitated transshipment points. Computers and Operations Research **34**, 3112–3125 (2007)

[16] Klose, A., Drexl, A.: Facility location models for distribution system design. European Journal of Operational Research **162**, 4–29 (2005)

[17] Leonardi, G.: The use of random utility theory in building location-allocation models. In: Thisse, J.F., Zoller, H.G. (eds.) Locational analysis of public facilities, pp. 357–383. North Holland (1983)

[18] Leonardi, G.: The structure of random utility models in the light of the asymptotic theory of extremes. In: Florian, M. (ed.) Transportation planning models, pp. 107–133. Elsevier (1984)

[19] Leonardi, G., Tadei, R.: Random utility demand models and service location. Regional Science and Urban Economics **14**, 399–431 (1984)

[20] Ozdemir, D., Yucesan, E., Herer, Y.T.: Multi-location transshipment problem with capacitated production and lost sales. In: Perrone, L.F., Wieland, F.P., Liu, J., Lawson, B.G., Nicol, D.M., Fujimoto, R.M. (eds.) Proceedings of the 2006 Winter Simulation Conference, pp. 1470–1476 (2006)

[21] Ozdemir, D., Yucesan, E., Herer, Y.T.: Multi-location transshipment problem with capacitated transportation. European Journal of Operational Research **175**, 602–621 (2006)

[22] ReVelle, C.S., Eiselt, H.A., Daskin, M.S.: A bibliography for some fundamental problem categories in discrete location science. European Journal of Operational Research **184**, 817–848 (2008)

[23] Ricciardi, N., Tadei, R., Grosso, A.: Optimal facility location with random throughput costs. Computers and Operations Research **29**, 593–607 (2002)

[24] Shakeel, J., Gupta, S.N.: Capacitated stochastic transshipment problem with prohibited routes. Int. J. Agric. Stat. Sci. **4**, 127–142 (2008)

[25] Shannon, C., Weaver, W.: The Mathematical Theory of Communication. University of Illinois Press (1949)

[26] Snyder, L.V.: Facility location under uncertainty: a review. IIE Transactions **18**, 547–564 (2006)

[27] Snyder, L.V., Daskin, M.S., Teo, C.: The stochastic location model with risk pooling. European Journal of Operational Research **179**, 1221–1238 (2007)

[28] Tadei, R., Ricciardi, N., Perboli, G.: The stochastic p-median problem with unknown cost probability distribution. Operations Research Letters **37**, 135–141 (2009)

[29] Unnikrishnan, A., Valsaraj, V., Waller, S.T.: Stochastic and dynamic shipper carrier network design problem. Networks and Spatial Economics **DOI: 10.1007/s11067-008-9094-3** (2009)

[30] Wilson, A.G.: Entropy in Urban and Regional Modeling. Pion (1970)

[31] Yanga, L., Ji, X., Gao, Z., Li, K.: Logistics distribution centers location problem and algorithm under fuzzy environment. Journal of Computational and Applied Mathematics **208**, 303–315 (2007)

[32] Zhao, L., Sen, S.: A comparison of sample-path-based simulation-optimization and stochastic decomposition for multi-location transshipment problems. In: Perrone, L.F., Wieland, F.P., Liu, J., Lawson, B.G., Nicol, D.M., Fujimoto, R.M. (eds.) Proceedings of the 2006 Winter Simulation Conference, pp. 238–245 (2006)

[33] Zhou, J., Liu, B.: New stochastic models for capacited location allocation problem. Computers & Industrial Engineering **45**, 111–125 (2003)